Using Strong Triadic Closure to Characterize Ties in Social Networks

Stavros Sintos*, Panayiotis Tsaparas
University of Ioannina

* Currently at Duke University

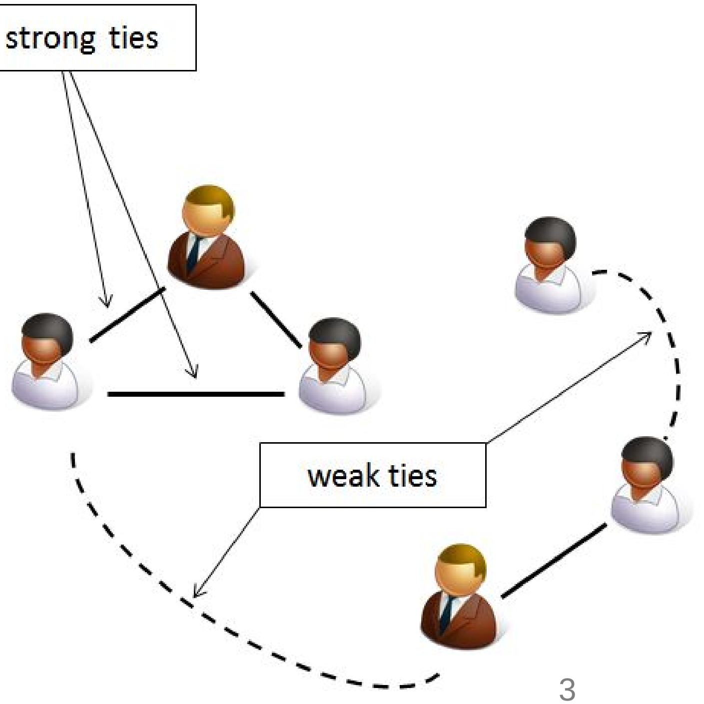
Introduction

- Who are true friends in Online Social Networks?

Introduction

- Understanding the strength and nature of online relationships is critical for many applications
  - Online Marketing
  - Friendship Suggestions
  - Sociology

- In this work: Use the graph structure to characterize social ties



Our Goal

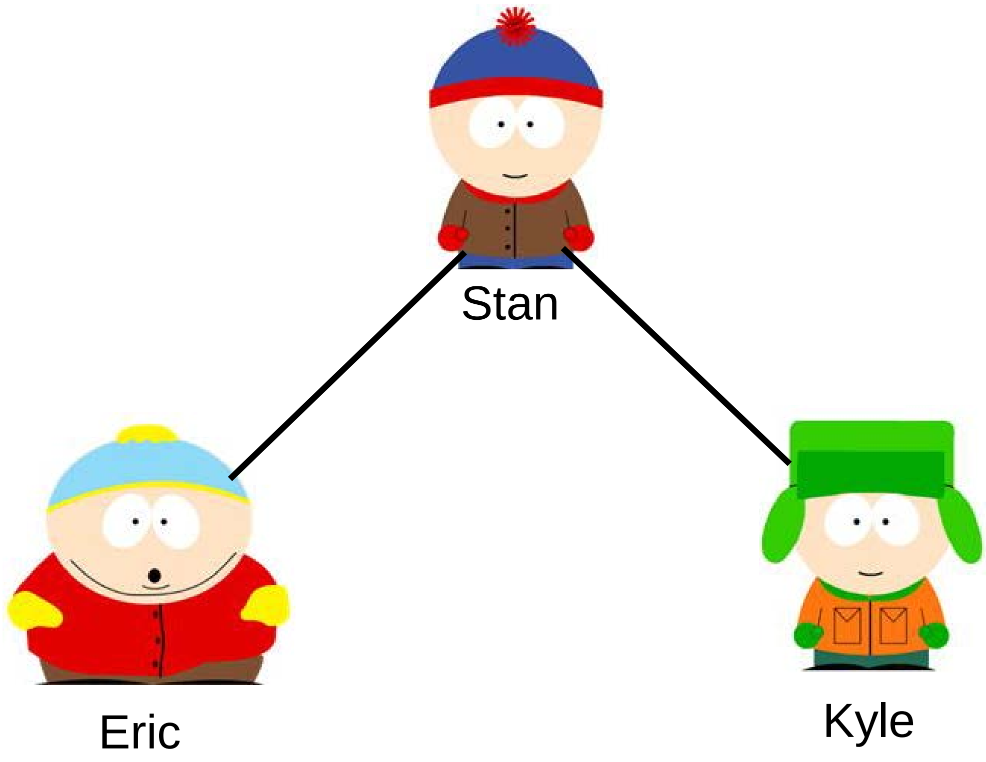
- Characterize relationships in a social network as Strong or Weak using the Strong Triadic Closure property.

- Strong Triadic Closure
  - [M. Granovetter. The strength of weak ties. American Journal of Sociology 1973]
  - [D. Easley, J. Kleinberg. Networks, Crowds, and Markets: Reasoning About a Highly Connected World. Cambridge University Press 2010]

INTRODUCTION > **PROBLEM DEFINITION** > ALGORITHMS > EXPERIMENTS > CONCLUSION

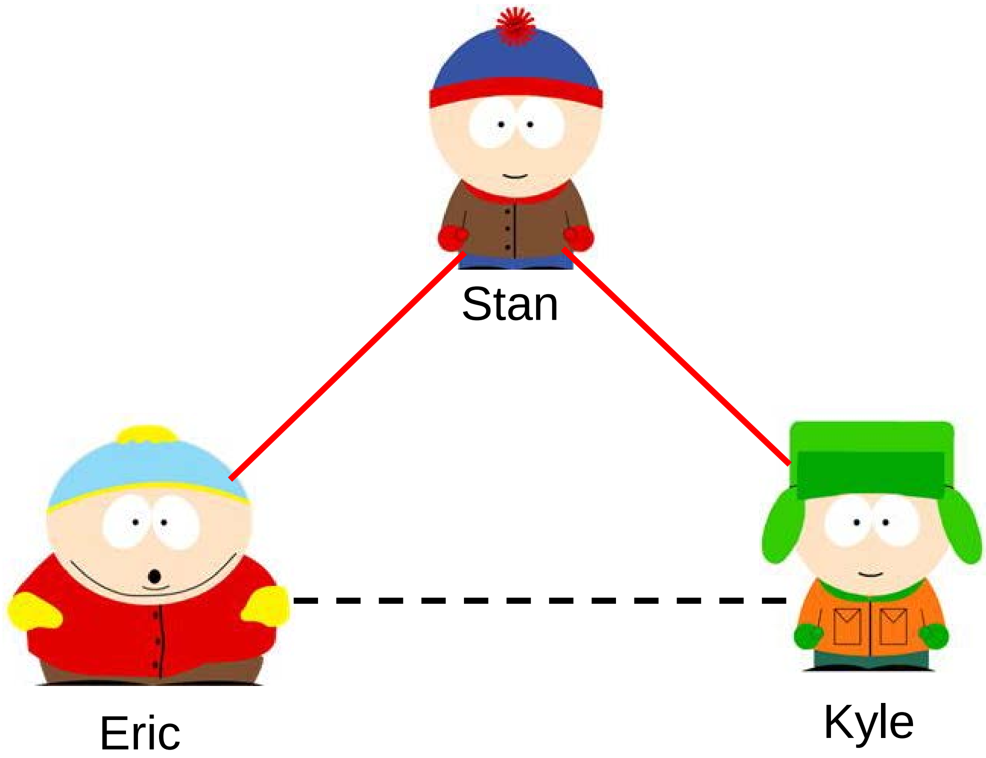
Strong Triadic Closure intuition

- If a person has strong relationships with two of his friends, these friends must know each other.



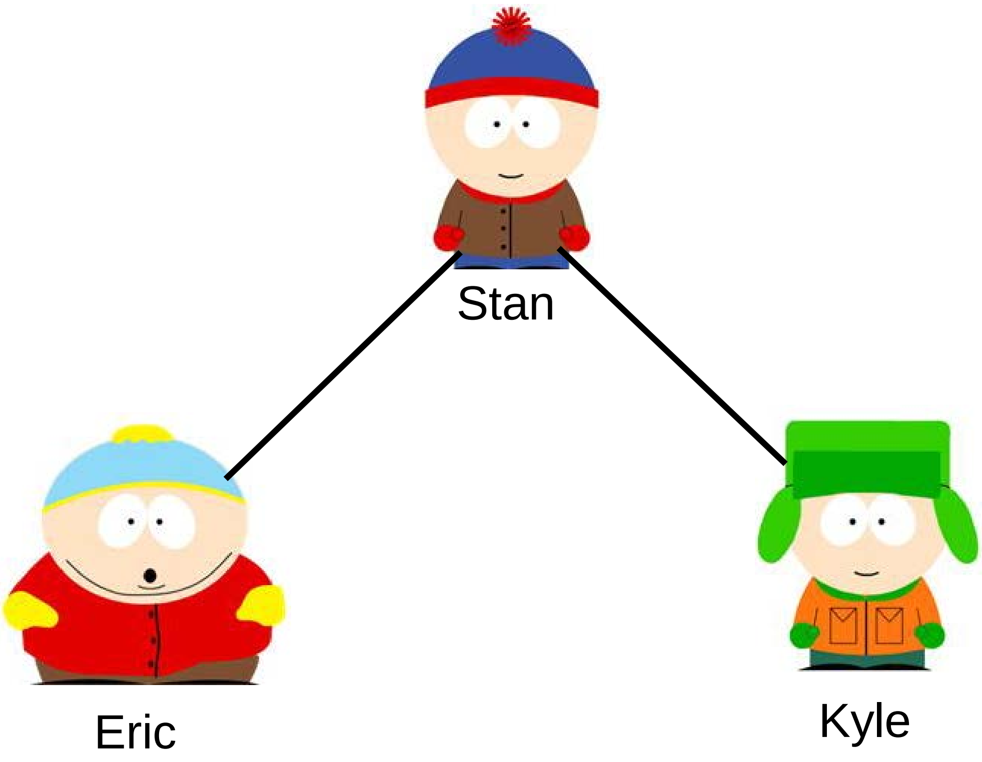
Strong Triadic Closure intuition

- If a person has strong relationships with two of his friends, these friends must know each other.



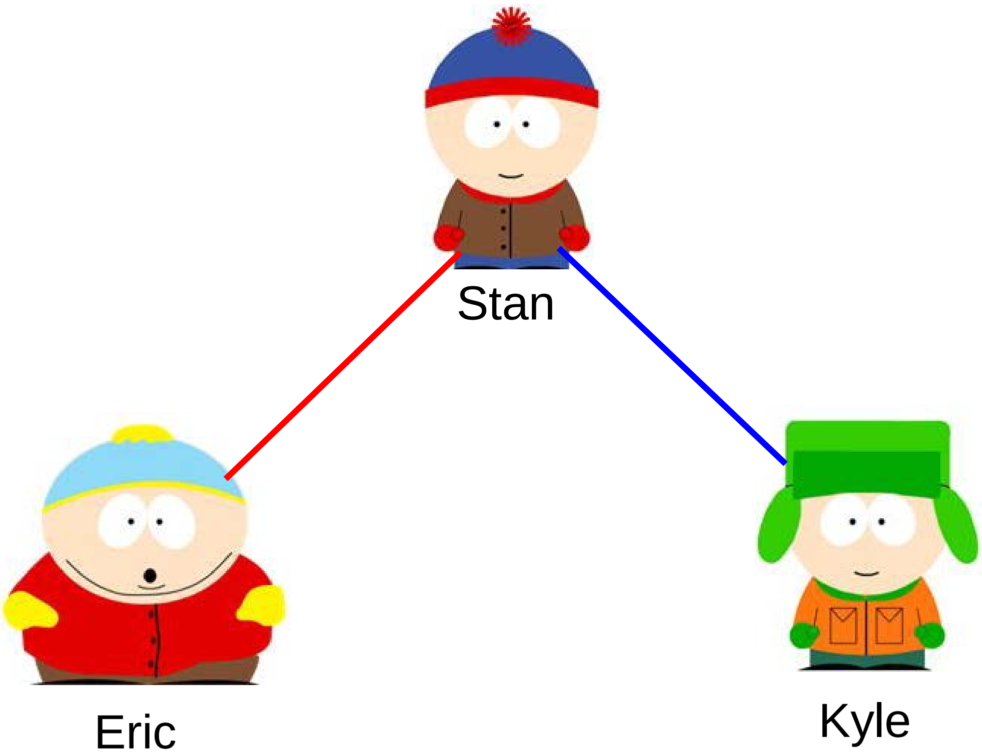
Strong Triadic Closure intuition

- If a person has strong relationships with two of his friends, these friends must know each other.



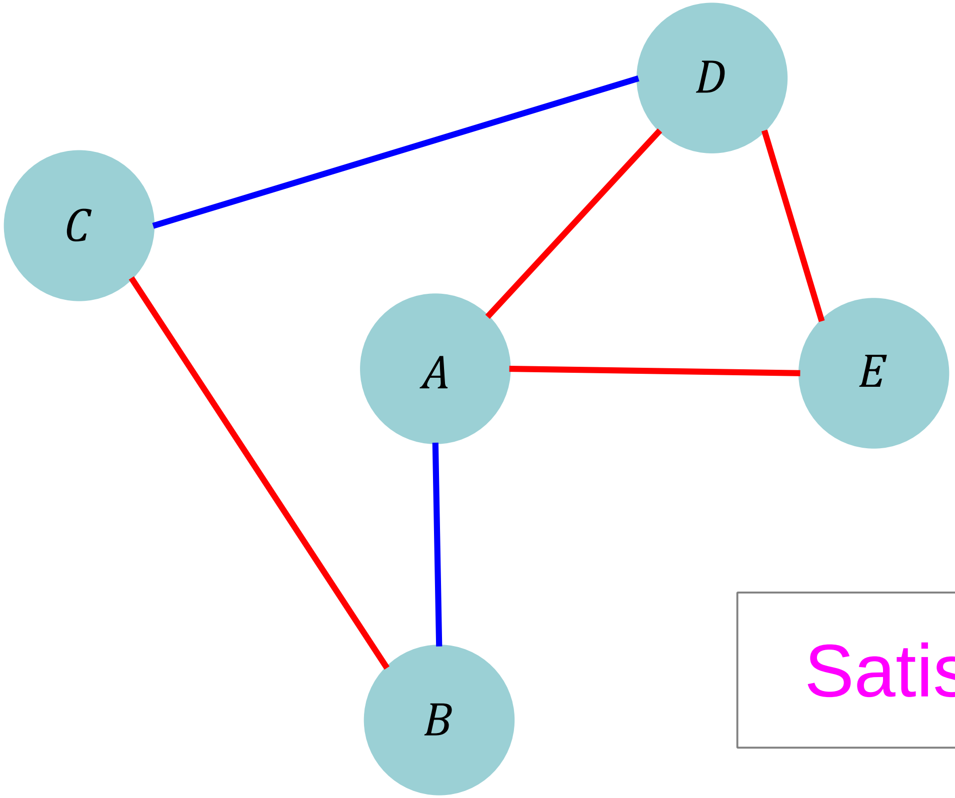
Strong Triadic Closure intuition

- If a person has strong relationships with two of his friends, these friends must know each other.



Strong Triadic Closure for graphs

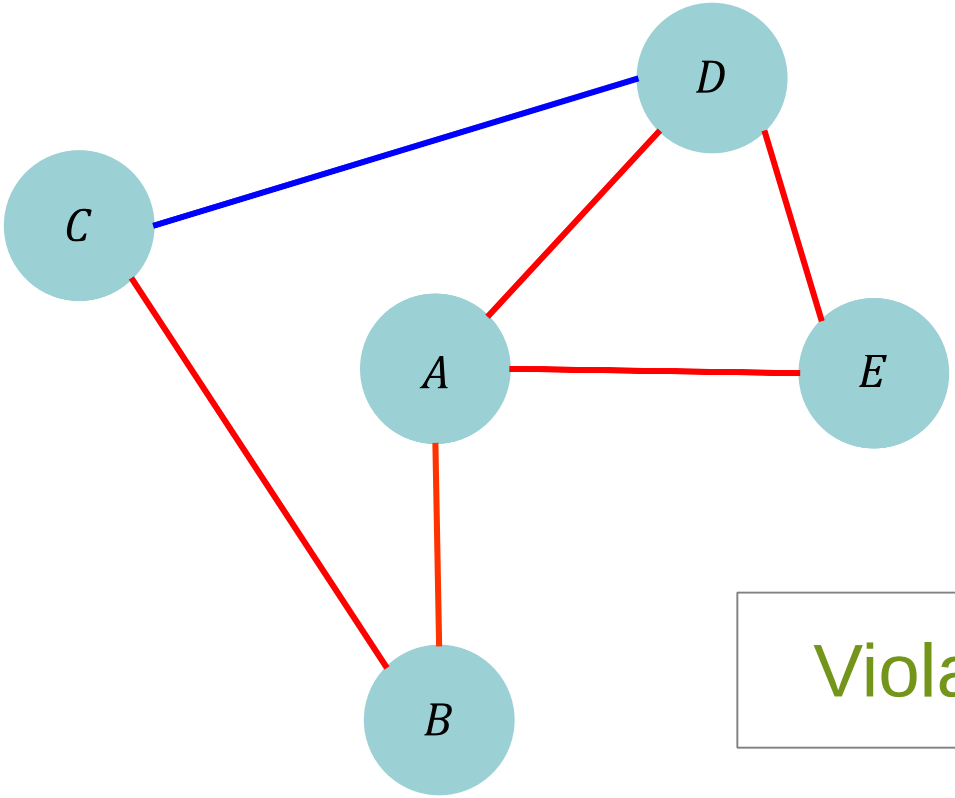
- Consider a labeling of the edges of a graph as Strong or Weak.
- The labeling satisfies the Strong Triadic Closure property if there is no open triangle with two Strong edges.



Satisfies the STC property

Strong Triadic Closure for graphs

- Consider a labeling of the edges of a graph as Strong or Weak.
- The labeling satisfies the Strong Triadic Closure property if there is no open triangle with two Strong edges.



Violates the STC property

Problem Definition

- Label ties of a social network as Strong or Weak so that the Strong Triadic Closure property holds.
- MaxSTC: Find an edge labeling (S, W) that satisfies the STC property and maximizes the number of Strong edges.
- MinSTC: Find an edge labeling (S, W) that satisfies the STC property and minimizes the number of Weak edges.

INTRODUCTION > PROBLEM DEFINITION > **ALGORITHMS** > EXPERIMENTS > CONCLUSION

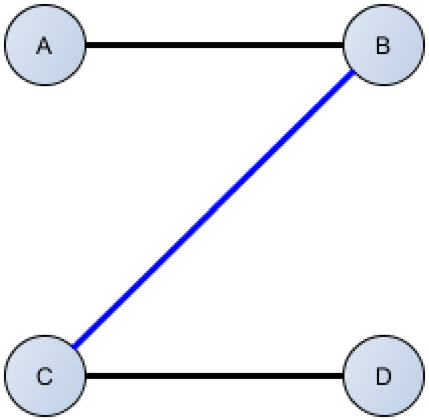
Complexity

- Bad News: MaxSTC and MinSTC are NP-hard problems!
  - Reduction from MaxClique to the MaxSTC problem.

Approximation

- Approximation Algorithms
  - Bad News: MaxSTC is hard to approximate.
  - Good News: There exists a 2-approximation algorithm for the MinSTC problem.

- Intuition:
  - No open triangle with 2 strong edges.
  - For every open triangle: a weak edge must cover this triangle.



- MinSTC can be mapped to the Minimum Vertex Cover problem.

Vertex Cover

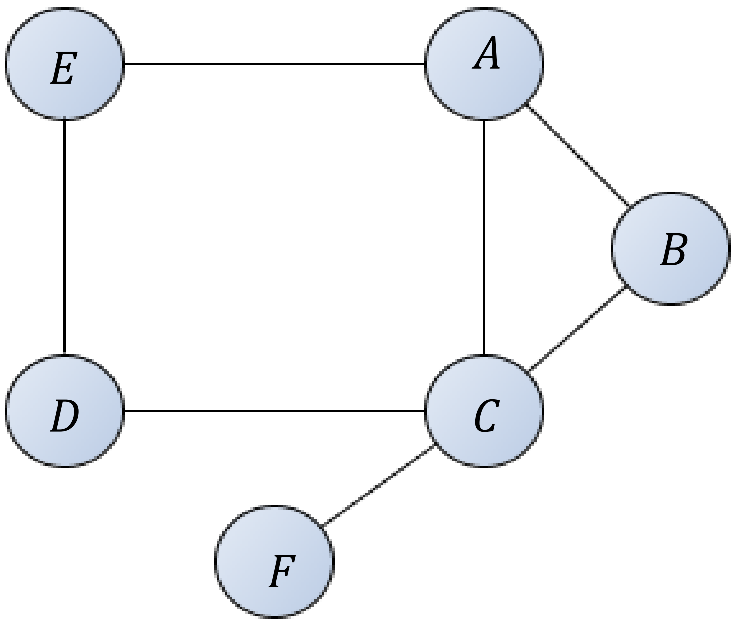
- Vertex Cover: For graph $G = (V, E)$ a vertex cover is a subset of vertices $C \subseteq V$, such that for every edge $(u, v) \in E, u \in C$ or $v \in C$.
- Minimum Vertex Cover Problem: Find a vertex cover with the smallest number of vertices.

Dual Graph

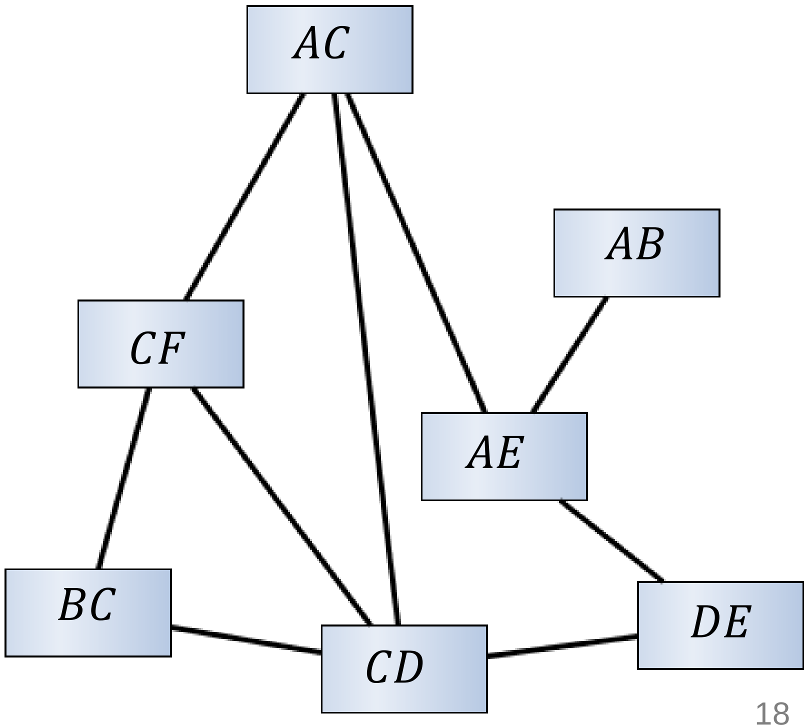
Given a graph G, we create the dual graph D:

- For every edge in G we create a node in D.
- Two nodes in D are connected if the corresponding edges in G participate in an open triangle.

Initial Graph G

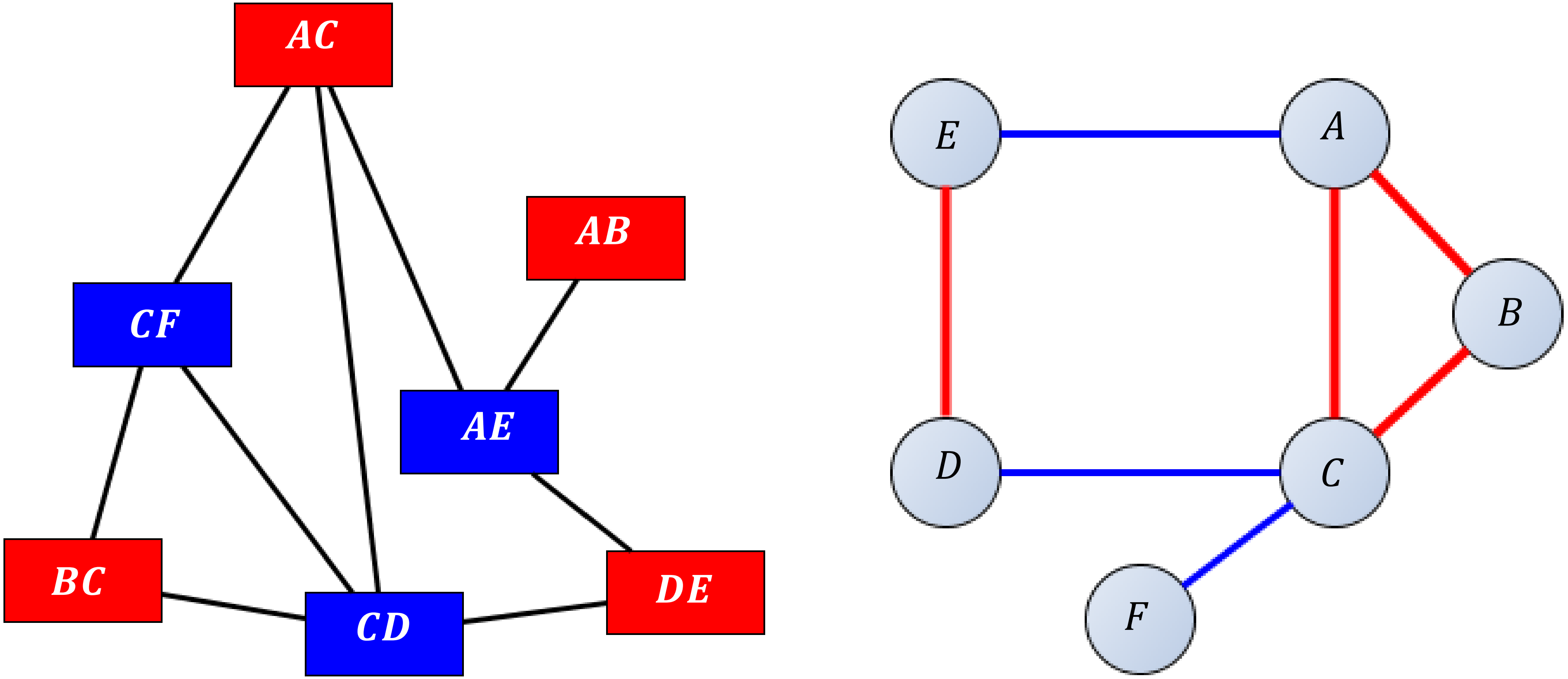


Dual Graph D



Minimum Vertex Cover - MinSTC

- Solving MinSTC on G is reduced to solving a Minimum Vertex Cover problem on D.



Approximation Algorithms

Approximation algorithms for the Minimum Vertex Cover problem:

Maximal Matching Algorithm	Greedy Algorithm
▪ Output a maximal matching • Maximal Matching: A collection of non-adjacent edges of the graph where no additional edges can be added. Approximation Factor: 2	▪ Greedily select each time the vertex that covers most uncovered edges. Approximation Factor: log n

- Given a vertex cover for dual graph D, the corresponding edges of G are labeled Weak and the remaining edges Strong.

INTRODUCTION > PROBLEM DEFINITION > ALGORITHMS > **EXPERIMENTS** > CONCLUSION

Experimental Goal

- Does our labeling have any practical utility?

Datasets

- Actors: Collaboration network between movie actors. (IMDB)
- Authors: Collaboration network between authors. (DBLP)
- Les Miserables: Network of co-appearances between characters of Victor Hugo's novel. (D. E. Knuth)
- Karate Club: Social network of friendships between 34 members of a karate club. (W. W. Zachary)
- Amazon Books: Co-purchasing network between books about US politics. (http://www.orgnet.com/)

Dataset	Number of Nodes	Number of Edges
Actors	1,986	103,121
Authors	3,418	9,908
Les Miserables	77	254
Karate Club	34	78
Amazon Books	105	441

Experiments

	Greedy		Maximal Matching	
	Strong	**Weak**	**Strong**	**Weak**
Actors	11,184	91,937	8,581	94,540
Authors	3,608	6,300	2,676	7,232
Les Miserables	128	126	106	148
Karate Club	25	53	14	64
Amazon Books	114	327	71	370

Measuring Tie Strength

- Question: Is there a correlation between the assigned labels and the empirical strength of the edges?
- Three weighted graphs: Actors, Authors, Les Miserables.
  - Strength: amount of common activity.

Mean activity intersection for Strong, Weak Edges

	Strong	Weak
Actors	1.4	1.1
Authors	1.34	1.15
Les Miserables	3.83	2.61

- The differences are statistically signicant

Measuring Tie Strength

- Frequent common activity may be an artifact of frequent activity.
- Fraction of activity devoted to the relationship
  - Sterngth: Jaccard Similarity of activity

$$Jaccard\ Similarity = \frac{Common\ Activities}{Union\ of\ Activities}$$

Mean Jaccard similarity for Strong, Weak Edges

	Strong	Weak
Actors	0.06	0.04
Authors	0.145	0.084

- The differences are statistically signicant

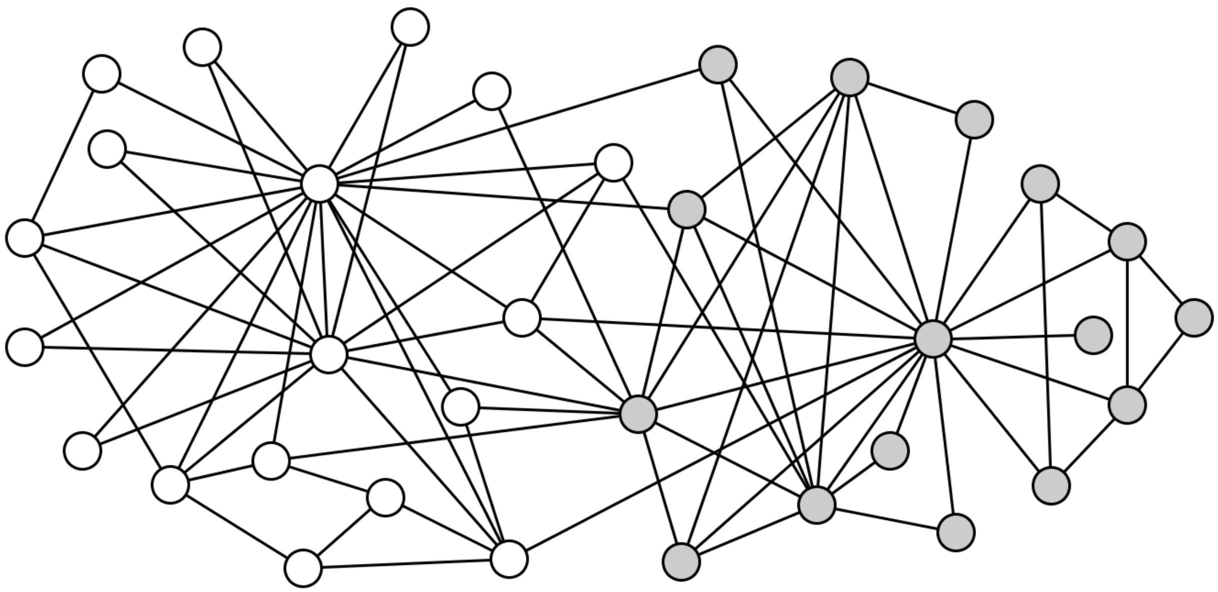
The Strength of Weak Ties

- [Granovetter] People learn information leading to jobs through acquaintances (Weak) rather than close friends (Strong).

- [Easly and Kleinberg] Graph theoretic formalization:
  - Acquaintances (Weak) act as bridges between different groups of people with access to different sources of information.
  - Close friends (Strong) belong to the same group of people, and are exposed to similar sources of information.

Datasets with known communities

- Amazon Books
  - US Politics books : liberal, conservative, neutral.
- Karate Club
  - Two fractions within the members of the club.

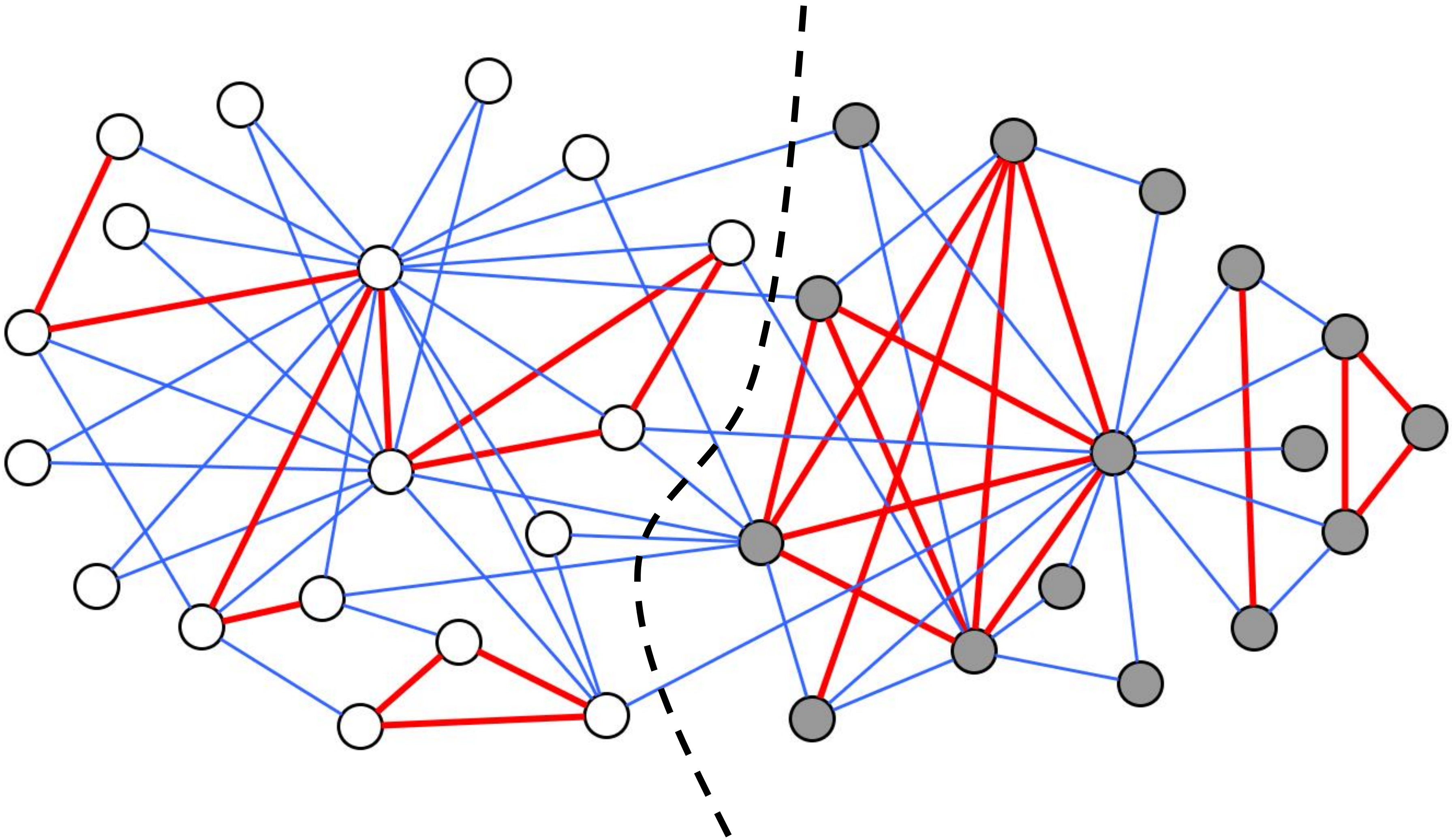


Weak Edges as Bridges

- Edges between communities (inter-community) ⇒ Weak label
  - R_W = Fraction of inter-community edges that are labeled Weak.
- Strong label ⇒edges within the community (intra-community).
  - P_S = Fraction of Strong edges that are intra-community edges

	P_S	R_W
Karate Club	1	1
Amazon Books	0.81	0.69

Karate Club graph



INTRODUCTION > PROBLEM DEFINITION > ALGORITHMS > EXPERIMENTS > **CONCLUSION**

Related Work

- Assessing the link strength in a social network
  - [I. Kahanda, and J. Neville. Using transactional information to predict link strength in online social networks. ICWSM 2009]
  - [R. Xiang, J. Neville, and M. Rogati. Modeling relationship strength in online social networks. WWW 2010]
  - [E. Gilbert, and K. Karahalios. Predicting tie strength with social media. CHI 2009]
- Characterizing the type of a relationship between users
  - [J. Tang, T. Lou, and J. Kleinberg. Inferring social ties across heterogenous networks. WSDM 2012]
  - [L. Backstrom, and J.M. Kleinberg Romantic partnerships and the dispersion of social ties: a network analysis of relationship status on facebook. CSCW 2014]

Conclusion

- Using only the graph structure and the idea of Strong Triadic Closure characterized ties in social network as Strong or Weak.
  - Also: Variations of the STC property

- Future Work
  - Consider relaxations of the STC property.
  - Further study in case of more than one strong type.

Thank you!!!

Questions?