

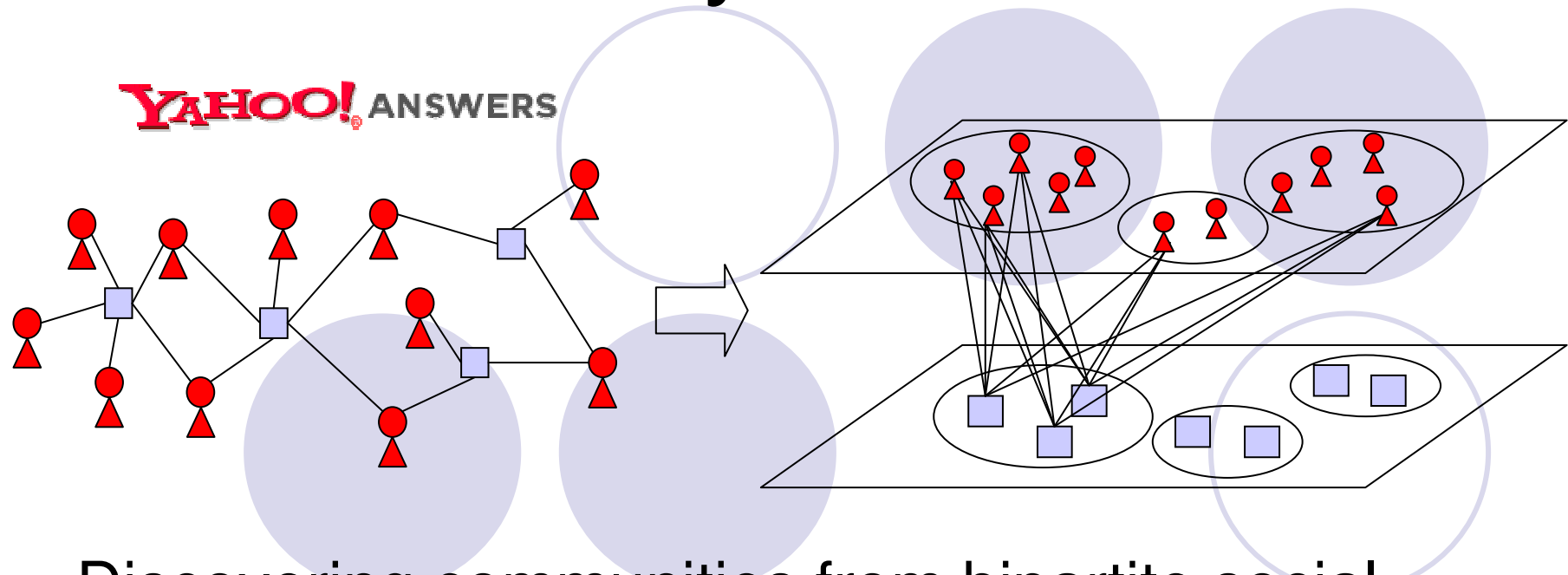
Discovering and Visualizing Communities from Bipartite Social Networks



Tsuyoshi Murata and Tomoyuki Ikeya

Tokyo Institute of Technology,
Japan

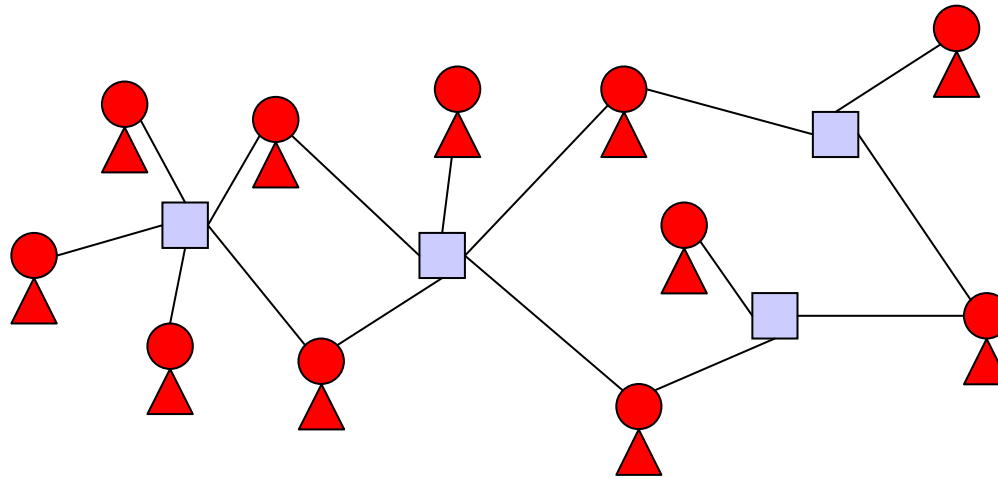
Summary of this talk



- Discovering communities from bipartite social networks
- Measuring correspondence between communities

Social Interactions as Bipartite Networks

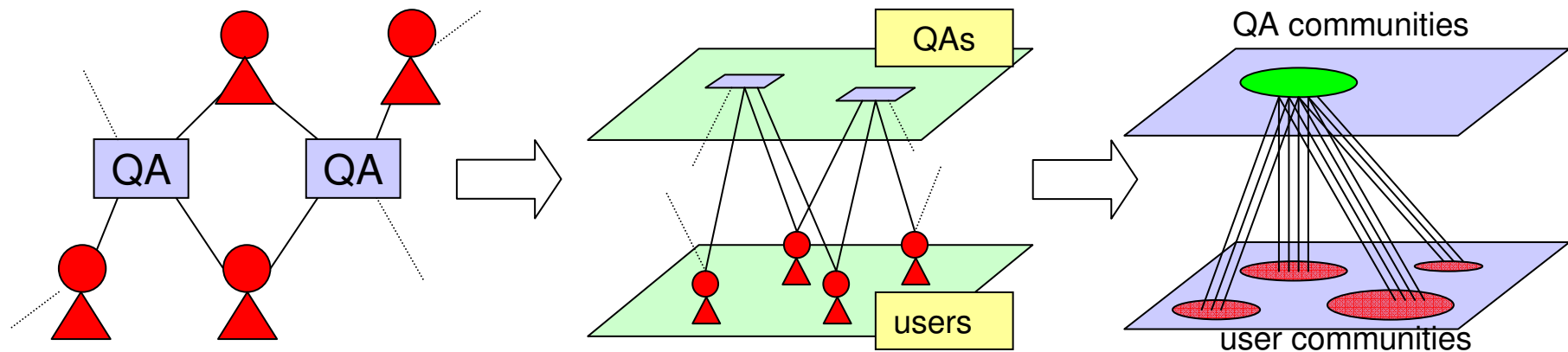
- Real interactions are often represented as bipartite networks
 - actors and movies
 - authors and papers



Goals

Discovering Communities from bipartite Social Networks

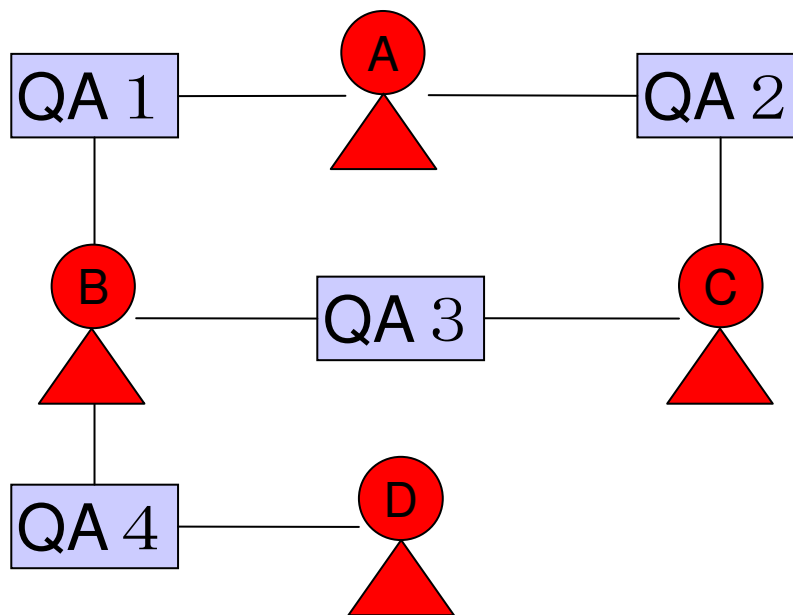
- methods for discovering communities
 - transforming to homogeneous networks
- criteria for evaluating “close-knitness” between communities of different nodes



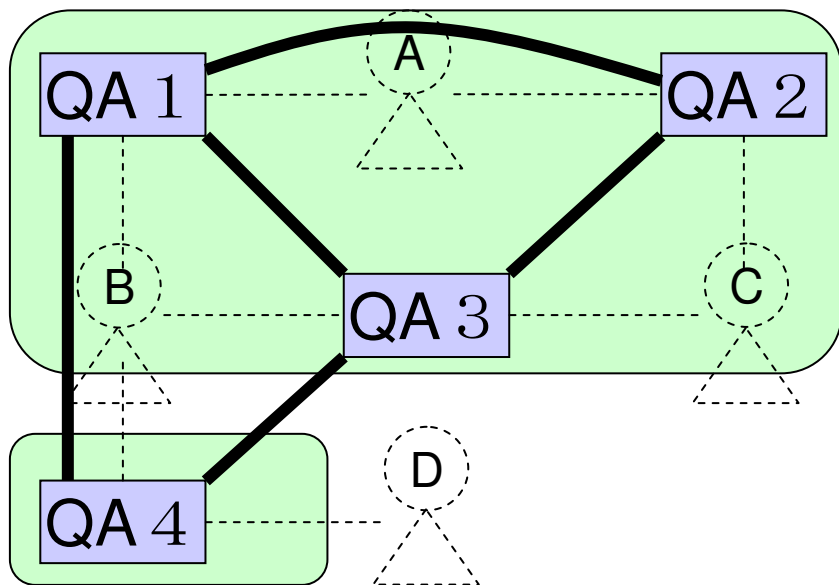
- Transforming networks

- QA network and user network are extracted

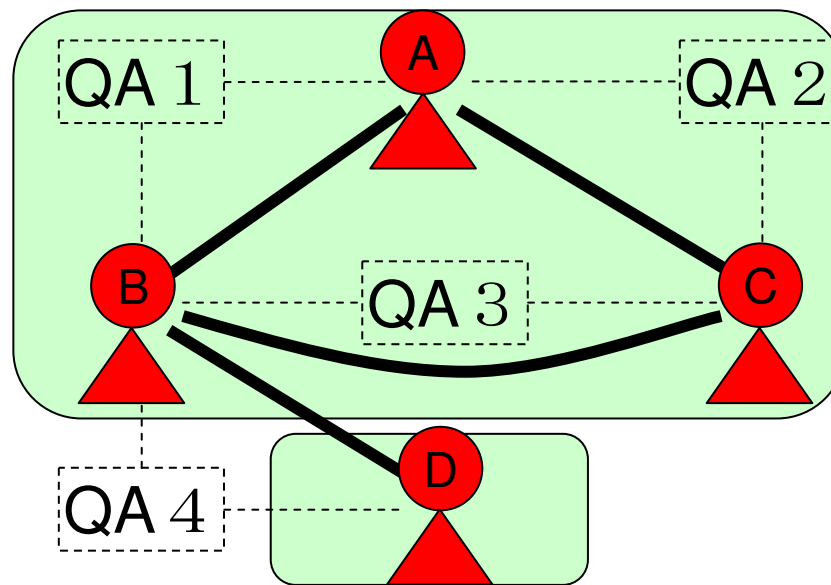
- [Clauset04] is used for extracting communities



QA network



user network





Data of Online Social Interactions

- Japanese Yahoo! Answers

- online questing-answering forum
<http://chiebukuro.yahoo.co.jp>

- QAs posted on September 2005 (One month)

- Nodes: QAs and users

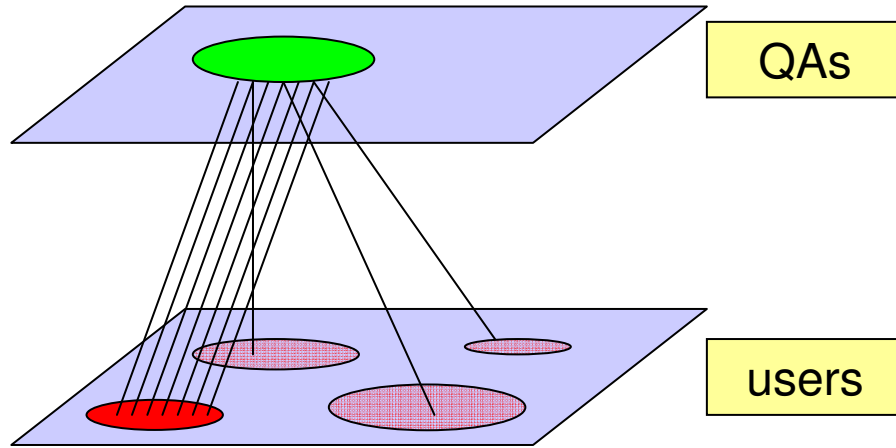
- Edges: between QAs and contributed users

- # of nodes: 246,849

- # of edges: 357,834

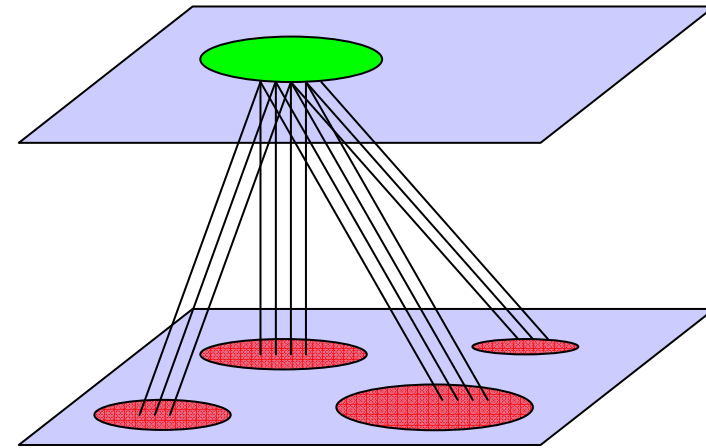
“Close-knitness” between communities

case1: unevenly distributed



QAs attract limited users

case2: evenly distributed



QAs attract several users

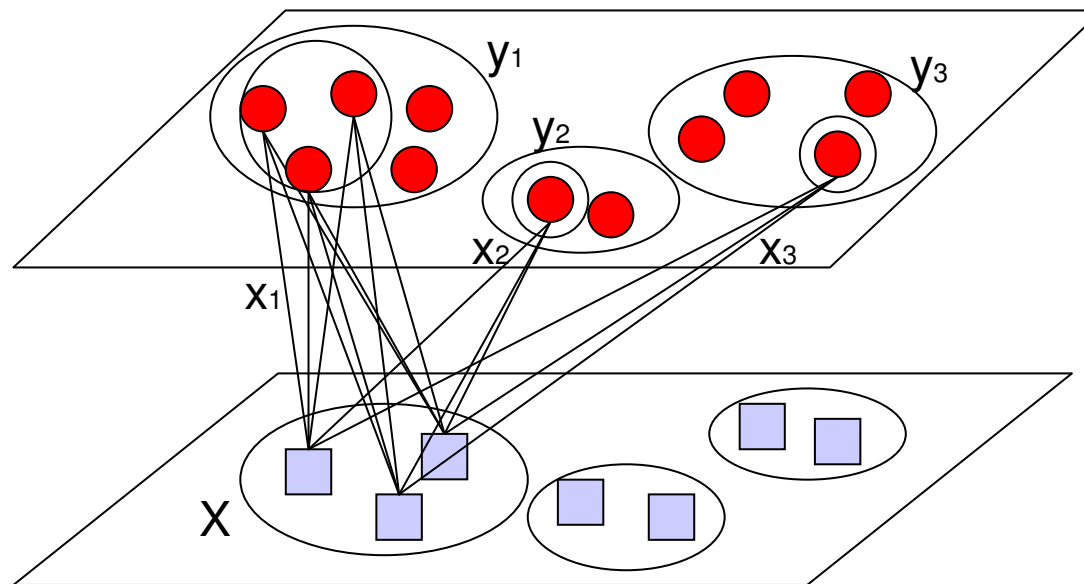
- Properties of communities in one layer can be explained by the connections to the other layer

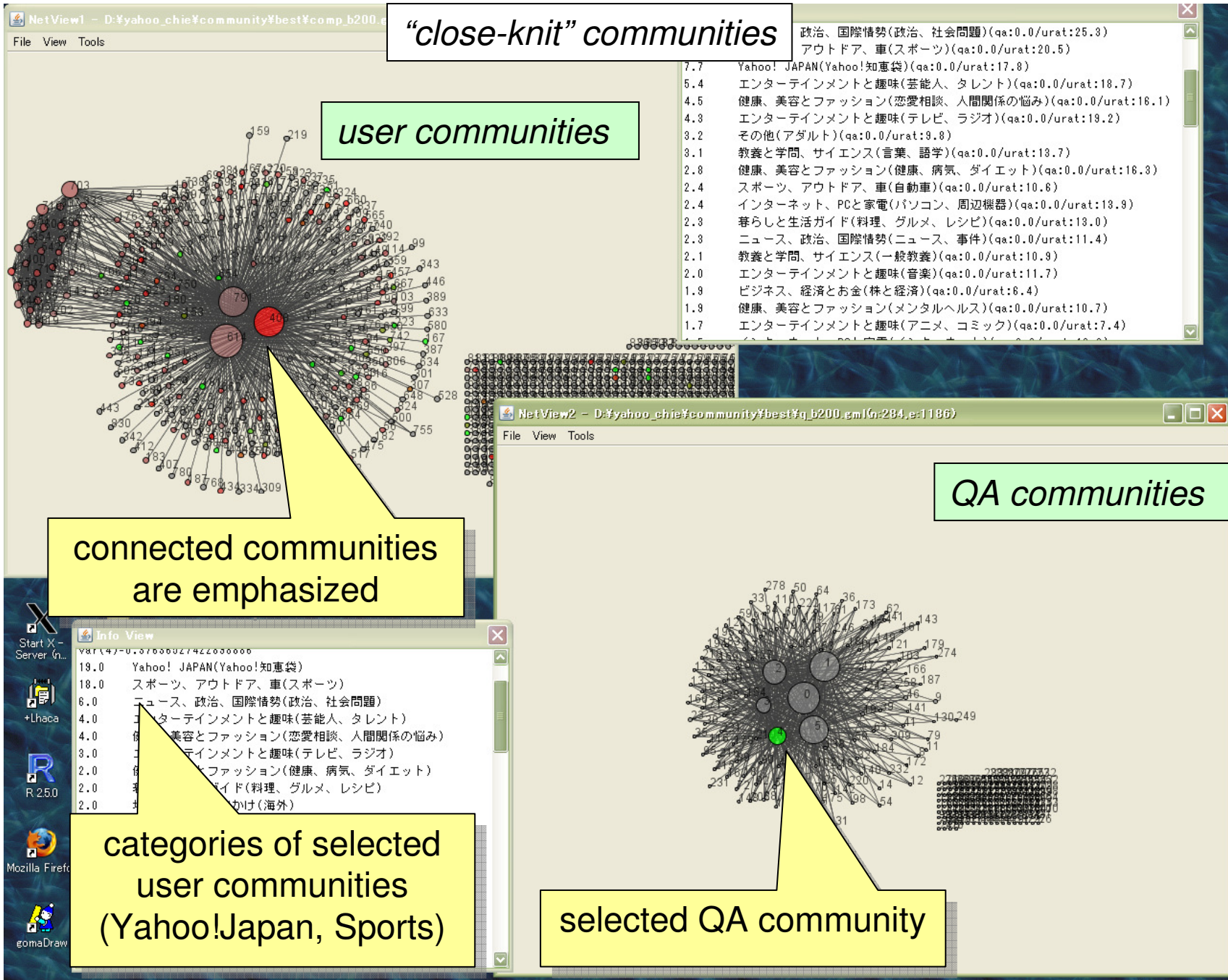
Community variance

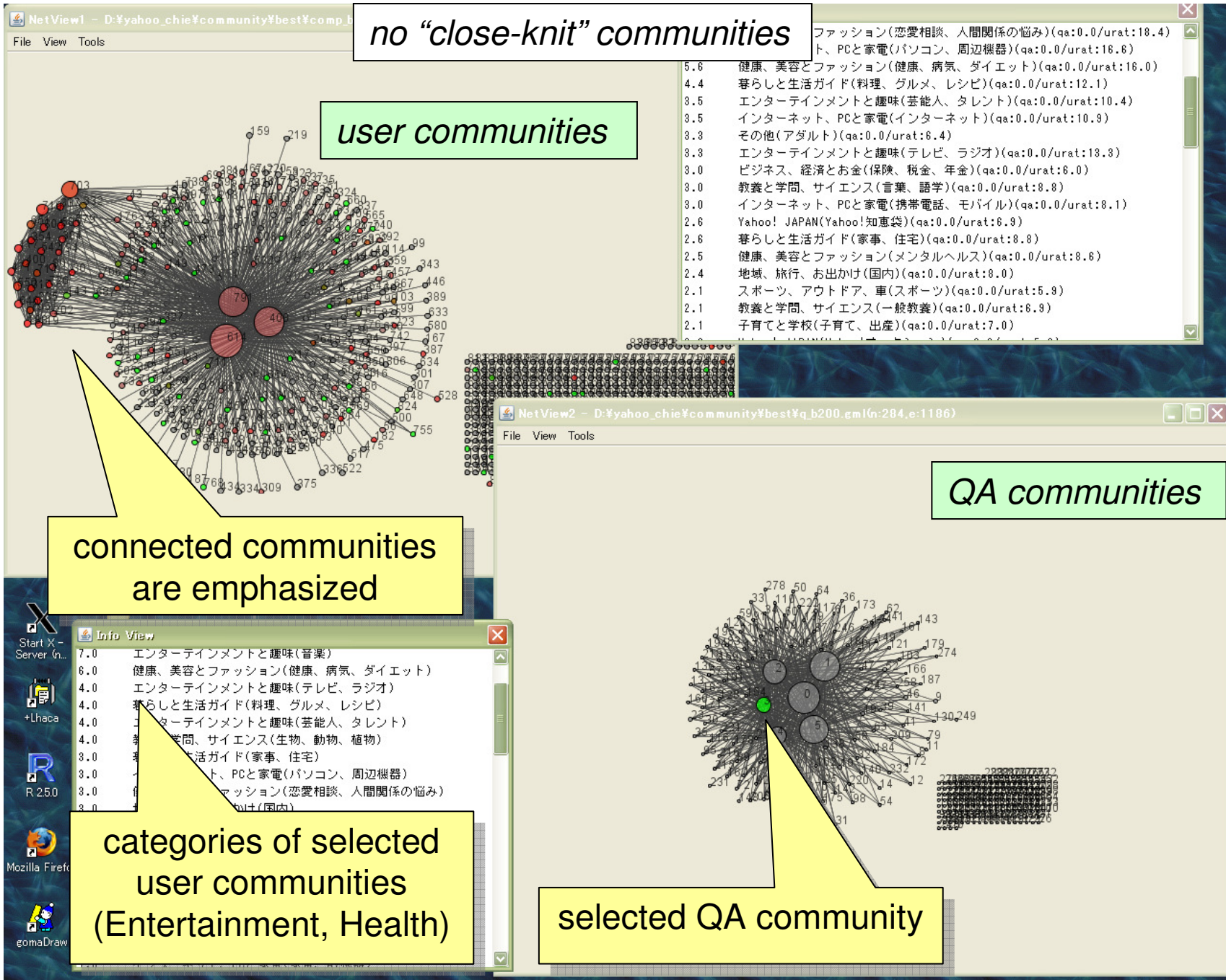
- “close-knitness” of communities in different layers

$$v = \frac{1}{X^2} \sum_i (x_i - \frac{y_i}{Y} X)^2 \quad X = \sum_i x_i \quad Y = \sum_i y_i$$

- no correlation \rightarrow community variance = 0









Evaluation

- showing validity of communities is not easy
- similarities between people
 - personal information (such as gender, age) cannot be obtained because of privacy reasons
- similarities between QAs
 - # of QAs in each category is shown for the sake of convenience

top/bottom-3 QA communities

nodes	com-var	Categories (# of posts)
13	0.9320	Internet,PC(61),Sports(7),Yahoo! Japan(7)
18	0.7646	Internet,PC(72),Sports(11),Life(5)
19	0.6851	Jobs(50),Life(20),Entertainment(20)
726	0.0089	Entertainment(17),Health(16),Internet,PC(12)
26	0.0054	Internet,PC(34),News(23),Yahoo! Japan(15)
83	0.0027	Yahoo!Japan(21),Finance(15),Health(12)



Concluding Remarks

- discovering communities from bipartite social networks
- Measuring the correspondence between communities of different layers
- Future work: discovering communities from tripartite, ..., n-partite networks