



ARISTOTLE  
UNIVERSITY  
OF THESSALONIKI



**Team:**  
**Eleftherios Spyromitros-  
Xioufis**  
**Emmanouela Stachtari**

# Data Mining Cup 2011



1

Aristotle University of Thessaloniki

# Presentation Outline

- Graph based recommendations
- Model generation
- Making recommendations
- Evaluation & alternatives

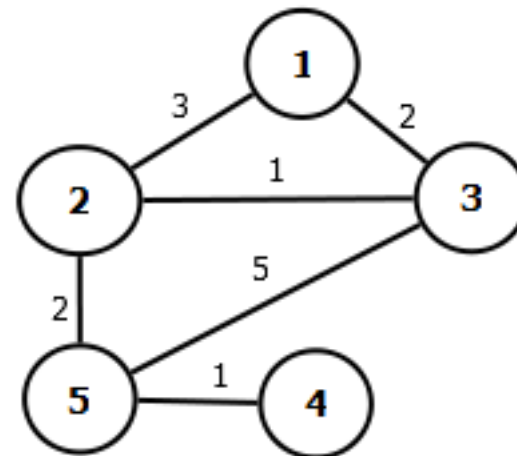
# Graph Based Recommendations

- What we didn't have:
  - Information about users
  - Information about the items' content
- What we had:
  - Transitions from item to item
- Formulation as a graph
  - Items are the nodes
  - Transitions are the edges

# Model Generation

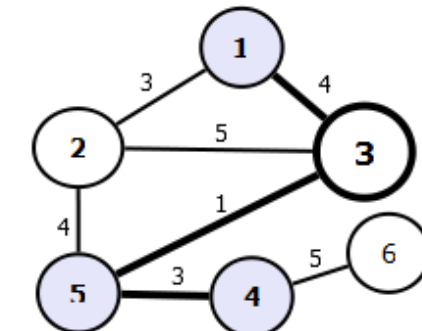
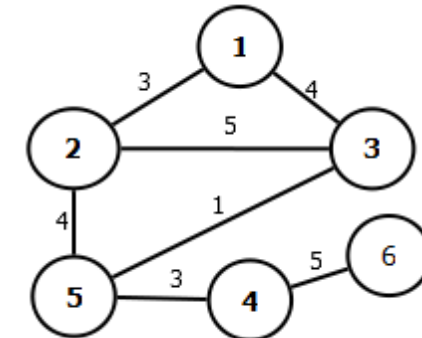
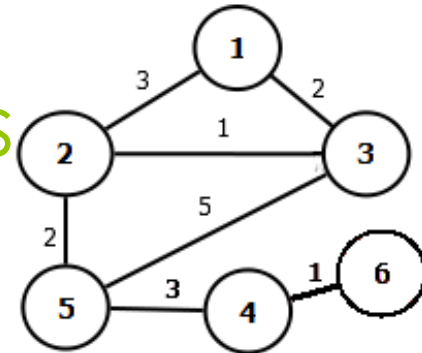
- Both training and test sessions were used to create a similarity graph
- For all sessions:
  - Add the session's items in the graph
  - Add an undirected edge for each item to item transition
  - Increase edge weight by 1 when a transition reappears

Session Id	Item Id	Transaction
1000	4	0
1000	5	0
1000	4	1
1000	6	0



# Making recommendations

- Recommendations are based on Dijkstra's shortest path algorithm:
- Transform weights to distances
  - $maxW = W(x, y) : W(x, y) \geq W(x', y'), \forall x', y' \in Nodes$
  - $dist(x, y) = maxW - W(x, y) + 1$
- Recommendations for a test session are the 3 nearest neighbors of the last item in the session
  - e.g.  $S_{200} = \langle 17, 2, 3 \rangle$   
 $Pred_{200} = \langle 5, 1, 4 \rangle$



# Evaluation & alternatives

- Transact\_train was splitted into:
  - Train (90%) – Test (10%)
- The following alternatives were tested:
  - Using a directed graph
  - Using edges of different weights for different transaction types
  - Recommending the nearest neighbors of all test session's items instead of recommending the nearest neighbors of the last item

# The end

- THANK YOU!
- QUESTIONS OR COMMENTS?