

# Social and Information Networks

## Tutorial #4: Approximately Balanced Networks

University of Toronto CSC303  
Winter/Spring 2021  
Ian Berlot-Attwell

Week 5: Feb 8-12 (2021)

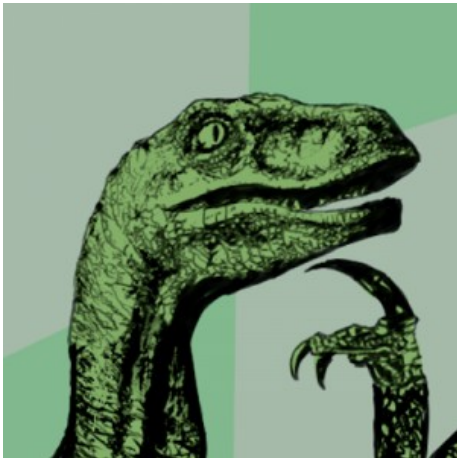
# Today's agenda

In lecture we've covered Chapter 5 of the textbook looking at Structural Balance.

Today:

- Questions from Lecture
- Recap of Structural Balance
- Approximately Balanced Networks (Ch 5.5b of E&K)
- Quercus Quiz

# Questions?

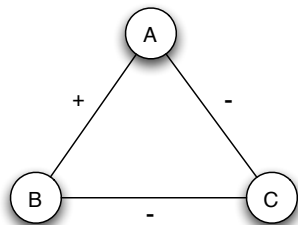
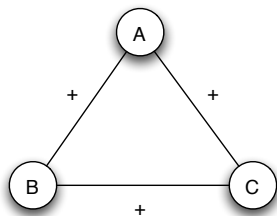


## Recap: Stable Triangles

A complete (i.e. fully connected) graph is stable if all of its triangles are stable.

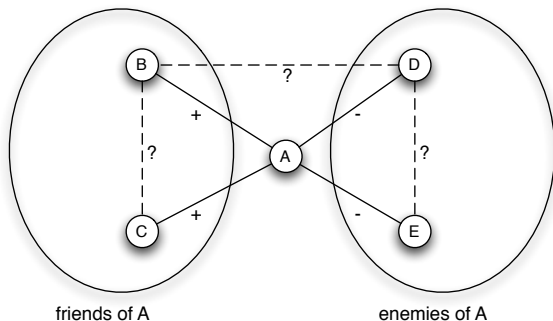
## Recap: Stable Triangles

A complete (i.e. fully connected) graph is stable if all of its triangles are stable.



## Recap: Harary Balance Theorem

If a labelled complete graph is balanced, then either everyone is friends, or the nodes can be partitioned into 2 groups that mutually loath each other and are internally purely friendly.



**Recap: Testing if a graph can be completed into a balanced graph**

## Recap: Testing if a graph can be completed into a balanced graph

- 1 Find the positive communities (i.e. BFS on only the positive edges to find the positive connected components)
- 2 Confirm that these do not contain any negative edges
- 3 Collapse the positive communities into supernodes (and we collapse the negative edges between these communities accordingly)
- 4 Check that the graph of supernodes connected by negative edges is bipartite (via modified BFS)



## Recap: Testing if a graph can be completed into a balanced graph

- 1 Find the positive communities (i.e. BFS on only the positive edges to find the positive connected components)
- 2 Confirm that these do not contain any negative edges
- 3 Collapse the positive communities into supernodes (and we collapse the negative edges between these communities accordingly)
- 4 Check that the graph of supernodes connected by negative edges is bipartite (via modified BFS)

How would we actually complete the graph given the positive communities if we wanted a true labeling?

## Recap: Testing if a graph can be completed into a balanced graph

- 1 Find the positive communities (i.e. BFS on only the positive edges to find the positive connected components)
- 2 Confirm that these do not contain any negative edges
- 3 Collapse the positive communities into supernodes (and we collapse the negative edges between these communities accordingly)
- 4 Check that the graph of supernodes connected by negative edges is bipartite (via modified BFS)

How would we actually complete the graph given the positive communities if we wanted a true labeling?

What if we only wanted to complete to a weakly balanced graph?

## Approximately Balanced Networks

Harary Balance Theorem only works when there are exactly no unbalanced triangles. Can we weaken this result to allow for some imbalanced triangles?

# Approximately Balanced Networks

Harary Balance Theorem only works when there are exactly no unbalanced triangles. Can we weaken this result to allow for some imbalanced triangles?

## Theorem

Let  $\epsilon$  such that  $0 \leq \epsilon \leq \frac{1}{8}$ , and  $\delta := \sqrt[3]{\epsilon}$ . If at least  $1 - \epsilon$  of the triangles in a complete labeled graph  $G = (V, E)$  are balanced, then either:

- 1  $\exists V' \subseteq V$  such that  $\frac{|V'|}{|V|} \geq 1 - \delta$  and the proportion of hostile pairs in  $|V'|$  is at most  $\delta$
- 2 We can partition  $V$  into  $X$  and  $Y$  such that:
  - ▶ The proportion of pairs in  $X$  that are hostile is at most  $\delta$
  - ▶ The proportion of pairs in  $Y$  that are hostile is at most  $\delta$
  - ▶ The proportion of edges between  $X$  &  $Y$  that are friendly is at most  $\delta$

# Proof

Can we modify the Proof of the Harary Balance Theorem? What problems could there be?

# Proof

Can we modify the Proof of the Harary Balance Theorem? What problems could there be?

The proof from lecture relies on dividing the graph based on the friends & enemies of a node  $A$ . Intuitively, we want to choose a node that is involved in a small number of violations. How do we formalize this intuition?

## Proof

Let's look at the number of violations that a node is involved in. We know that the proportion of violating triangles out of all triangles is at most  $\epsilon$ . Let  $N := |V|$ . How many triangles do we have?

## Proof

Let's look at the number of violations that a node is involved in. We know that the proportion of violating triangles out of all triangles is at most  $\epsilon$ . Let  $N := |V|$ . How many triangles do we have?

$${}_N C_3 = \binom{N}{3} = \frac{N(N-1)(N-2)}{6}$$

Therefore we have at most  $\epsilon \frac{N(N-1)(N-2)}{6}$  violating triangles.



## Proof

Let  $\text{viol} : V \rightarrow \mathbb{N}$  count the number of violations that a node is involved in. Each triangle will be counted 3 times, therefore:

$$\sum_{v \in V} \text{viol}(v) \leq 3 \times \epsilon \frac{N(N-1)(N-2)}{6}$$

## Proof

Let  $\text{viol} : V \rightarrow \mathbb{N}$  count the number of violations that a node is involved in. Each triangle will be counted 3 times, therefore:

$$\sum_{v \in V} \text{viol}(v) \leq 3 \times \epsilon \frac{N(N-1)(N-2)}{6}$$

Note that we immediately know that  $\exists A \in V$  such that:

$$\text{viol}(A) \leq 3 \times \epsilon \frac{N(N-1)(N-2)}{6} / |V| = \epsilon \frac{(N-1)(N-2)}{2} \leq \epsilon N^2 / 2$$

Let's use this node  $A$  to partition  $V$  into the sets  $X$  ( $A$ 's friends) and  $Y$  ( $A$ 's enemies).

## Case 1: $|X| \geq (1 - \delta)N$

Note  $\delta = \sqrt[3]{\epsilon} < \sqrt[3]{\frac{1}{8}} = \frac{1}{2}$ , therefore  $|X| > \frac{1}{2}N$ .

Assuming  $N$  is even then  $|X| \geq \frac{1}{2}N + 1$ .

## Case 1: $|X| \geq (1 - \delta)N$

Note  $\delta = \sqrt[3]{\epsilon} < \sqrt[3]{\frac{1}{8}} = \frac{1}{2}$ , therefore  $|X| > \frac{1}{2}N$ .

Assuming  $N$  is even then  $|X| \geq \frac{1}{2}N + 1$ .

Therefore, the number of edges between nodes in  $X$  is

$$\binom{|X|}{2} \geq \binom{\frac{1}{2}N+1}{2} = \frac{(\frac{1}{2}N + 1)(\frac{1}{2}N)}{2} \geq \frac{(\frac{1}{2}N)^2}{2} = N^2/8$$

## Case 1: $|X| \geq (1 - \delta)N$

Note  $\delta = \sqrt[3]{\epsilon} < \sqrt[3]{\frac{1}{8}} = \frac{1}{2}$ , therefore  $|X| > \frac{1}{2}N$ .

Assuming  $N$  is even then  $|X| \geq \frac{1}{2}N + 1$ .

Therefore, the number of edges between nodes in  $X$  is

$$\binom{|X|}{2} \geq \binom{\frac{1}{2}N+1}{2} = \frac{(\frac{1}{2}N + 1)(\frac{1}{2}N)}{2} \geq \frac{(\frac{1}{2}N)^2}{2} = N^2/8$$

We know that  $\text{viol}(A) \leq \epsilon N^2/2$ , and it's clear that any negative edge between nodes in  $X$  cause a violated triangle with  $A$ .

## Case 1: $|X| \geq (1 - \delta)N$

Note  $\delta = \sqrt[3]{\epsilon} < \sqrt[3]{\frac{1}{8}} = \frac{1}{2}$ , therefore  $|X| > \frac{1}{2}N$ .

Assuming  $N$  is even then  $|X| \geq \frac{1}{2}N + 1$ .

Therefore, the number of edges between nodes in  $X$  is  $\binom{|X|}{2} \geq \binom{\frac{1}{2}N+1}{2} = \frac{(\frac{1}{2}N + 1)(\frac{1}{2}N)}{2} \geq \frac{(\frac{1}{2}N)^2}{2} = N^2/8$

We know that  $\text{viol}(A) \leq \epsilon N^2/2$ , and it's clear that any negative edge between nodes in  $X$  cause a violated triangle with  $A$ .

Therefore, the proportion of edges between nodes in  $X$  causing a violation is at most:

$$\frac{\text{viol}(A)}{\binom{|X|}{2}} \leq \frac{\epsilon N^2/2}{N^2/8} = 4\epsilon = 4\delta^3 < \delta$$

Therefore we satisfy the theorem. Note that the final inequality holds as  $\delta < \frac{1}{2}$ .

**Case 2:**  $|Y| \geq (1 - \delta)N$

The same proof as Case 1 applies.

### Case 3: $|X| < (1 - \delta)N$ and $|Y| < (1 - \delta)N$

As  $X$  and  $Y$  partition  $V$ , we know that  $|X| + |Y| = |V| = N$ , therefore  $|X| = N - |Y| > \delta N$ .

Assuming  $\delta N \in \mathbb{N}$  then the number of edges within  $X$  is:

$$\binom{|X|}{2} \geq \binom{\delta N + 1}{2} > \delta^2 N^2 / 2$$



### Case 3: $|X| < (1 - \delta)N$ and $|Y| < (1 - \delta)N$

As  $X$  and  $Y$  partition  $V$ , we know that  $|X| + |Y| = |V| = N$ , therefore  $|X| = N - |Y| > \delta N$ .

Assuming  $\delta N \in \mathbb{N}$  then the number of edges within  $X$  is:

$$\binom{|X|}{2} \geq \binom{\delta N + 1}{2} > \delta^2 N^2 / 2$$

As we showed before, any unfriendly edge within  $X$  will cause an unstable triangle with  $A$ , and  $\text{viol}(A) \leq \epsilon N^2 / 2$ , therefore the proportion of unfriendly edges in  $X$  is at most:

$$\frac{\text{viol}(A)}{\binom{|X|}{2}} \leq \frac{\epsilon N^2 / 2}{\delta^2 N^2 / 2} = \delta$$

The same argument holds for unfriendly edges in  $Y$ . All that remains is to show the same bound on the proportion of edges between  $X$  and  $Y$  that are friendly.

### Case 3: $|X| < (1 - \delta)N$ and $|Y| < (1 - \delta)N$ cont'd

The number of edges between  $X$  and  $Y$  is  $|X||Y| = |X|(N - |X|)$  for  $\delta N < |X| < (1 - \delta)N$ .

As this is a concave quadratic function maximized at  $|X| = 0.5N$  and  $\delta < 0.5$ , it's clear  $|X|(N - |X|) > \delta(1 - \delta)N^2 > \delta(1 - 0.5)N^2 = \delta N^2/2$ .

### Case 3: $|X| < (1 - \delta)N$ and $|Y| < (1 - \delta)N$ cont'd

The number of edges between  $X$  and  $Y$  is  $|X||Y| = |X|(N - |X|)$  for  $\delta N < |X| < (1 - \delta)N$ .

As this is a concave quadratic function maximized at  $|X| = 0.5N$  and  $\delta < 0.5$ , it's clear  $|X|(N - |X|) > \delta(1 - \delta)N^2 > \delta(1 - 0.5)N^2 = \delta N^2/2$ .

Therefore, using the same bound on violations as before we can see that the proportion of edges between  $X$  and  $Y$  that are friendly is at most:

$$\frac{\text{viol}(A)}{|X||Y|} \leq \frac{\epsilon N^2/2}{\delta N^2/2} = \delta^2 < \delta$$

Therefore  $X$  and  $Y$  satisfy the theorem.

# Quercus Quiz