
Dynamic and Social Network Analysis

Lecture 7

Miray Kas

Bilkent University

Computer Engineering Department

Network Formation & Evolution

Network Formation & Evolution

- Real networks evolve!
- A network forming from scratch is an evolution too!
- At some point, they may grow, shrink, or dissolve completely!

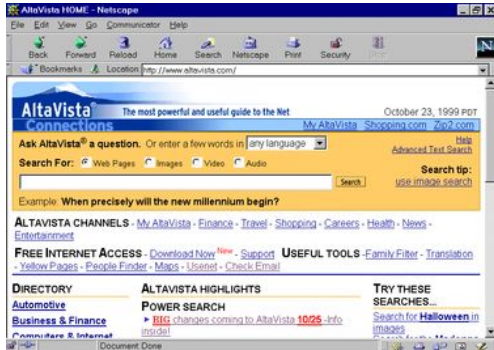
If Barabasi-Albert Preferential Attachment Model were to hold for all, then...

- How and why do trends and popularity change?

Google was a late comer!

Google founded in 1998

6 years after the earlier search engines were there

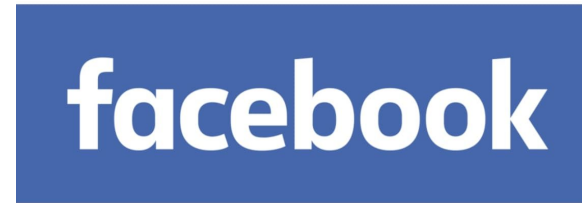


AltaVista

Founded in 1995, it was #1
Lost ground to Google
Shut down by Yahoo! In 2013

The Google logo, consisting of the word 'Google' in its characteristic multi-colored font (blue, red, yellow, blue, green, red).

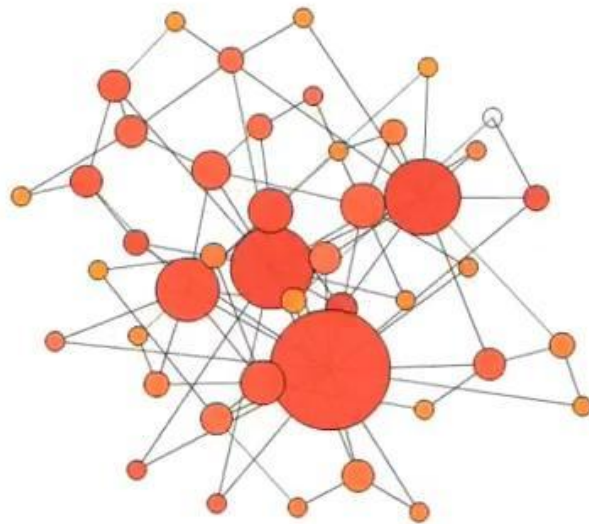
Founded in 1998
Became the most connected
node of WWW by 2000



Founded in 2004
Became the most connected
node of WWW by 2011

Networks Evolve Over Time

- Real networks evolve over time!
 - Dynamic Network Analysis
 - Temporal Network Analysis
 - Time-varying Network Analysis
 - Over-time Network Analysis



Temporal Networks:

- Have another dimension apart from nodes and edges: **Time!**
- Representations of over time networks include
 - Incremental updates
 - Series of full snapshots

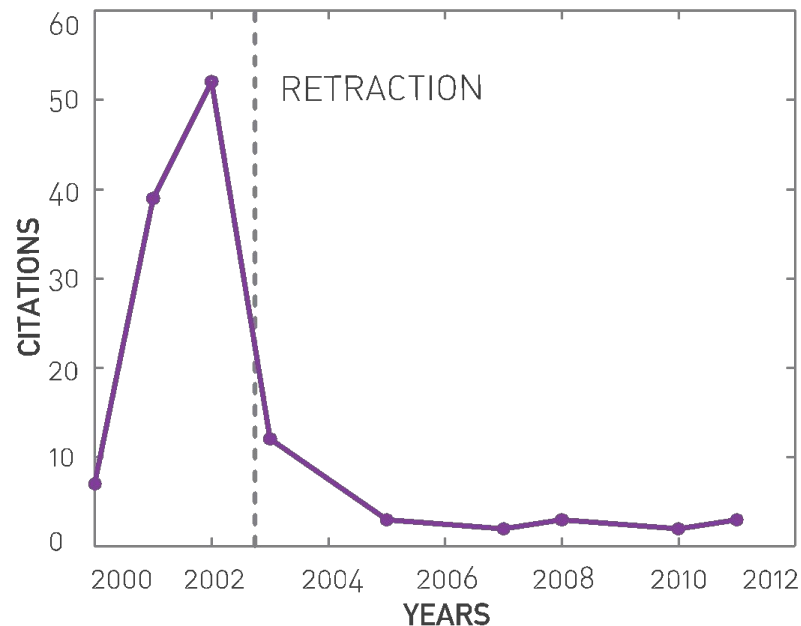
Mechanisms Networks Evolve Through (1)

- **Addition of nodes (E.g. new members joining)**
 - Starting a school
 - After getting a new job, you join the company's network
 - A new paper getting published in citation network
- **Removal of nodes (E.g. Members exiting the network)**
 - Graduation removes you from active students' network
 - Retirement
 - Some archival networks do not support node removal (e.g. citation networks).

The Impossibility of Node Deletion in Citation Networks

● Schön Scandal

- Publishing in Science and Nature
- Listed as a co-author of a new article every 8 days
- Semiconductor breakthroughs which were later debunked to be achieved with falsified data
 - Bell Labs where he worked started formal investigations
 - University of Konstanz revoked his PhD Degree, and called it the biggest fraud in Physics in 50 years



[23] J.H. Schön, Ch. Kloc, R.C. Haddon, and B. Batlogg. A superconducting field-effect switch. *Science*, 288: 656–8. 2000.

Mechanisms Networks Evolve Through (2)

- **Addition of edges (E.g. new relationships forming)**

- In School x Student network, a new edge is added when a student starts a new school
- Becoming colleagues with people on the same team
- A new paper citing existing papers

- **Removal of edges (E.g. new relationships forming)**

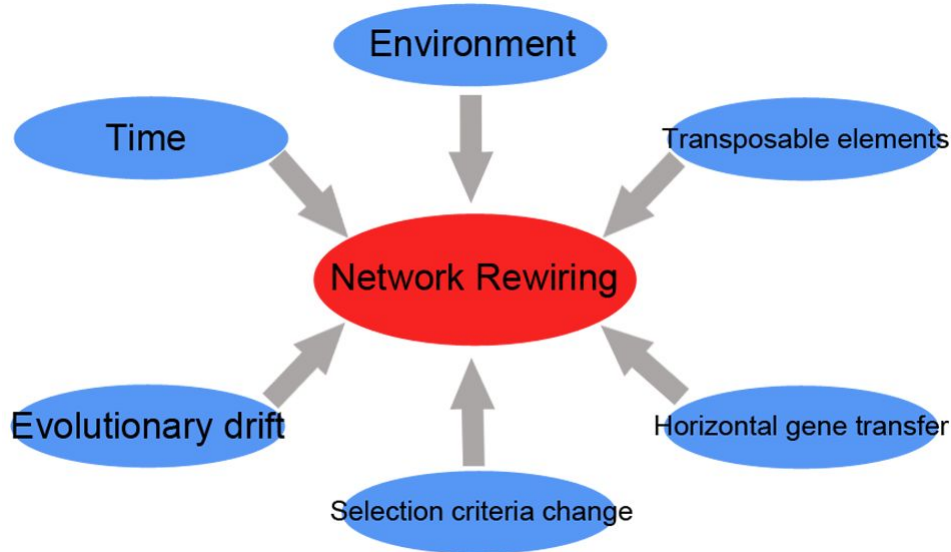
- Closing your account finishes your affiliation with a bank.

Mechanisms Networks Evolve Through (3)

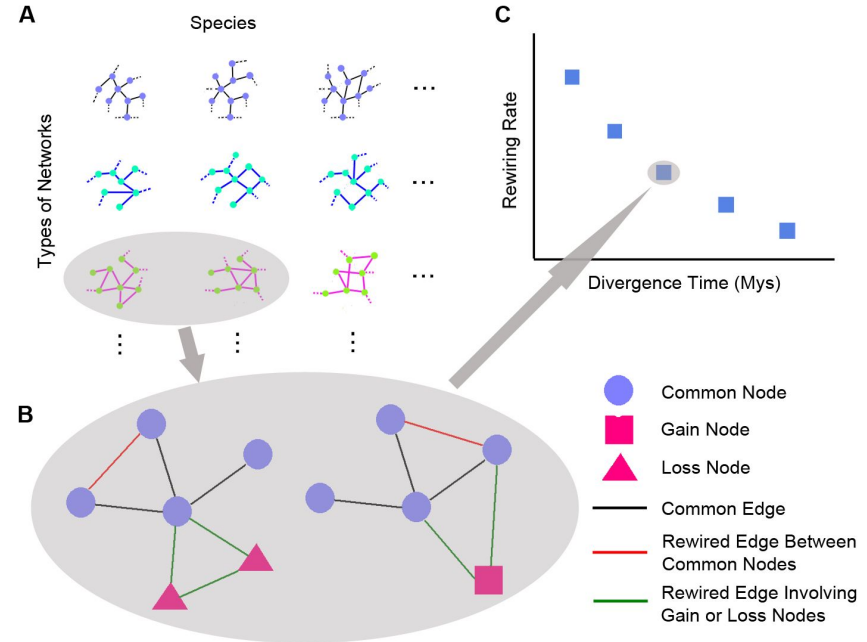
- **Rewiring of links ($A \rightarrow B$ becomes $A \rightarrow C$)**
 - Can be represented as two separate events back to back:
 - Removal of one edge ($A \rightarrow B$)
 - Addition of another edge ($A \rightarrow C$)
 - A student changing advisors
 - Changing jobs

Evolutionary Rewiring of Biological Networks

Factors affecting rewiring



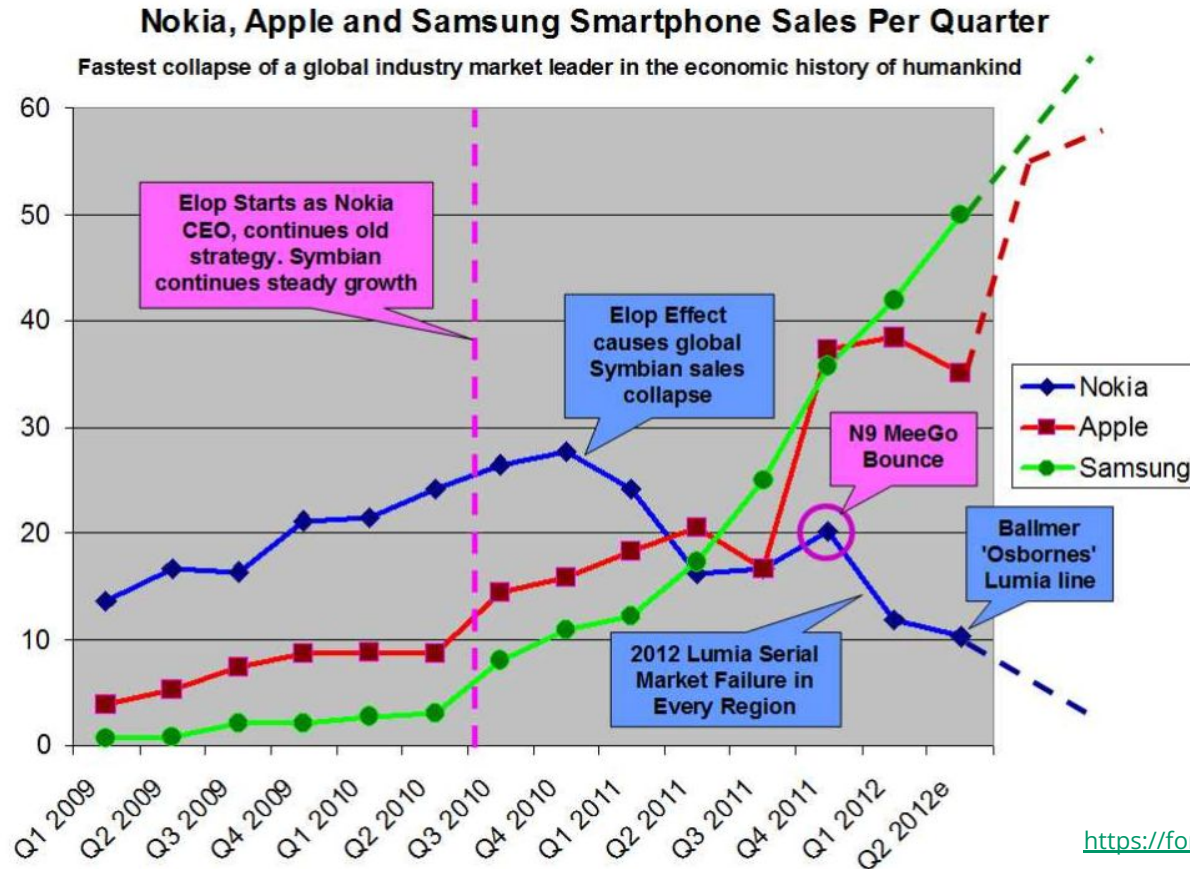
Measuring rewiring



Mechanisms Networks Evolve Through (4)

- **Changes in the link weights (E.g. The level of importance or the intensity of relationship changes)**
 - More collaboration can happen between authors (Number of papers published increase)
 - Alliances can turn into conflicts (+1 weight can become -1)
 - Trade value between two countries can be different every year.

Rise of Smartphones, Fall of Nokia



Imagine a 2-mode network:
(Phone model x Consumer)

It is so big, and it is evolving so fast!

Sales people want to avoid returned phones

- **Sales reps quickly learn which phones get returned**
 - An unhappy customer takes a lot of time and energy
 - You have to payback the commission you made last week - *Smaller paycheck for you this week*
- **Self protection in action!**
 - Sales reps will quickly adapt to push for other phones
 - Frequently returned brand/model will lose popularity even more quickly
 - The Lumia line of smartphones? *The highest return rate for any Nokia smartphone, ever.*

(Further reading for the curious:

<https://seekingalpha.com/article/916271-how-stephen-elop-destroyed-nokia>)

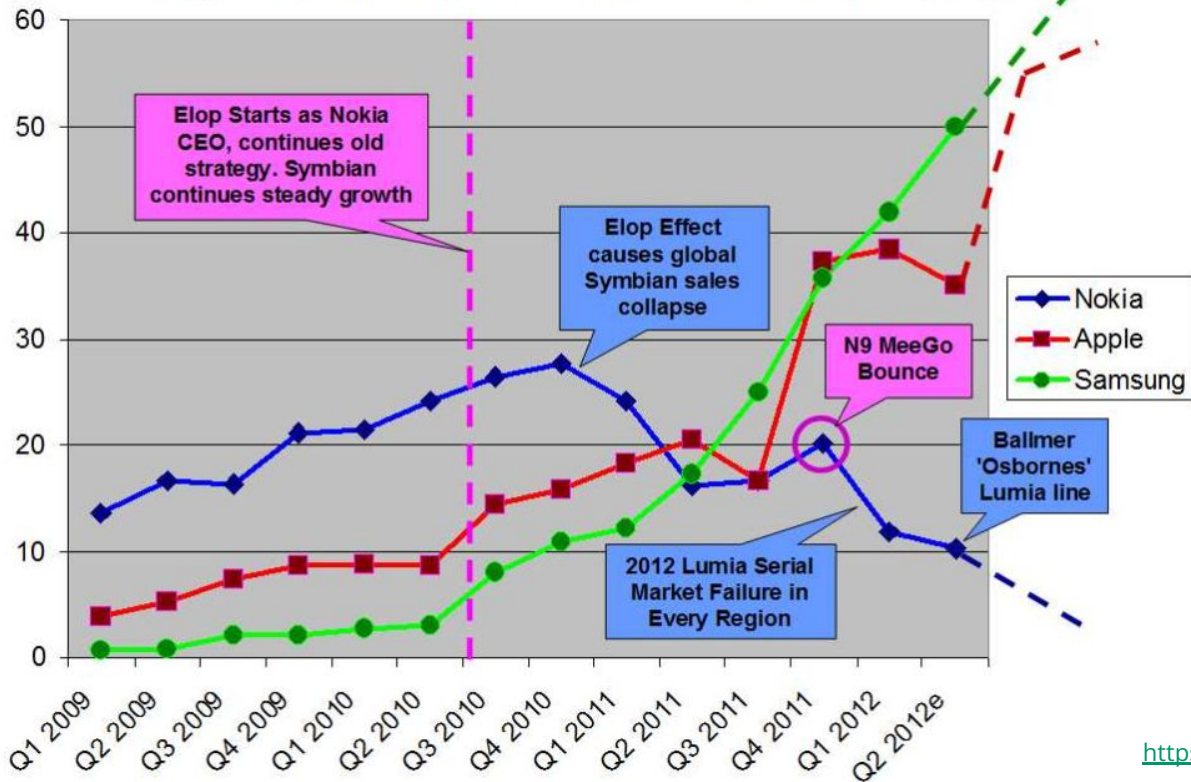
Let's remember Lumia again!

NOKIA LUMIA



Nokia, Apple and Samsung Smartphone Sales Per Quarter

Fastest collapse of a global industry market leader in the economic history of humankind



Imagine a 2-mode network:
(Phone model x Consumer)

It is so big, and it is evolving so fast!

Secondary Mechanisms that Impact Network Evolution (1)

- **Initial Attractiveness**

- If only rich gets richer, how do you even start building connection or get popular?
- Finding a niche is critically important

- **Changes in Node Attributes**

- New constraints added (E.g. A doctor who is overbooked cannot accept new patients)
- Losing attributes that were attractions (E.g. A company losing its designer cannot attract customers anymore)
- Resistance

Secondary Mechanisms that Impact Network Evolution (2)

- **Node Constraints**

- Capacity saturation (E.g. A doctor who is overbooked cannot accept new patients)
- Blacklisting (E.g. Person-X cannot send emails to a group due to prior spamming)
- Feature based limitations (E.g. You have to be 55+ to buy property from senior living facility)

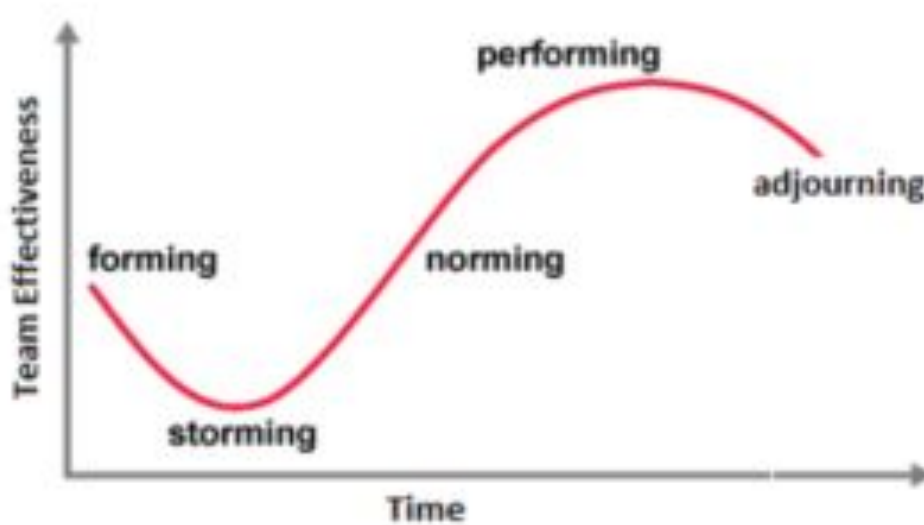
- **Edge Constraints**

- Forbidding relationships between two parts of the organization (E.g. Goldman Sachs internal Chinese Wall - [info for the curious](#))

Secondary Mechanisms that Impact Network Evolution

- **Time-Related Constraints and Updates (e.g. expirations)**
 - Aging
 - Renewal
 - Accelerated Growth
- **Changes in Objective Functions**
 - Changes in goals will change which links are likely to form and which nodes are attractive

All these changes take networks through phases



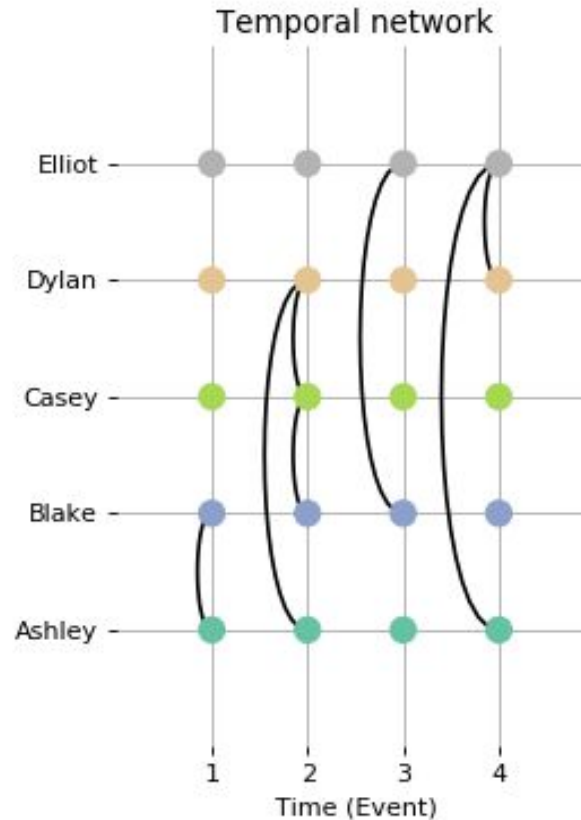
[Tuckman's Article](#) (1965)

- A team is a small network of people
- Bruce Tuckman proposed this model in 1965
- **Fun Fact:** Bruce Tuckman was in the field of educational psychology and studied Procrastination of college students and came up with an index for it!

Temporal Network Data: Adding a Time Dimension

Tuple
 (i, j)

For undirected,
unweighted
static
networks



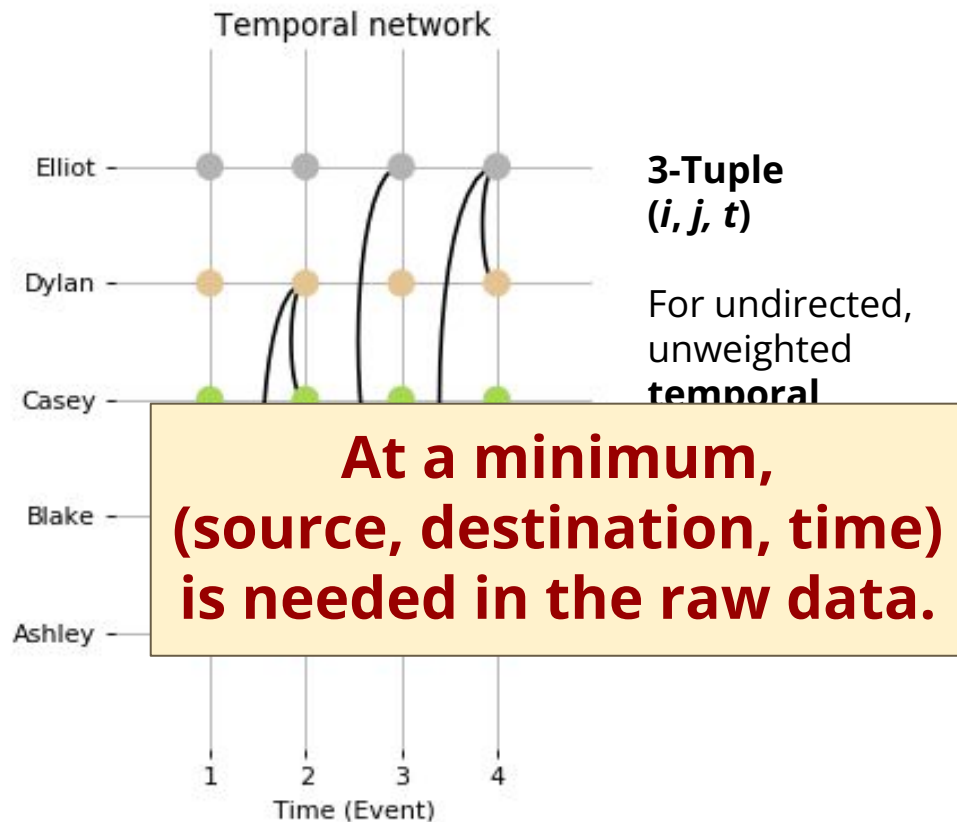
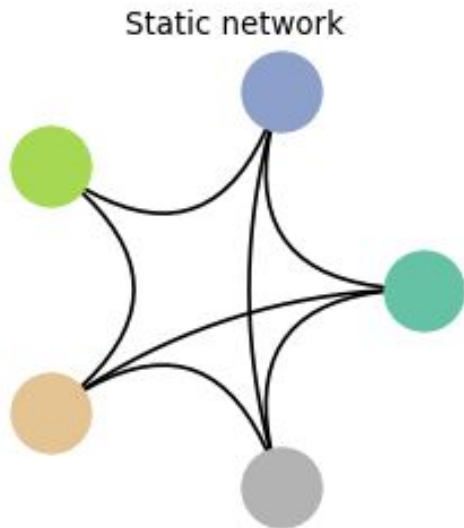
3-Tuple
 (i, j, t)

For undirected,
unweighted
temporal
networks

Temporal Network Data: Adding a Time Dimension

Tuple
 (i, j)

For undirected,
unweighted
static
networks



Temporal Network Data: Adding a Time Dimension

- **Incremental updates (Finer Granularity)**

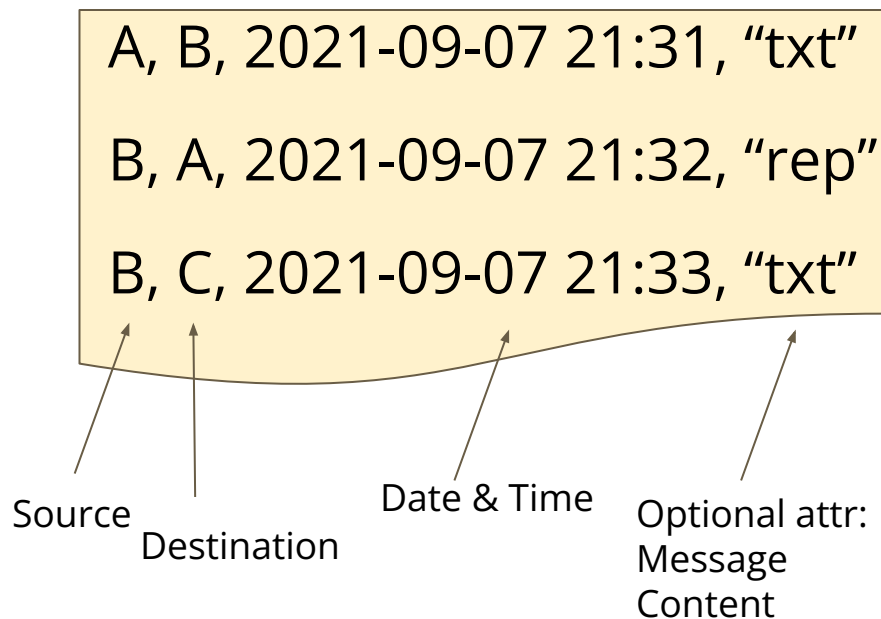
- **Example**: Chat messages/tweets come with full timestamp
- Access to more precise ordering/sequencing of events
- By aggregation, incremental updates can be turned into full snapshots.

- **Full snapshots (Coarser Granularity)**

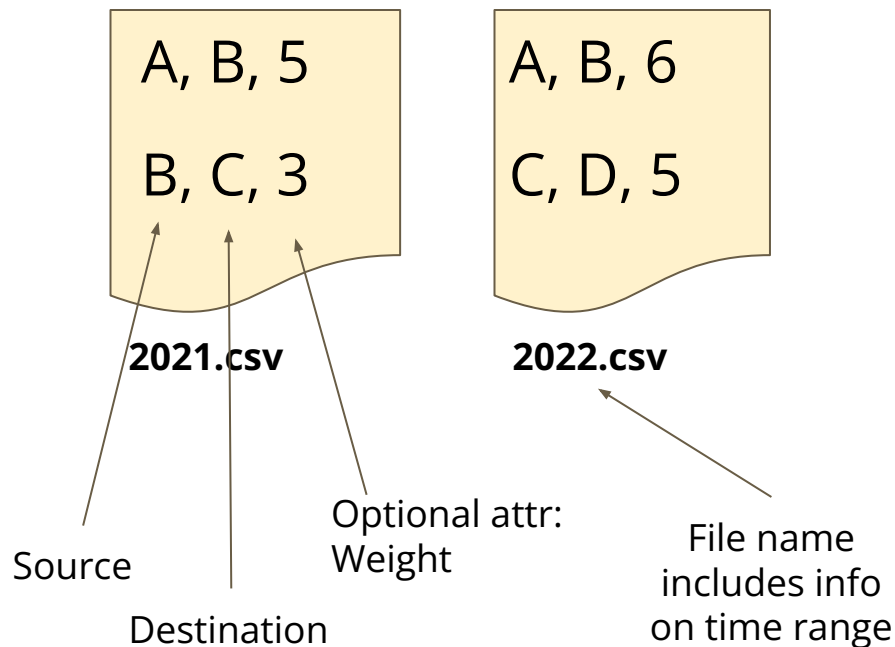
- **Example**: Citation networks are usually represented in yearly snapshots
- Incremental updates can be grouped into snapshots
- Details on the individual events' timing may be lost

Temporal Network Data Representation

Incremental updates



Full snapshots



Temporal Network Data Representation: Single File

- Most tools support converting data from n-tuples
- Good for processing incremental data all at once or in fine grain detail
- One row per edge activity in the form of n-tuples. In each row:
 - Source
 - Destination
 - Time
 - Other optional attributes (including but not limited to)
 - Weight
 - Event type

Temporal Network Data Representation: Multiple Files

- Full snapshots benefit from use of multiple files
- One file per snapshot
 - Example: One file per year in multi-year data.
- Data might be readily in the form of multiple files
- May process (aggregate/split) and arrive at multiple files
- **Tip:** *Using a consistent file name convention and including the time range information in the file names is a life saver!*

Temporal Network Data Representation

A, B, 2021-09-07 21:31, "txt"

B, A, 2021-09-07 21:32, "rep"

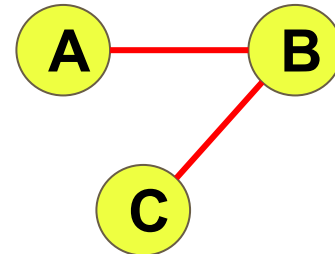
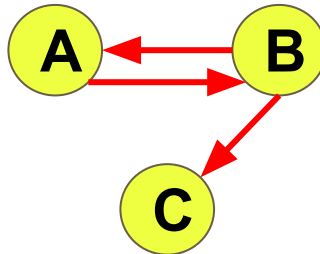
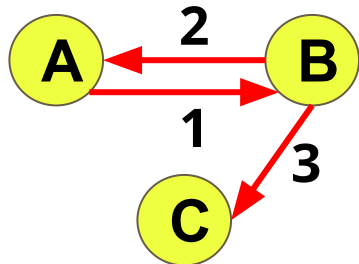
B, C, 2021-09-07 21:33, "txt"

A shares some information with B

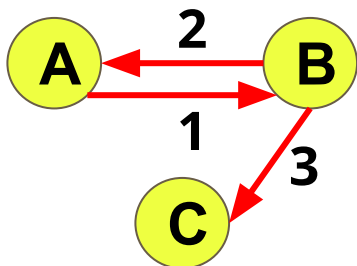
B replies back to A

B relays this information to C

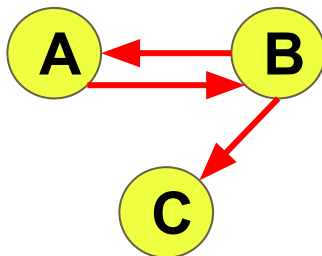
We can tell that A started the gossip chain, B relays this information



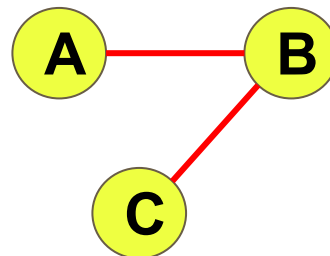
Temporal to Static Information Loss



We can tell that A started the gossip chain, then B relays this information



We **cannot** tell if B relays the info from A to C. Or if B is the conversation starter telling both A and C at the same time.



We **cannot** tell if it is A or B or C that is the conversation starter.

Aggregating Temporal Data

- Full snapshots benefit from aggregation & use of multiple files
 - Data might be readily in the form of multiple files
 - May process (aggregate/split) and arrive at multiple files
- One raw data file per snapshot (e.g. yearly, monthly, daily, etc)
 - Combine (usually summation) data from days in a month to arrive at monthly data
 - Remove the date information, keep month/year only

Aggregate & Split: Example of Oversimplification

A, B, 2021-09-07 21:31, "txt"
B, A, 2021-09-07 21:32, "rep"
B, C, 2021-09-08 21:33, "txt"
B, A, 2021-09-08 21:32, "rep"
B, C, 2021-09-08 21:33, "txt"

texts_2021_09.csv

- *Convert from directed to undirected*
- *Drop actual content & time*
- *Combine interaction counts into a weight column*
- *Process each day separately for trends*

A, B, 2

texts_2021_09_07.csv

B, C, 2

B, A, 1

texts_2021_09_08.csv

Aggregation can come with over-simplification which is sometimes useful for finding out key strengths.

Representing Different Event Types

- Networks like citation networks are ever growing
 - They don't have delete or modify; they are add-only.
- Non-archival networks have people joining/leaving groups

**What if you need to
represent
delete or modify events?**

Representing Different Event Types: Delete

- **Explicit Representation**

- In the raw data, the event type is called out explicitly
- **Example:** Add a new column in your n-tuple file to say **A**(dd) ,**D**(elete), **M**(odify)

- **Implicit Representation**

- Works for multi-file full snapshot representations.
- **Assumption:** Each snapshot contains complete network, not only the newest updates.
- **Implicit Assumption:** If X is not in the snapshot, it is deleted.

Representing Different Event Types: Modifies

- Similar arguments hold for modify events
- If modification is on the weight, the weight update could be represented as the delta or the new value

Next Week:

Dynamic Network Analysis

