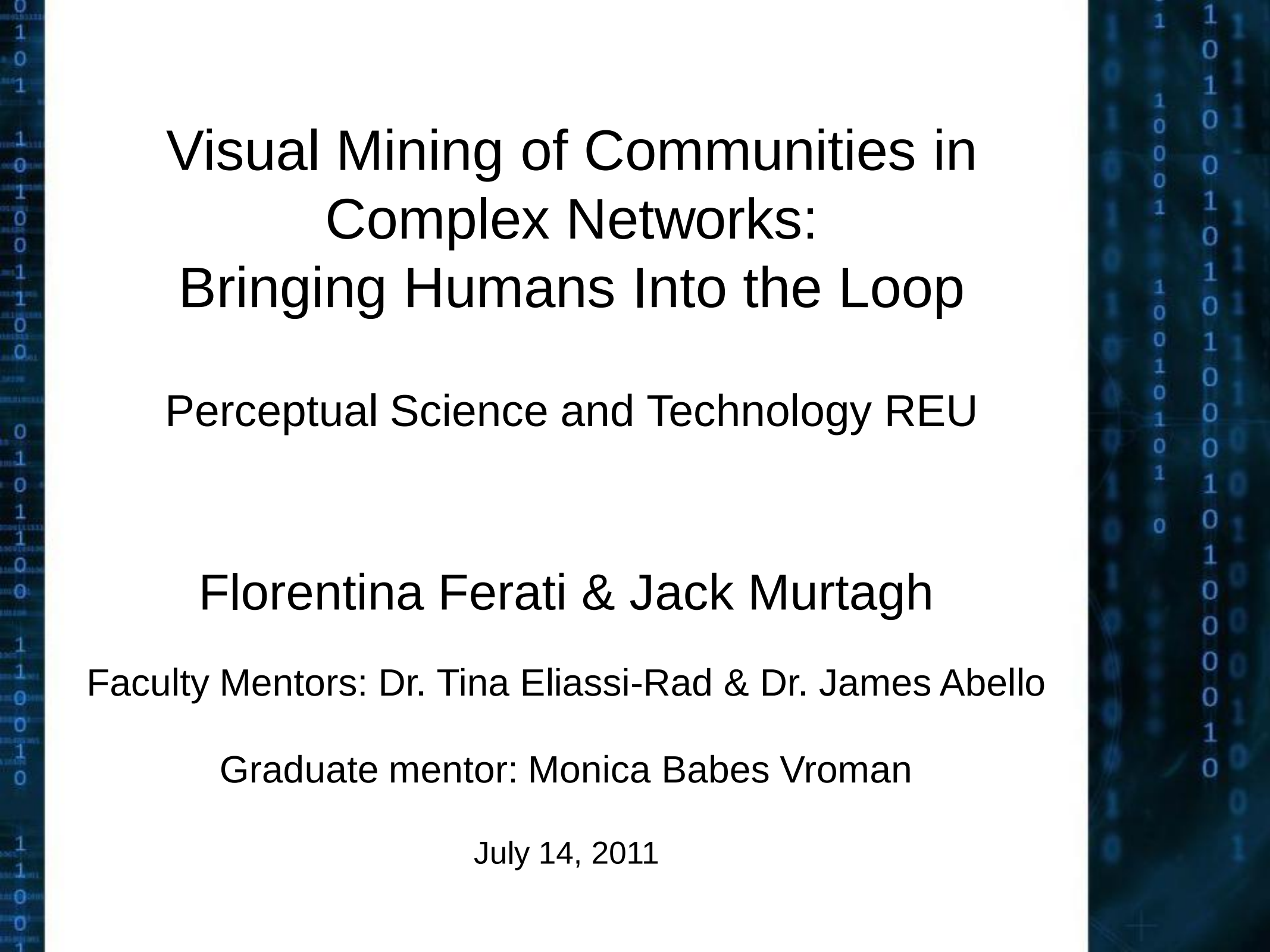




Visual Mining of Communities in Complex Networks: Bringing Humans Into the Loop

Perceptual Science and Technology REU

Florentina Ferati & Jack Murtagh

Faculty Mentors: Dr. Tina Eliassi-Rad & Dr. James Abello

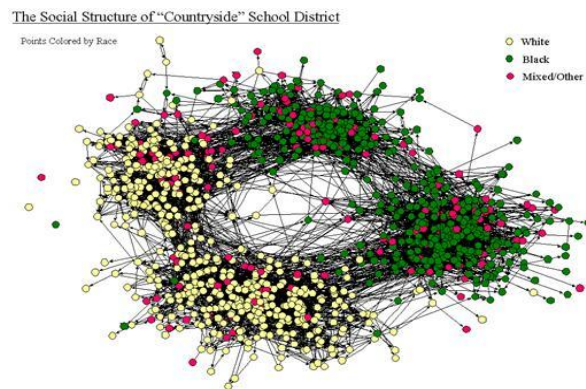
Graduate mentor: Monica Babes Vroman

July 14, 2011

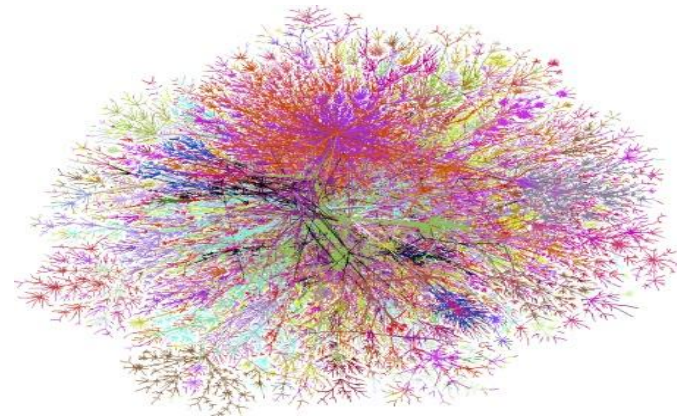
Motivation

- Complex networks are ubiquitous
 - Examples: social networks, Internet, WWW, etc.

Friendship



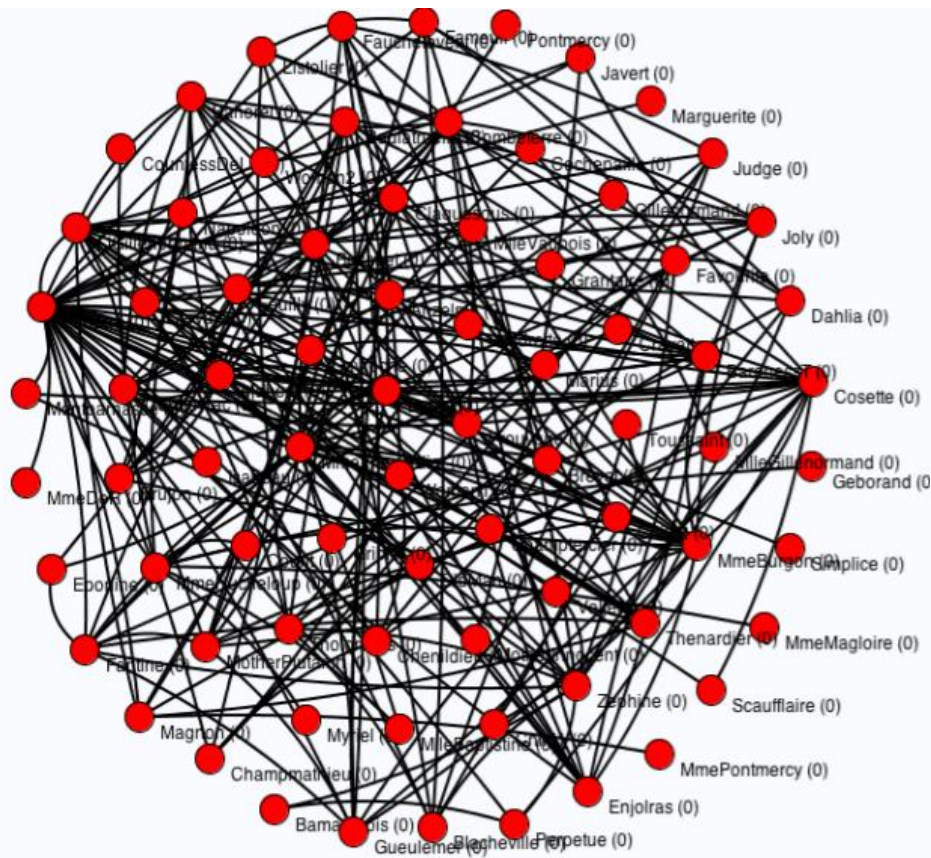
Internet



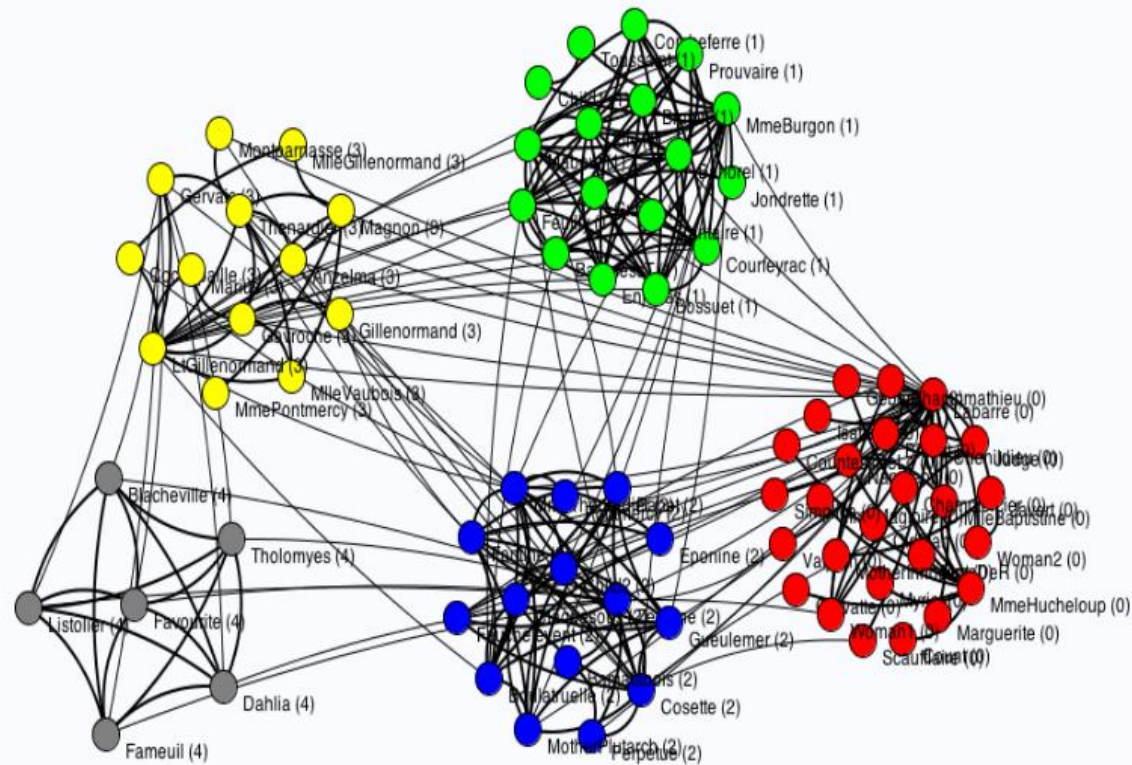
- Community discovery & visual analytics allow us to “make sense” of the underlying structure of the network

Les Miserables Character Co-occurrence Network

Before

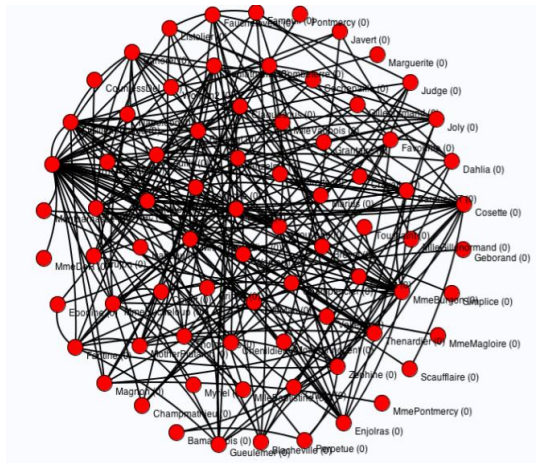


After

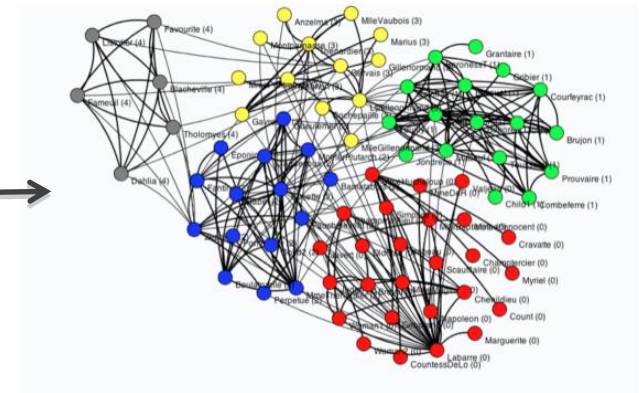


Goal: Bring Human in the Loop

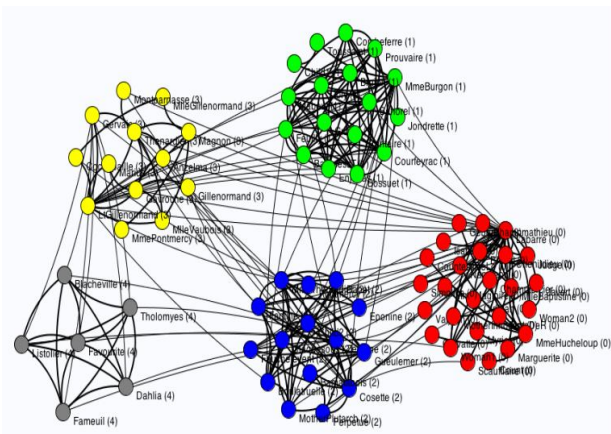
Input Network



Community Discovery
(*Fast Modularity**)



Visual Analytics
(*our contribution*)



Output Network



* [Clauset, et al. 2004]

Fast Modularity

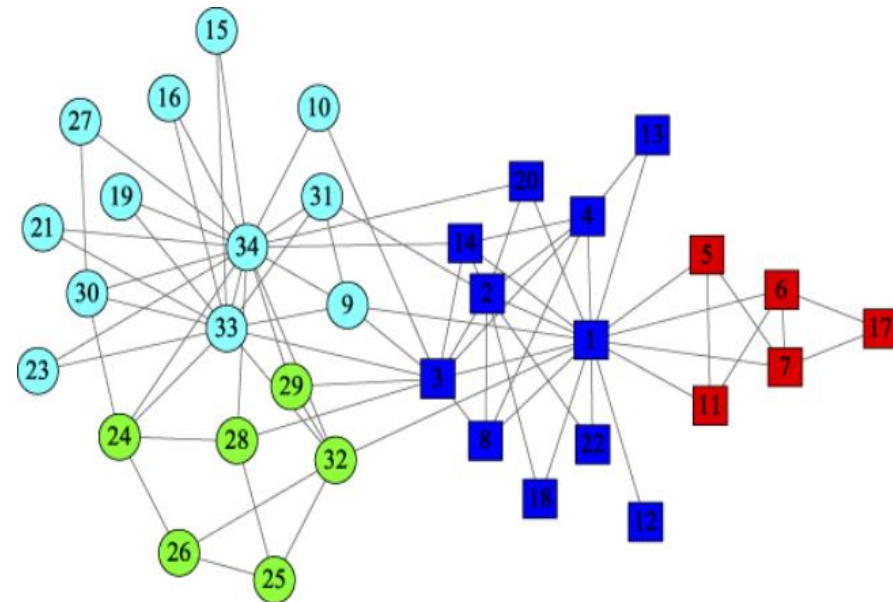
[Clauset, et al., *Phys. Rev. E*. 2004]

- Maximizes modularity, Q : measures the **fraction of all edges within communities** minus **the expected number in a random graph with the same degrees**

$$Q = \frac{1}{2m} \sum_{vw} \left[A_{vw} - \frac{k_v k_w}{2m} \right] \delta(c_v, c_w)$$

- m = number of edges in the graph
- $A_{vw} = 1$ if $v \rightarrow w$; 0 otherwise
- k_v = degree of vertex v
- $\delta(i, j) = 1$ if $i == j$; 0 otherwise

- Part of the Community Analysis Tool (CAT)

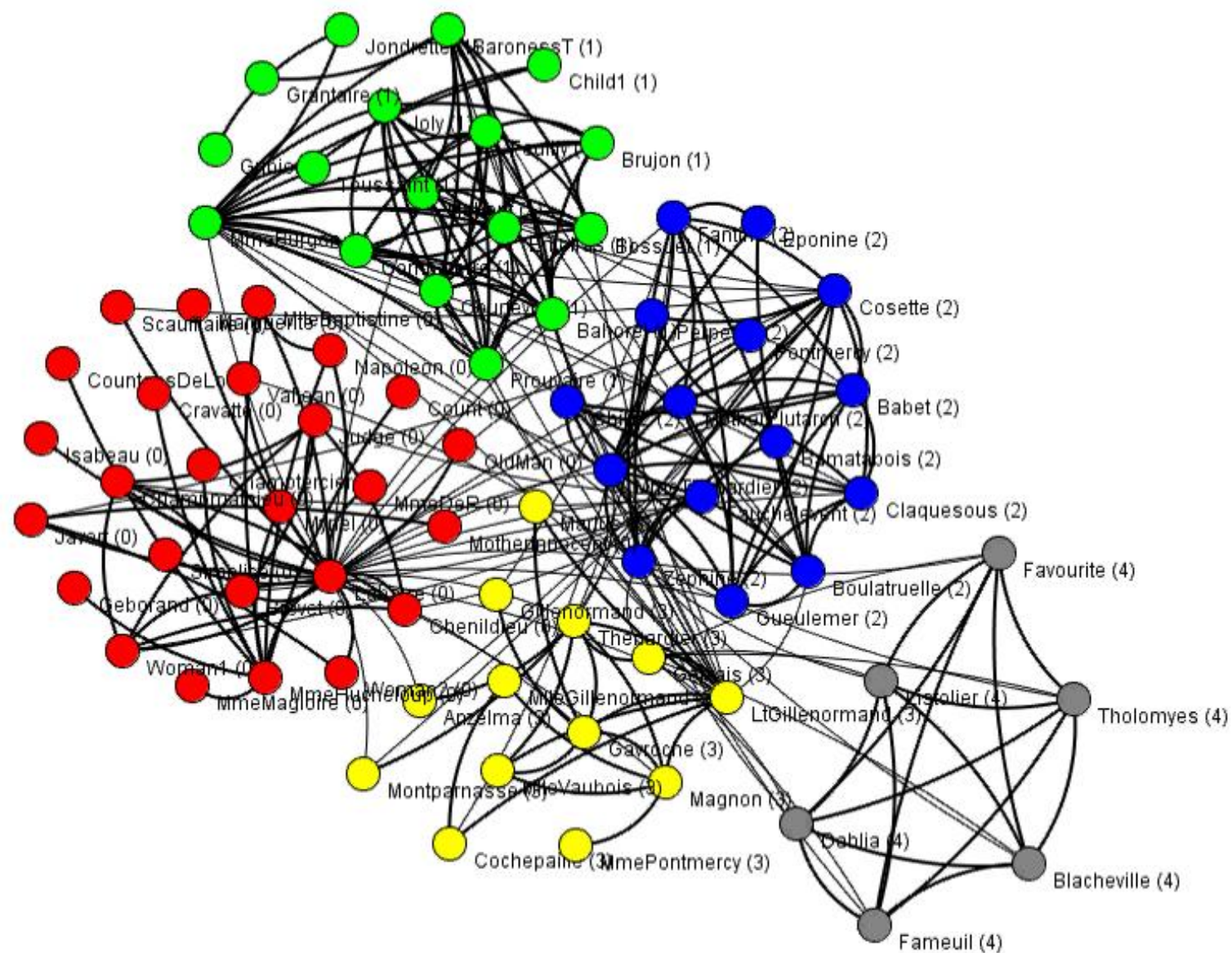


Our Contribution So Far (I)

- Allow user to change the community of a node
 - Update its color and position
 - Track changes
 - Re-compute modularity
 - Warn user if his/her action decreases modularity score too much
- Add new communities

Our Contribution So Far (II)

- Suggest changes
 - Which nodes are least “comfortable” in their community
 - Nodes blink the color of their desired community
- Provide a more detailed history of changes



Graph

Communities = 5 Modularity = 0.50059676 Change = 0.0

Next Steps

- Find a quicker way to determine a node's "comfort"
- Make size of nodes proportional to their "comfort"
- Incorporate constraints from the user before clustering

References

- [1] A. Clauset, M.E.J. Newman and C. Moore, "Finding community structure in very large networks."
Phys. Rev. E 70, 066111 (2004).
- [2] Keith Henderson, Tina Eliassi-Rad, Spiros Papadimitriou, Christos Faloutsos: HCDF: A Hybrid Community Discovery Framework. SDM 2010: 754-765.
- [3] J. Abello, F. van Ham, and N. Krishnan, "Ask-graphview: A large scale graph visualization system", IEEE TVCG journal, Vol. 12, No. 5, pp. 669–676, 2006.
- [4] J. Leskovec, K. Lang, M. Mahoney: Empirical Comparison of Algorithms for Network Community Detection. WWW 2010:631-640.
- [5] Zeqian Shen, Kwan-Liu Ma, Tina Eliassi-Rad: Visual Analysis of Large Heterogeneous Social Networks by Semantic and Structural Abstraction. IEEE Trans. Vis. Comput. Graph. 12(6): 1427-1439 (2006).

Thank You!

- Questions?

