Community Detection and Interpretation

Günce Keziban Orman

_____

Galatasaray University, Computer Engineering Department

May 26, 2018

CV

- Education
  - B.Sc. Galatasaray University, Computer Engineering, Istanbul, 2005
  - M.Sc. Galatasaray University, Computer Engineering, Istanbul, 2010
  - Ph.D. INSA de Lyon, Computer Science&Informatics, Lyon, 2014
- Experiences
  - Database Manager, Is Bankasi, Istanbul, 2005-2006
  - Researcher, Tübitak, Gebze, 2006-2009
  - Research and Teaching Assistant, Galatasaray University, Computer Engineering, Istanbul, 2009-2016
  - Research Consultant, Vodafone, Istanbul, 2014-2015
  - Assistant Professor, Galatasaray University, Computer Engineering, Istanbul, 2016-...
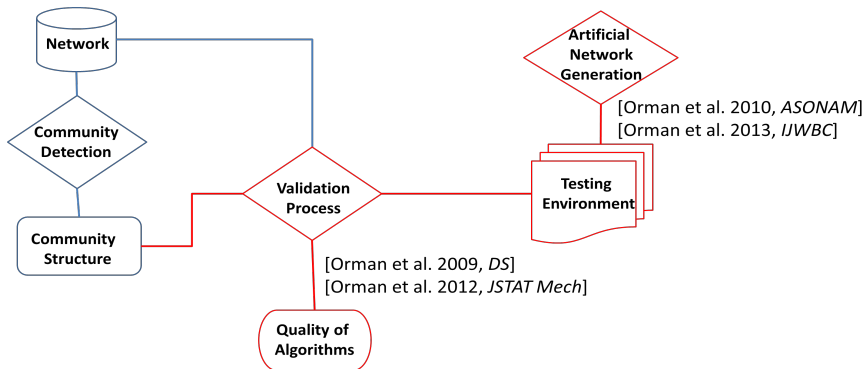
Previous Studies

[Fortunato 2010, *Phys. Report*]
[Zhou et al. 2009, *VLDB*][Yang et al. 2013, *ICDM*]
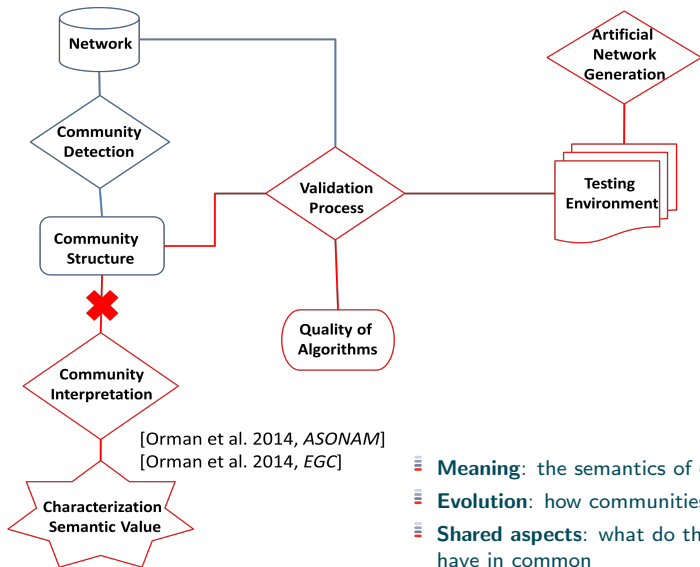[Aynaud et al. 2011, *Dynamics on and of Complex Networks Vol. 2*]

- Hundreds of methods for plain networks
- Considering homogeneity for attributed networks
- Consensual or evolving community structures for dynamic networks

Previous Studies

- Artificial networks with realistic properties
- Evaluating algorithms performance by using artificial networks

Previous Studies

[Orman et al. 2014, *ASONAM*]
[Orman et al. 2014, *EGC*]

- **Meaning**: the semantics of communities
- **Evolution**: how communities change through time
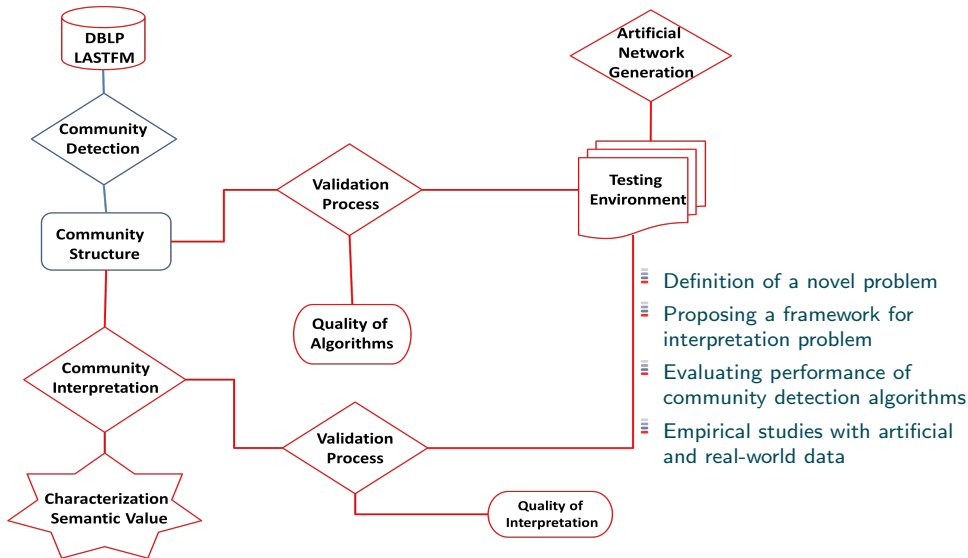- **Shared aspects**: what do their constituting nodes have in common
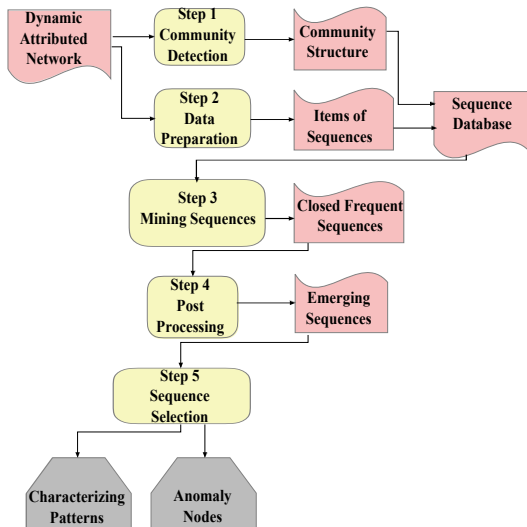
Community in Sociology

Groups of people sharing

- A common territory (neighborhood, town, city, etc.)
  [Gusfield 1975, *Harper & Row*]

- Having common relationships (social relationships, family, etc.) or emotions
  [McMillan and Chavis 1986, *Journal of Community Psychology*]

- Common <u>behavior</u> and <u>property</u>
  [McMillan and Chavis 1986, *Journal of Community Psychology*]

Community Interpretation Methodology

- Definition of a novel problem
- Proposing a framework for interpretation problem
- Evaluating performance of community detection algorithms
- Empirical studies with artificial and real-world data
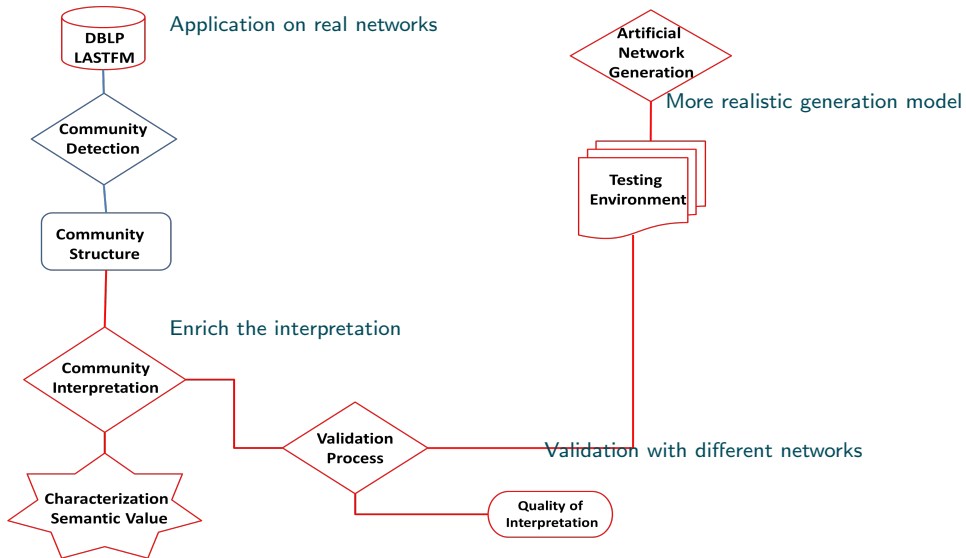
Community Interpretation Framework

General Results

Lessons from the empirical validation

- Both topology and attributes support the interpretation process
- Question of Homogeneity
  - Homogeneity low level
    - Node groups are homogeneous
    - Homogeneous groups tend to be together as time goes by
  - Homogeneity high level
- Question of Community Evolution
  - Homogeneous node groups tend to merge or expand

Perspectives

Application on real networks

DBLP LASTFM

Community Detection

Community Structure

Enrich the interpretation

Community Interpretation

Characterization Semantic Value

Validation Process

Quality of Interpretation

Artificial Network Generation

More realistic generation model

Testing Environment

Validation with different networks

Perspectives

Artificial Network Generation

- Model with realistic attributes
- Model with combination of community evolution events

Applications

- Networks from different domains
- Different types of attributes
 - Textual, image, etc.

Interpretation

- Distance between communities
  - Which descriptors specify more the communities?
- The most representative nodes
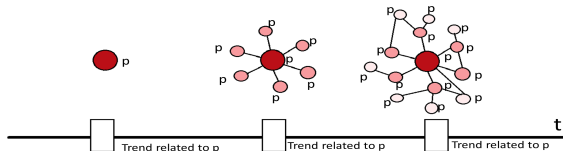
Roles of Nodal Changes in Dynamic Networks

▤ Assumption that the neighborhood changes reflects the role and
  position of the node

▤ Characterizes each individual node by studying the evolution of
  its direct neighborhood

▤ Define the concept of neighborhood event

  • Six types of such events: birth, death, merge, split, expansion
    and shrink

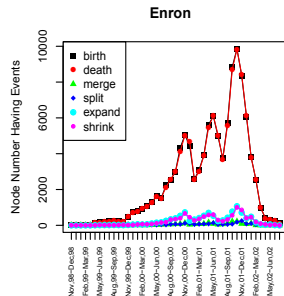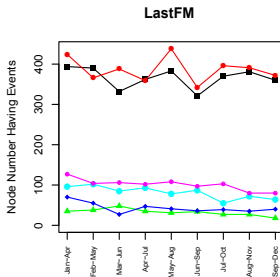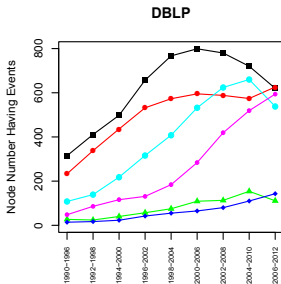▤ Study the affects of the evolution of those events

How to find the events
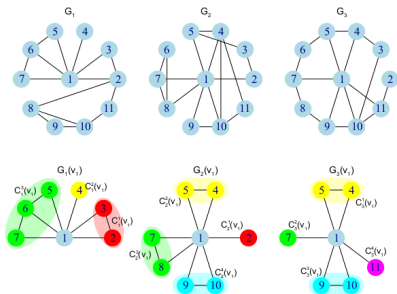
Experimental Results

Table: Real-world networks used in the experimental evaluation

Network	Nodes	Active nodes	Time slices	Time span
DBLP	2145	2046	10	1990-2012
LastFM	1701	1269	10	Jan-Dec 2013
Enron	28802	28649	46	1997-2002

THANK YOU