Network Theory and Dynamic Systems

10. Dynamics II

SOSE 2025

Dr. -Ing. Stefania Zourlidou

Institute for Web Science and Technologies
Universität Koblenz

Outline

- Ideas, Information, Influence
- Epidemic Spreading
  - Diseases
  - Rumor
- Opinion
- Search
  - Local Search
  - Searchability

Dynamics I (previous Lecture)

Dynamics II (Today Lecture!)

1. Opinion Dynamics

Opinion Dynamics (1/2)

- **Opinion dynamics:** The scientific study of how opinions form, evolve, and spread within society (social networks)
- Opinions change via interactions on social networks (friends, followers)
- Opinions can be:
  - **Discrete**: Categorical choices (e.g., "yes/no," specific candidate)
  - **Continuous**: A spectrum of beliefs (e.g., 0–1 rating, degree of agreement)

Opinion Dynamics (2/2)

Two Main Model Types

- **Discrete Opinion Dynamics**
  - Opinions are distinct, integer values
  - Example: A group decides where to eat; individuals adopt the most popular restaurant choice among their friends
- **Continuous Opinion Dynamics**
  - Opinions are real numbers (degrees of belief)
  - Example: People only discuss politics with those holding "similar enough" views, potentially leading to polarization or consensus
- **Why it is important:** Helps us understand social phenomena like *misinformation*, *political polarization*, and *collective decision-making*

Discrete Opinions (1/3)

- Definition: Dynamics where opinions are limited to a finite, usually small, number of distinct choices (e.g., right/left, buy/sell, agree/disagree)
- Opinions are typically represented by **integers**
- For simplicity, we often focus on **binary** opinions, where only two states are possible (e.g., 0 or 1, $\{-1,1\}$)
- Model: A discrete opinion model is defined by the specific **rules** governing how an individual's (node's) opinion updates based on the opinions of their connected neighbors

Discrete Opinions (2/3)

- **Initial Configuration**
  - Opinions are typically assigned randomly to the nodes (individuals) within the network
  - This usually implies an initial state of near-equal distribution of opinions, representing maximal *disagreement* or high heterogeneity
- **Opinion Update Process**
  - The defined opinion update rule is applied iteratively to all nodes in the network
  - Each iteration involves scanning through all nodes
  - **Update Mechanism**: Nodes are typically updated **asynchronously** (one at a time) and in random order to prevent biases and facilitate stable convergence to a final opinion state

Discrete Opinions (3/3)

- Two Primary Outcomes:
  - **Steady State (Equilibrium):** The system reaches a configuration where no node (individual) changes its opinion in subsequent iterations
    - This stable state can be
      - Consensus: All nodes eventually adopt the same single opinion
      - Polarization: The network splits into distinct subgroups, each holding a different, fixed opinion (e.g., in binary opinion models, two opposing factions emerge)
  - **Non-Stationary State (Oscillation/Chaos):** The system fails to reach a global steady state, with some or all nodes continually changing their opinions across iterations
    - Even in non-stationary states, some global properties of the opinion configuration, such as the average opinion or other statistical moments, may exhibit long-term stabilization or periodic behavior

Variables: Average Opinion

- The **average opinion** is the **arithmetic mean of all individual opinions across the network's nodes**
- When beginning with a random distribution of two binary opinions (e.g., 0 and 1), the initial average opinion will typically be approximately 0.5, reflecting an equal proportion of each opinion
- The average opinion usually changes throughout the simulation, and its value is often tracked after each iteration to monitor the system's evolution
- If the system converges to a stationary (steady) state, the average opinion will also converge to a stable, precise value
    - In the case of consensus, this converged average will be either 0 or 1, depending on which opinion became globally dominant

Variables: Exit Probability

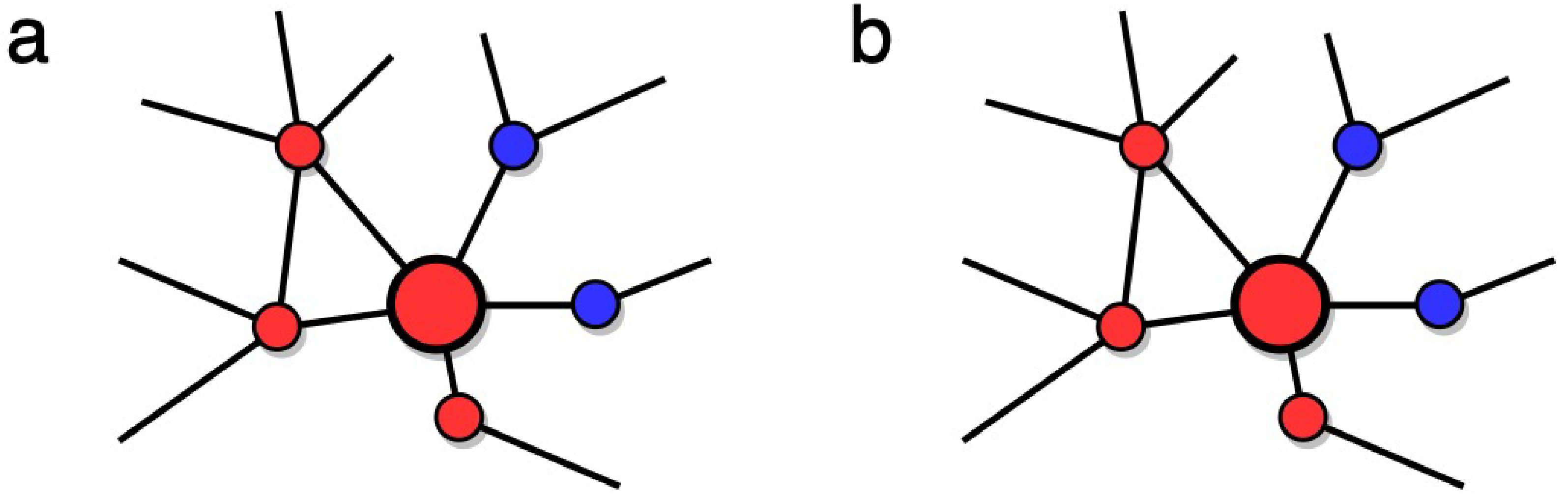
- The **exit probability** quantifies the likelihood of the network reaching consensus on a specific opinion (e.g., opinion '1'), as a function of the initial fraction of nodes holding that opinion.
- Example:
  - **Simulation Setup**: The model dynamics are executed multiple times (e.g., 100 runs), each starting from a distinct random initial configuration
  - **Initial Condition**: For each run, opinion '1' is assigned to every node with a given probability, e.g., P(opinion 1)=0.4, resulting in approximately 40% of nodes initially holding opinion '1'
  - **Calculation**: If all 100 runs converge to a consensus state, and 30 of these runs achieve consensus on opinion '1', then the exit probability for an initial P(opinion 1)=0.4 is 30/100=0.3

Discrete Opinion Dynamics Models

- Two models
  - The majority model
  - The voter model

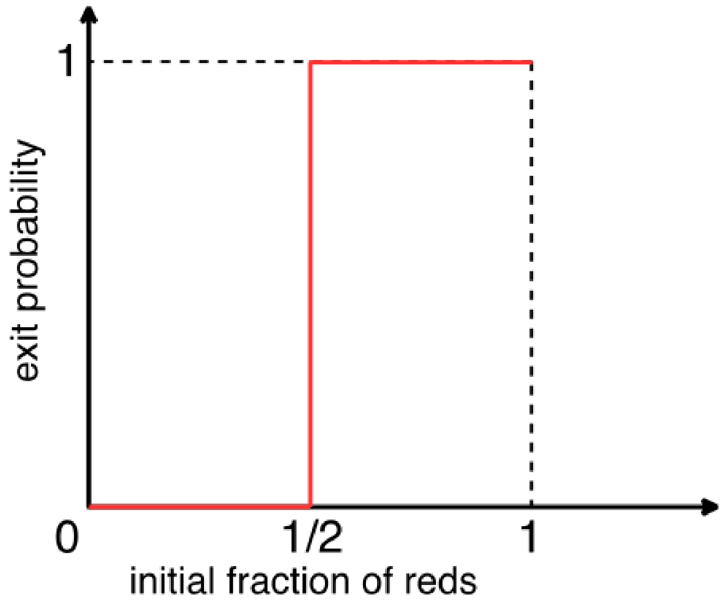
Majority Model (1/2)

- **Majority rule:** each node adopts the opinion of the majority of its neighbors
- If the number of neighbors is even (2, 4, etc.) and there is an equal number of them with either opinion, then we flip a coin to decide which opinion will be taken by the node
- Equivalent to fractional threshold model of information diffusion, with threshold 1/2



Majority Model (2/2)

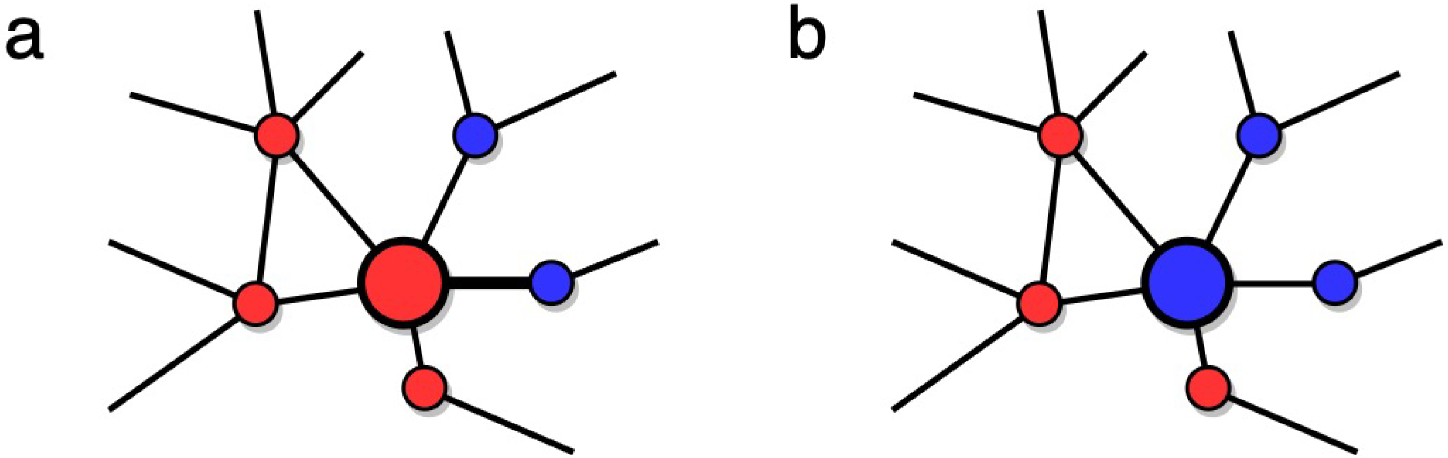
- **Stable States (Equilibria)**
  - **Consensus:** All nodes eventually adopt the same single opinion
  - **Coexistence** (Polarization): Different opinions persist stably in the network, with each node holding the majority opinion within its local neighborhood
- **Consensus in Majority Model**
  - On most complex networks (e.g., random networks, scale-free networks), the Majority Model typically does not reach full consensus, often resulting in a coexistence of opinions
  - However, consensus can be achieved on highly structured topologies such as **one-** and **two-dimensional grids**



- **Exit Probability Profile**
  - For the runs that do lead to consensus, the exit probability (likelihood of reaching consensus on a specific opinion) typically exhibits a step-like profile as a function of the initial fraction of that opinion (see figure)

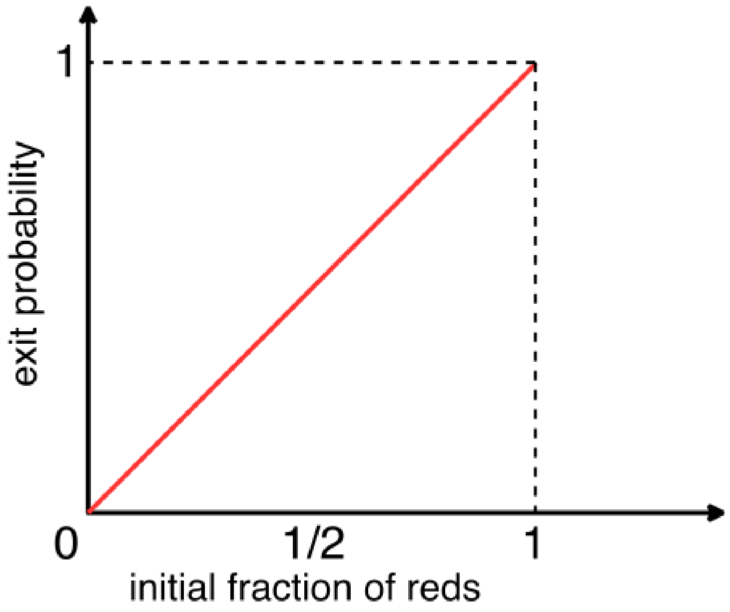
Voter Model (1/2)

- **Update Rule:** At each step, a randomly chosen node adopts the opinion of one of its randomly chosen neighbors
- **Consensus** (a state where all nodes share the same opinion) is the sole stable state of the Voter Model and is reached regardless of the underlying network topology
- **Convergence**: In any state where diverse opinions coexist, interactions between neighbors holding different opinions will always occur, facilitating the propagation of opinions until a single opinion dominates



Voter Model (2/2)

- **Exit Probability:** For the Voter Model, the probability of reaching consensus on a specific opinion (e.g., opinion '1') is precisely equal to the initial fraction of nodes holding that opinion
  - This is often referred to as a **diagonal function**
- **Consensus:** Even when starting from an initial configuration with a majority of one opinion (e.g., opinion '1'), the system can still converge to consensus on the minority opinion (e.g., opinion '0') due to the stochastic nature of the update process
- **Example:** If 30% of the nodes initially hold opinion '1', then we expect that in 30% of independent simulation runs, the entire network will ultimately reach consensus on opinion '1'



Voter Model: Variants

- Several extensions and modifications to the basic Voter Model have been developed to incorporate more complex social dynamics
  - **Zealots:** Introduction of nodes that never change their opinion
    - If all zealots share the same opinion, they can drive the system towards consensus on that opinion
    - If zealots hold different opinions, full consensus is typically prevented, leading to coexistence
  - **Multi-State Opinions with Bounded Confidence:** Models with more than two opinion states, where interactions are limited to nodes whose opinions are sufficiently "close"
    - Example: Three opinions (1,2,3), where only neighboring opinions can interact (e.g., 1 with 2, and 2 with 3, but not 1 with 3)
  - **Spontaneous Opinion Changes (Noise):** Introduction of a probability for nodes to spontaneously change their opinion at each iteration, independent of neighbor influence
    - This adds stochastic "noise" to the system dynamics

Continuous Opinions (1/2)

- Dynamics where opinions are represented by **real** numbers, allowing them to vary smoothly along a continuous spectrum of possible values
  - **Example**: Political alignment on a scale from very progressive (e.g., -1) to very conservative (e.g., $+1$)
- **Initial Configuration:** Opinions are typically assigned *randomly* from a uniform distribution within the predefined range of possible values
- **Stopping Criterion:** Simulations are usually terminated when the maximum change in any individual opinion between successive iterations falls below a small, predefined threshold
  - This signifies that the system has reached a stationary state (equilibrium)

Continuous Opinions (2/2)

- **Possible Stationary States**: Depending on the model and parameters, the system can converge to
  - **Consensus**: All opinions converge to a single, shared value
  - **Polarization**: Opinions cluster around two distinct, opposing values
  - **Fragmentation**: Opinions disperse and cluster around multiple distinct values, forming several factions

Bounded Confidence Model (1/4)

- **Core Principle**
  - Interactions between two individuals (nodes) result in opinion influence only if the absolute difference between their current opinions is less than or equal to a predefined value
    - This value is known as the confidence bound, or tolerance (often denoted as ϵ)
    - This implies that individuals only engage with, or are influenced by, those whose opinions are "close enough" to their own
- **Model Parameters**
  - The confidence bound ϵ (epsilon): Defines the maximum opinion difference for interaction
  - The convergence parameter μ (mu): Determines the extent to which opinions are adjusted during an interaction (i.e., how much closer they move)

Bounded Confidence Model (2/4)

- Dynamics
  - At iteration *t*, each node *i* has opinion $o_i(t)$, which is a real number between, say, zero and one
  - An iteration consists of a sweep over all nodes, synchronously or in random order
  - At iteration *t*+1, for each node *i* we pick one neighbor *j* at random. If

$$|o_i(t) - o_j(t)| < \epsilon$$

  - the opinions of *i* and *j* are changed as

$$o_i(t+1) = o_i(t) + \mu[o_j(t) - o_i(t)]$$
$$o_j(t+1) = o_j(t) + \mu[o_i(t) - o_j(t)]$$

Bounded Confidence Model (3/4)

$$o_i(t+1) = o_i(t) + \mu[o_j(t) - o_i(t)]$$
$$o_j(t+1) = o_j(t) + \mu[o_i(t) - o_j(t)]$$

- If $\mu = 1/2$, the opinions converge to their average
- If $\mu = 1$, they switch! The parameter μ usually varies between 0 and 1/2
- If we sum the opinion update equations side by side and divide by two, we see that the average opinion of *i* and *j* is the same before and after the update —> **the average opinion of the population is preserved by the dynamics**
- **Consequence**
  - If the initial opinions are taken at random from the range [0, 1], their average is 1/2 (with possible small deviations)
  - So, **if the system eventually reaches consensus, the opinions of all nodes will cluster around 1/2**

Bounded Confidence Model (4/4)

- Starting from any random initial opinion configuration, the model dynamics are guaranteed to reach a stationary (equilibrium) state on any network topology
- **Role of Convergence Parameter (μ):** The convergence parameter μ solely influences the rate of convergence, specifically affecting the number of iterations required to reach a stationary state, but not the final outcome
- **Determinants of Final State:** The number and arrangement of opinion clusters in the stationary state fundamentally depend on:
  - The confidence bound ϵ
  - The structure of the underlying network
  - Inverse Relationship: A lower confidence bound (ϵ) generally leads to a larger number of distinct final opinion clusters (i.e., increased fragmentation or polarization)
- **Consensus Condition:** For confidence bounds $\epsilon > 1/2$, the system consistently reaches global consensus on any network, with the final opinions centered around 1/2 (assuming opinions are typically scaled from 0 to 1

Bounded Confidence Model: Variants

Several extensions and modifications to the basic Bounded Confidence Model have been developed to incorporate more realistic social phenomena

- **Heterogeneous Confidence Bounds (ϵ_i)**
  - Using individual-specific values for the confidence bound (ϵ_i) to reflect that individuals differ in their openness to persuasion
  - The confidence bound of a node can be dynamically linked to its current opinion value. For example, if an individual's opinion is close to the extremes of the opinion range, their confidence bound might be smaller, reflecting the higher resistance of "extremists" to persuasion compared to more moderate individuals
- **Spontaneous Opinion Changes (Noise)**
  - Introducing the possibility for individuals to change their opinion spontaneously, independent of direct social influence
  - Similar to other opinion dynamics models, this can be implemented by allowing nodes to randomly perturb their opinion with a certain probability at each iteration, introducing a stochastic "noise" component

Coevolution of Networks and Dynamics (1/7)

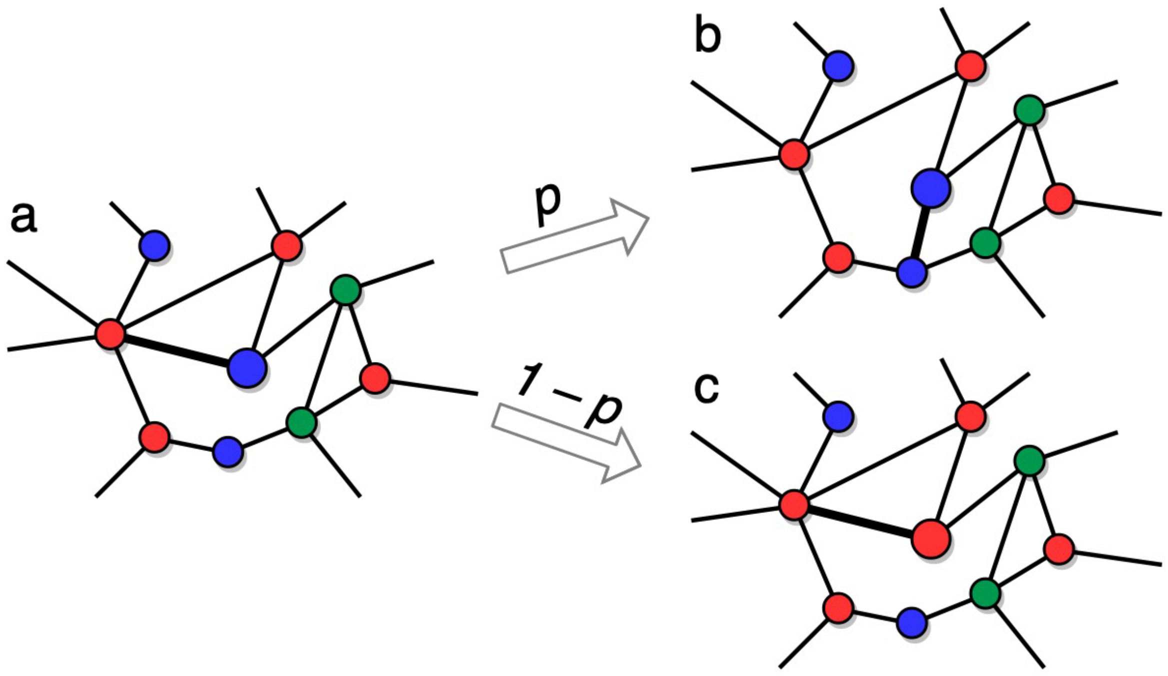
- **Network Assortativity:** A topological property where nodes (individuals) with similar characteristics or opinions tend to connect to each other more frequently than to dissimilar nodes
- **Mechanisms of Assortativity in Social Networks:** The observed assortativity in social networks is primarily driven by two interdependent mechanisms
  - **Selection (Homophily):** Individuals establish connections with others because they share similar attributes, opinions, or interests. ("birds of a feather flock together")
  - **Social Influence:** Individuals become more similar in their opinions or behaviors because they are connected and interact
    - This is the core mechanism studied in opinion dynamics, where network structure influences opinion spread

Coevolution of Networks and Dynamics (2/7)

- Limits of traditional opinion dynamics models
  - The network is fixed. Selection is not allowed, **nodes with very similar opinions do not have the option to become neighbors**, unless they already are
  - Neighbors with very dissimilar opinions cannot become disconnected
- **Coevolution models:** opinion changes may induce modifications in the network structure, which could in turn affect the opinions, and so on
- We consider a coevolution model with discrete opinions, initially randomly assigned to the nodes

Coevolution of Networks and Dynamics (3/7)

- **Dynamics**
  - Each iteration is a sweep over the nodes, synchronously or in random order
  - For each node *i* select a random neighbor *j* with different opinion from *i*:
  - With the probability *p*, the link between *i* and *j* is rewired from *i* to a randomly selected non-neighbor holding the same opinion as *i* **(selection)**
  - With probability 1 – *p*, *i* takes the opinion of *j* **(influence)**

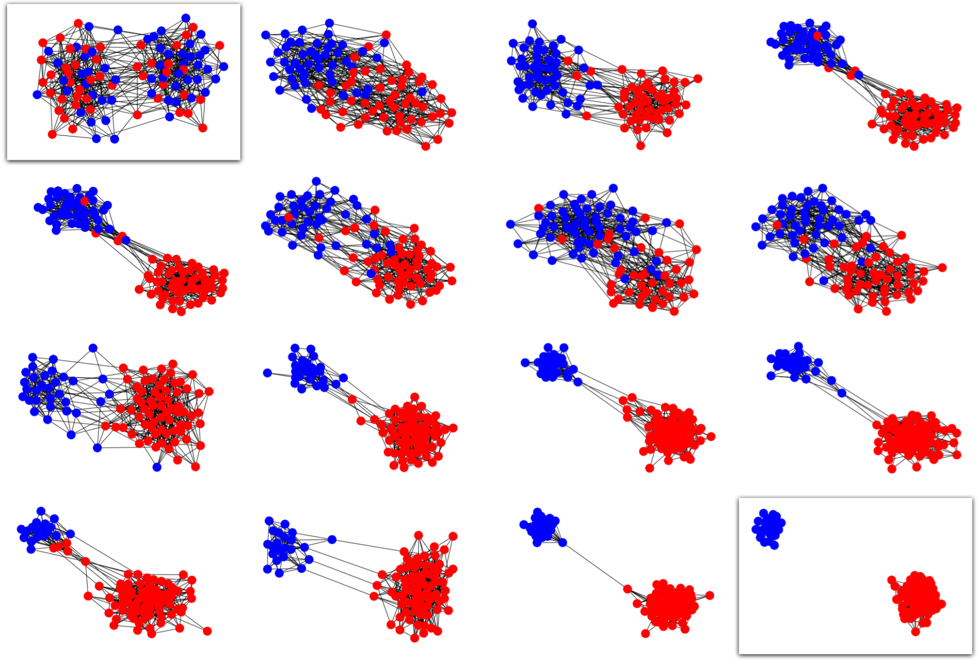


Coevolution of Networks and Dynamics (4/7)

- Both **selection** (nodes forming ties based on similarity) and **social influence** (nodes becoming similar due to existing ties) work synergistically
  - They collectively tend to reduce the number of neighboring node pairs holding different opinions
  - This drives the network towards a state where all connected neighbors share the same opinion
- Network Fragmentation (Polarization): As a consequence of homophilous alignment, the network eventually divides into a set of separate, disconnected components (clusters)
  - Within each component, all members hold the same opinion. However, the specific opinion may differ across different components, leading to global polarization

Coevolution of Networks and Dynamics (5/7)

- **Stable Equilibrium:** This final fragmented state represents a stable state (equilibrium) for the coevolutionary system
  - At this point, no further changes occur in either individual opinions or the network's structure, and the dynamics cease



Coevolution of Networks and Dynamics (6/7)

- Influence Dominance (When p≈0):
  - When the parameter p (representing the likeihood for network rewiring/selection) is close to zero, social influence largely dominates the dynamics, and the network structure undergoes minimal changes
  - In this scenario, opinions within the initially connected components of the network tend to homogenize
- Selection Dominance (When p≈1):
  - When the parameter p is close to one, selection (homophily) becomes the dominant mechanism, significantly reducing the impact of social influence on opinion changes
  - The final network components primarily consist of groups of nodes that shared the same opinion in the initial configuration, essentially preserving original opinion clusters

Coevolution of Networks and Dynamics (7/7)

- *What happens when the number of opinions is large?*
  - If we start from a random network with average degree larger than one, we know that it has a giant component
  - For p near zero in the long run there will be **a giant community holding the majority opinion**, and many small communities with different opinions
  - For p near one **the link dynamics will break the network into many small components**, each made mostly of nodes that were initially assigned one of the distinct opinions
  - There is an **abrupt transition** between the scenario with a large majority opinion and the scenario with many smaller opinion communities of comparable size

2. Network Search

Network Search (1/2)

- The process of locating a specific resource or target node within a given network
- To achieve this, it's essential to devise efficient strategies for exploring the network until the desired node is successfully identified and reached
- Exploration Method: Typically, the search starts from a designated "node of origin" and proceeds by progressively visiting its immediate neighbors, then their neighbors, and so on, radiating outwards

Network Search (2/2)

- Exhaustive Approaches
  - Breadth-First Search (BFS): An example of an exhaustive algorithm that systematically explores all reachable nodes layer by layer
  - While effective in some scenarios, particularly for smaller networks or when abundant computational and storage resources are available (e.g., in web crawlers for search engine indexing), BFS can be computationally intensive for large-scale networks
  - **Better Strategy**: In many real-world network scenarios, a local search strategy is often a more efficient and practical approach

Local Search

- Breadth-First Search (BFS) for Search
  - BFS can be used as a general search mechanism to find a target node, even if it's not typically categorized as "local search" in the context of taking advantage of network structure
  - Starting from a source node, the process involves exhaustively visiting all nodes in the first layer (immediate neighbors), then checking if any are the target
  - If not, the query is propagated to their neighbors, and so on, until the target node is reached
- Limitations of BFS
  - Breadth-First Search is generally **not** an efficient strategy for optimizing local search in complex networks
  - This is because BFS does not leverage specific properties of network structure (e.g., small-world phenomenon, presence of hubs, community structure) that could guide the search more directly towards the target, thus requiring unnecessary exploration of a vast number of nodes

Peer-to-Peer (P2P) Networks (1/5)

- A distributed network architecture where individual computers, or "peers," are directly connected to each other to share resources, such as files, without relying on a central server
- **Advantage (Decentralization)**
  - The absence of a central server for file location or control means the entire system is highly resilient
  - It cannot be easily compromised by the failure of any single node, or by external threats like legal challenges or Denial-of-Service (DoS) attacks directed at a central point

Peer-to-Peer (P2P) Networks (2/5)

- **Disadvantage (Resource Discovery)**
  - The exact location of a desired file or resource is initially unknown to any single peer
  - Consequently, when a user searches for a file, queries must be actively sent to other connected computers within the P2P network
  - If a peer does not possess the requested file, the query is then forwarded to one or more of its immediate neighbors, and this process continues iteratively until the file is located or the search scope is exhausted

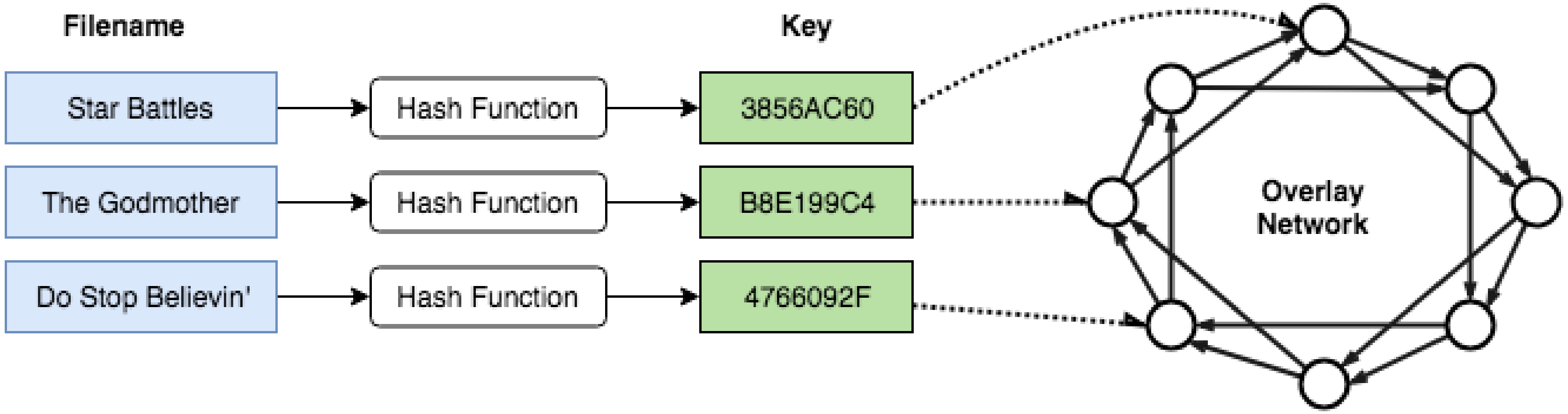
Peer-to-Peer (P2P) Networks (3/5)

P2P networks fundamentally rely on two interconnected components to facilitate decentralized resource sharing

- **Distributed Hash Table (DHT):** A decentralized system that maps unique file identifiers (keys, often cryptographic hashes of the file content) to the specific peer computers that store those files
  - This mechanism allows any peer to efficiently locate a file without requiring a central server
- **Overlay Network:** A virtual network built on top of the underlying physical network infrastructure. It connects the peer nodes that participate in the P2P system
  - This overlay dictates how queries and data requests propagate between peers, enabling the discovery and transfer of files

Peer-to-Peer (P2P) Networks (4/5)

- **File Storage Mechanism**
  - When a file needs to be stored within the P2P network, a unique key is generated for that file
  - This key is produced by a hash function, which is a cryptographic algorithm designed to output a fixed-size, unique signature from any arbitrary input data
  - Crucially, each generated key is mapped to a specific node (peer computer) within the network's address space, enabling the file to be efficiently routed to, and retrieved from, that particular peer
- **File Retrieval Process**
  - During a file search operation, the unique key associated with the desired file is utilized to forward the query through the overlay network
  - This query propagation continues iteratively until it successfully reaches the specific node that possesses the file corresponding to that key



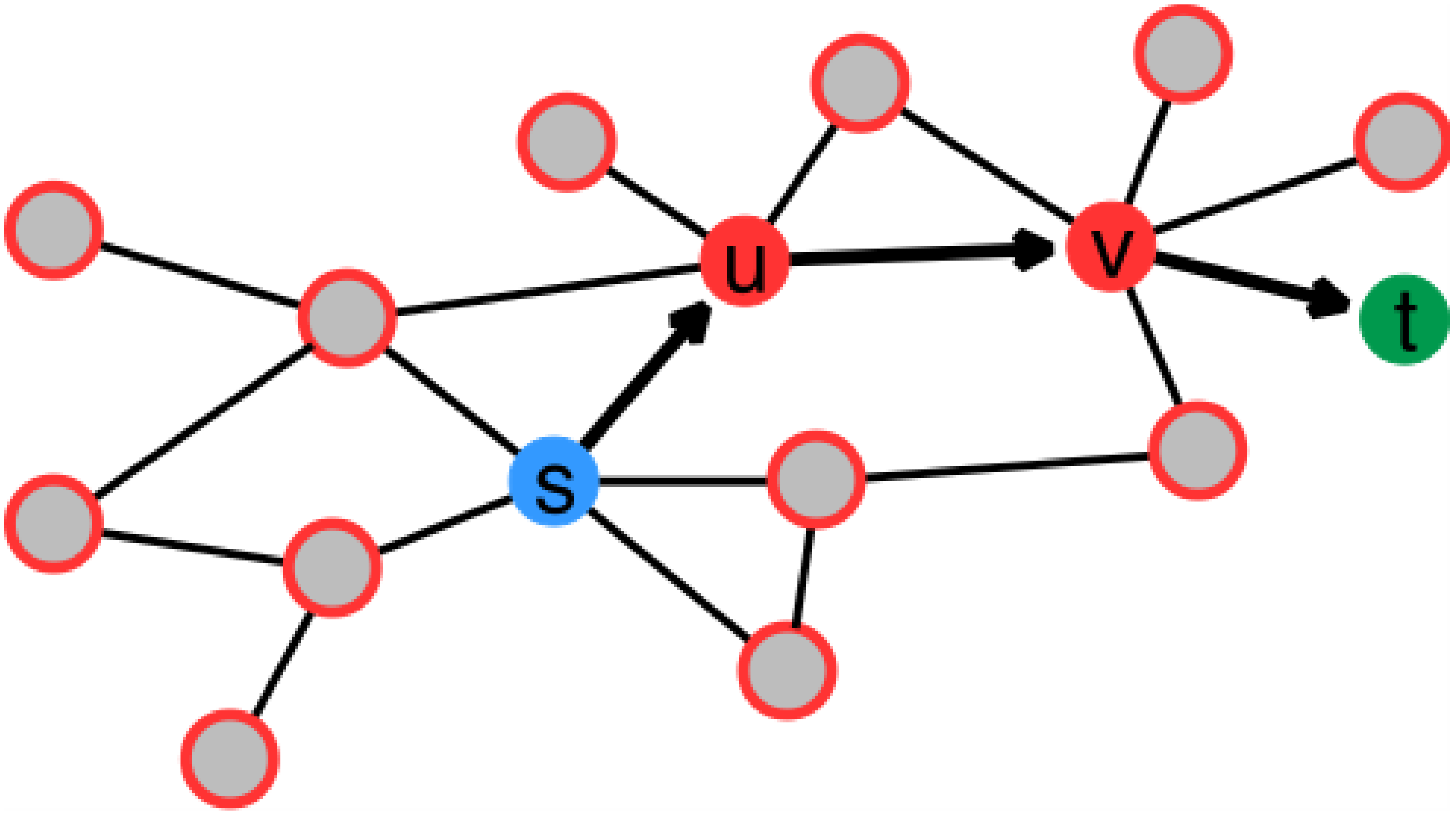
Peer-to-Peer (P2P) Networks (5/5)

- **Decentralized Routing**
  - Each peer (node) maintains a **local routing table**, which is a subset of links to its immediate neighbors within the overlay network
  - For any given file key, a node either possesses the target file itself, knows the specific node that owns that key, or holds a link to a neighbor that is demonstrably "closer" to the target key in the network's address space
  - A simple greedy routing algorithm can then be employed, directing messages to the neighbor determined to be closest to the target key in the distributed hash table
- **Dynamic Membership**
  - A significant and important property of P2P networks is their support for dynamic membership, allowing any computer to join or leave the network at any time
  - When a peer joins or leaves the network, only its directly connected neighbor peers require updates to their routing tables; **the rest of the vast network remains unaffected**

Local Search: Exploiting Network Structure (1/3)

- **Approach**: This strategy is designed to improve search efficiency by exploiting the presence of hubs (nodes with high connectivity) within the network
- **Dynamics of Search**
  - All information accessible to any given node is strictly local. Each node knows the degree of all its immediate neighbors, as well as the data (resources) stored within those neighbors
  - When a node is queried, starting from the initial source node, it forwards the request to its neighbor with the largest degree, unless the current node itself or any of its immediate neighbors is the target
  - This forwarding process is repeated iteratively until the message successfully reaches a neighbor of the target node, signifying the target's discovery
  - To avoid redundant queries and infinite loops, nodes that have already passed the request are marked as visited, ensuring that no node is queried more than once during a specific search procedure

Local Search: Exploiting Network Structure (2/3)



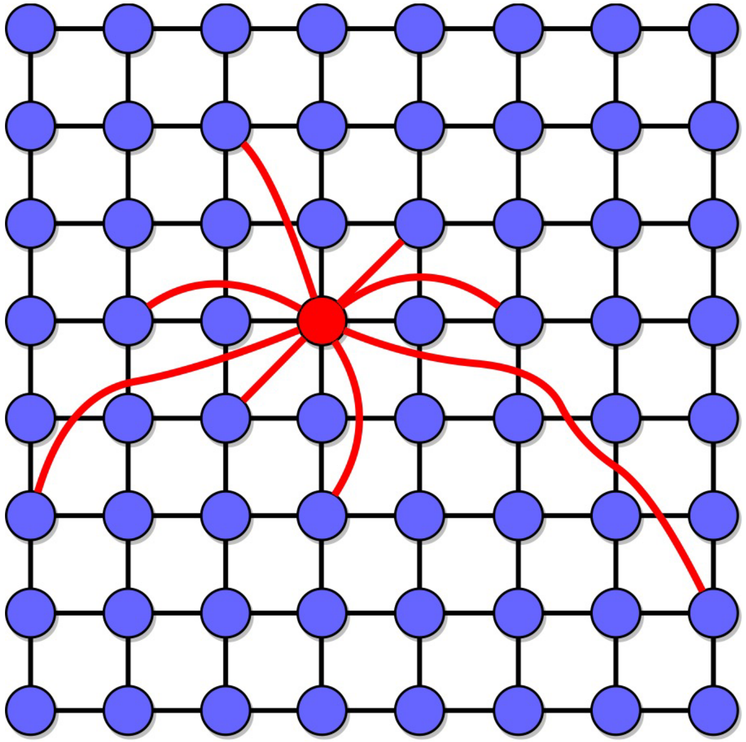
Local Search: Exploiting Network Structure (3/3)

- By preferentially exploring neighbors with a large degree (high connectivity), the search algorithm significantly increases the probability of encountering one of their neighbors that is also a major hub
  - During this initial "greedy" phase, the algorithm rapidly converges to a node with the highest degree in the local vicinity
  - Following this rapid transient phase, where visited nodes progressively exhibit larger degrees, the subsequent exploration effectively proceeds in an inverse order of the network's degree sequence—spreading downwards from the highly connected hub
  - This leads to a very rapid growth in the number of queried nodes (specifically, the neighbors of the hubs), resulting in the target being reached in a relatively small number of steps
- **Problem**: On average, the total number of nodes that must be queried by this local search strategy is comparable to that of a full Breadth-First Search
  - This is because, in principle, the target node can be located anywhere within the network

Network Searchability

- **Searchable Networks**: Networks that can be efficiently navigated to locate a specific target within a reasonably short computational time
- **Fundamental Question:** *Are all networks inherently searchable by decentralized, local strategies?*
- Insights from **Milgram's Small-World Experiment**
  - This seminal experiment revealed surprisingly short paths (the "six degrees of separation") between participants and their designated targets
  - Participants were remarkably effective at finding these paths even without comprehensive knowledge of the global network structure
  - The successful search strategy employed was largely based on **homophily**, meaning individuals tended to forward the message to contacts who were perceived to be more similar to the ultimate target person than the sender themselves (e.g., sharing geographic proximity like living in the same city, or having similar professional occupations)

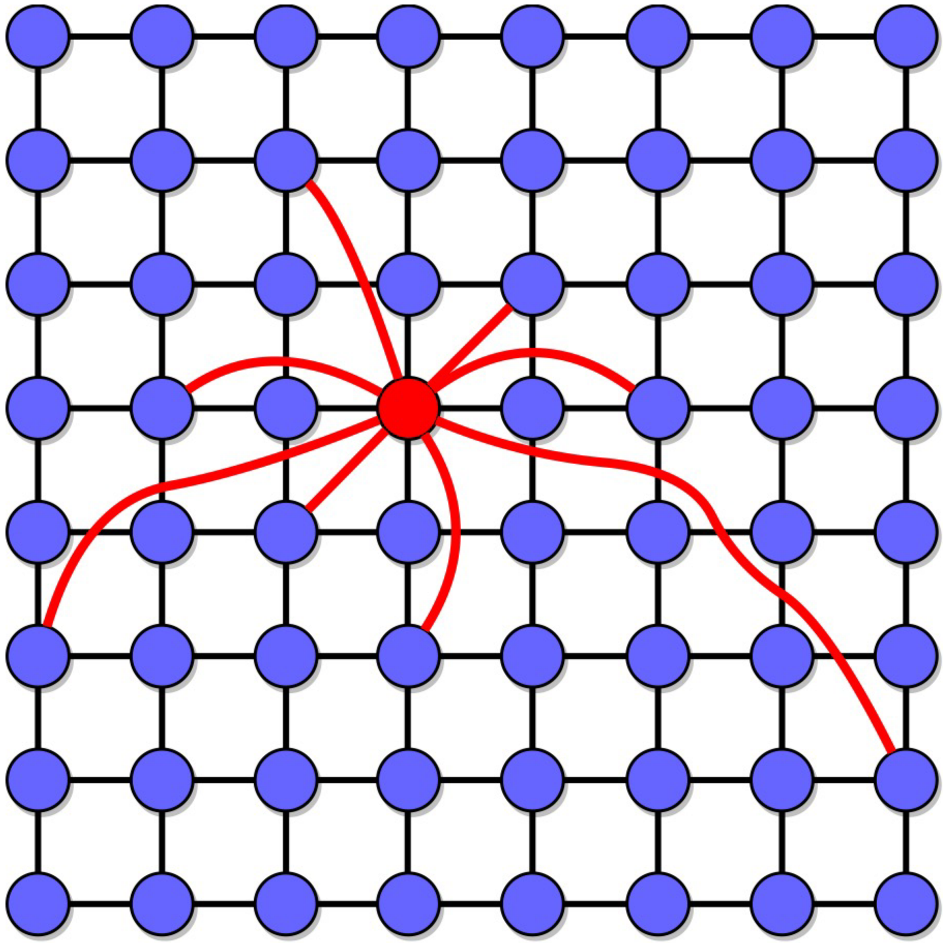
Geographic Searchability (1/4)



- The social network is conceptually embedded within a square grid to represent **geographic space**
- Each node is initially connected only to its nearest neighbors, forming a foundational grid network structure
- Additional "**shortcut**" links are subsequently added between pairs of nodes across the grid
  - Distance-Dependent Link Probability: In contrast to the standard small-world model, the probability of establishing these shortcut links decreases with increasing geographic distance between the nodes in the grid
  - This design choice accounts for the empirical observation that most relationships in real social networks are formed among individuals in geographic proximity

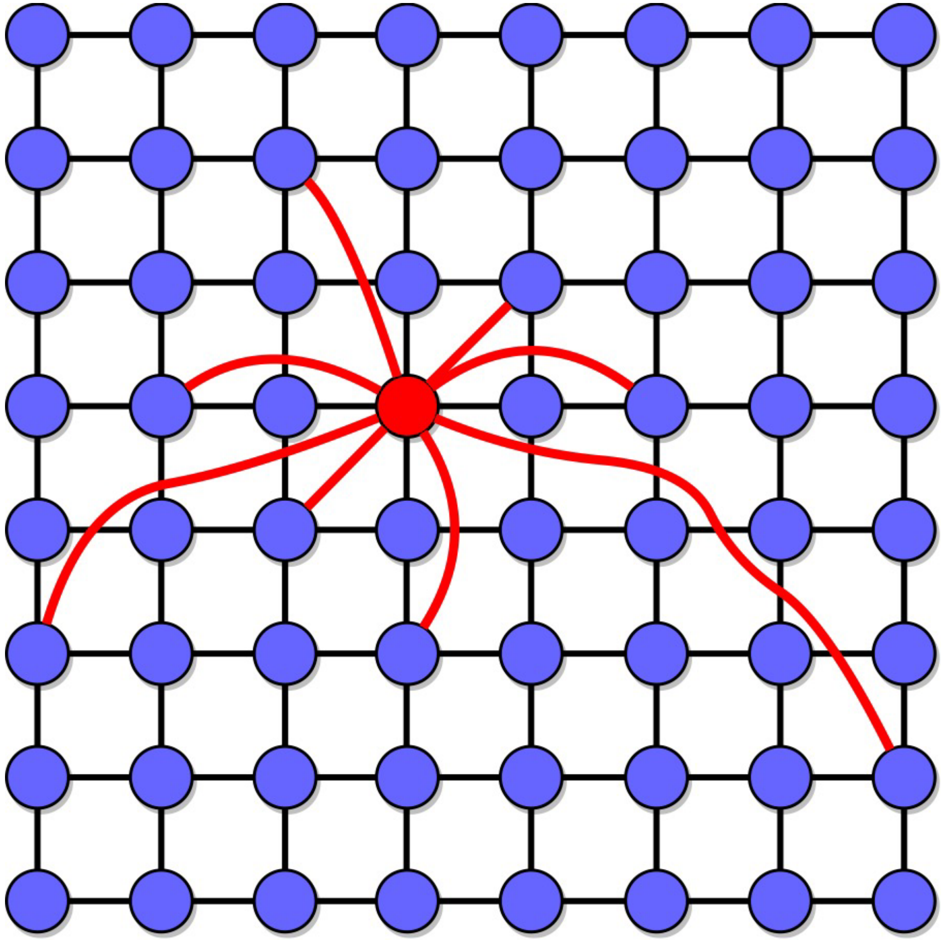
Geographic Searchability (2/4)

- Every individual knows exactly the geographic position of their neighbors, as well as the position of the target. Therefore **every individual can precisely determine which neighbor is geographically closest to the target**
- Source and target nodes are **chosen at random**
- **Greedy search algorithm:** each node forwards the message across a link that brings it as close as possible to the target
- **Delivery time:** number of times the message is passed between nodes until it reaches the target



Geographic Searchability (3/4)

- The delivery time is very short **only if the shortcut probability falls off in just the right way as a function of the geographic distance** between nodes
- In the case of a two-dimensional grid the probability of a shortcut must decay as the **inverse of the square of the distance**
- **Example:** a link between two nodes lying two steps apart from each other should be four times more likely than a link connecting two nodes that are twice as far (four steps)



Geographic Searchability (4/4)

- *What is the impact of shortcut distribution?*
  - If the probability of forming shortcut links decreases too rapidly with increasing geographic distance between nodes, the network **lacks sufficient long-range connections**
    - Consequently, the search algorithm is constrained to traverse numerous local links before it can successfully reach the target
  - Conversely, if the shortcut probability decays too slowly, the network contains an **abundance of long-range links**
    - **....** leading to many theoretically shortest paths. However, these numerous shortest paths become challenging to discover efficiently by a local search strategy

Web Searchability

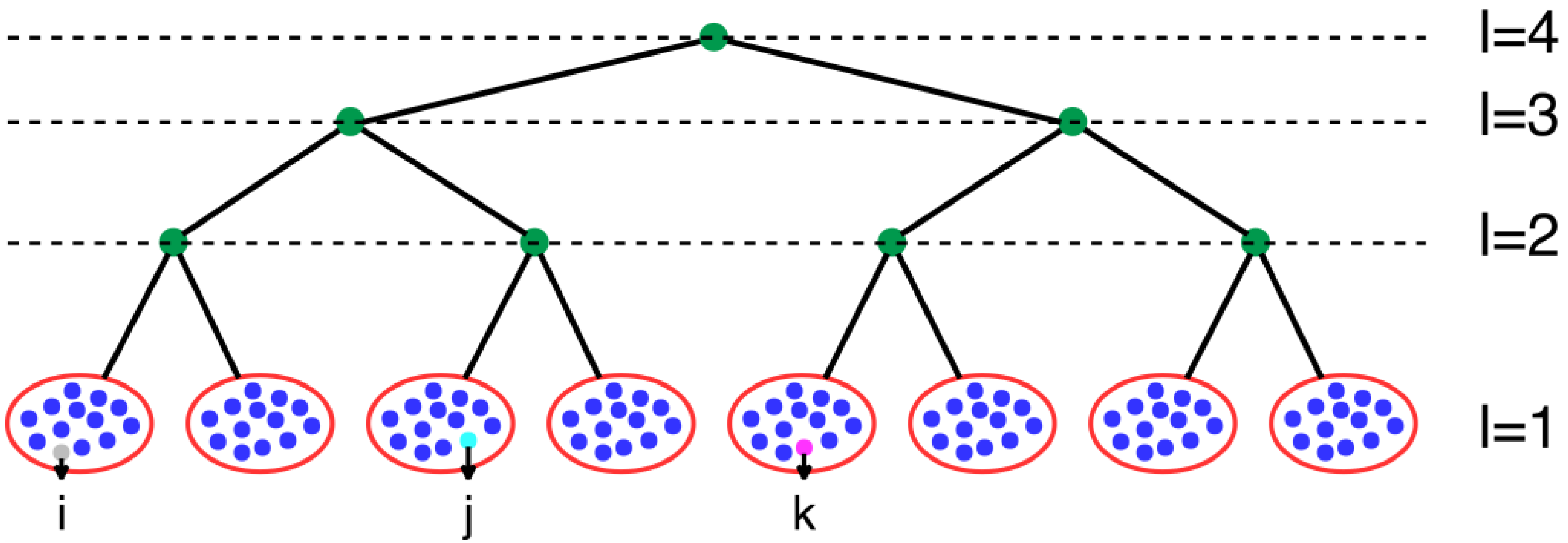
- For the World Wide Web, the concept of geographic homophily is replaced by topical locality
- This means that web pages with similar content or topics are highly likely to have common neighbors or be directly linked
- Even for distantly related (dissimilar) pages, the decay in link probability is compatible with the condition for geographic searchability
- As a result, despite its vastness, **the Web is empirically observed to be searchable**!

Topical Searchability (1/6)

- Geography constitutes just one of numerous possible attributes associated with nodes (individuals or entities) within a network undergoing a search process
- In real-world social networks, individuals can share a wide array of non-geographic similarities, such as having the same occupation, practicing the same hobby, attending the same educational institution, and so forth
- General Definition:
  - **Topical Searchability**: Refers to a network's property where any salient attribute of its nodes can be effectively reflected in network homophily, thereby facilitating the search process
  - Example
    - In Milgram's classic small-world experiment, the occupation of the target node served as a useful piece of information for participants to guide the forwarding of messages

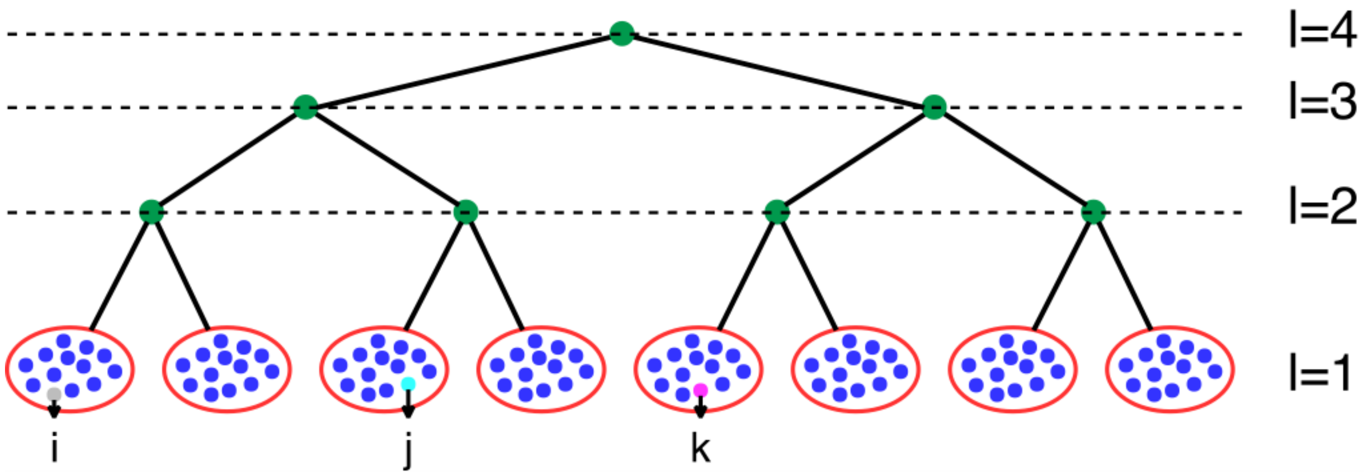
Topical Searchability (2/6)

- Network nodes can be organized **hierarchically** based on their topical attributes. The resulting structured representation is known as a **topical distance tree**
- The top of this hierarchy represents the most general topical category
- As one descends through the tree, this general category is progressively split into smaller, increasingly specific topical categories
- This process continues until the smallest identifiable groups of nodes with highly similar topical attributes are reached



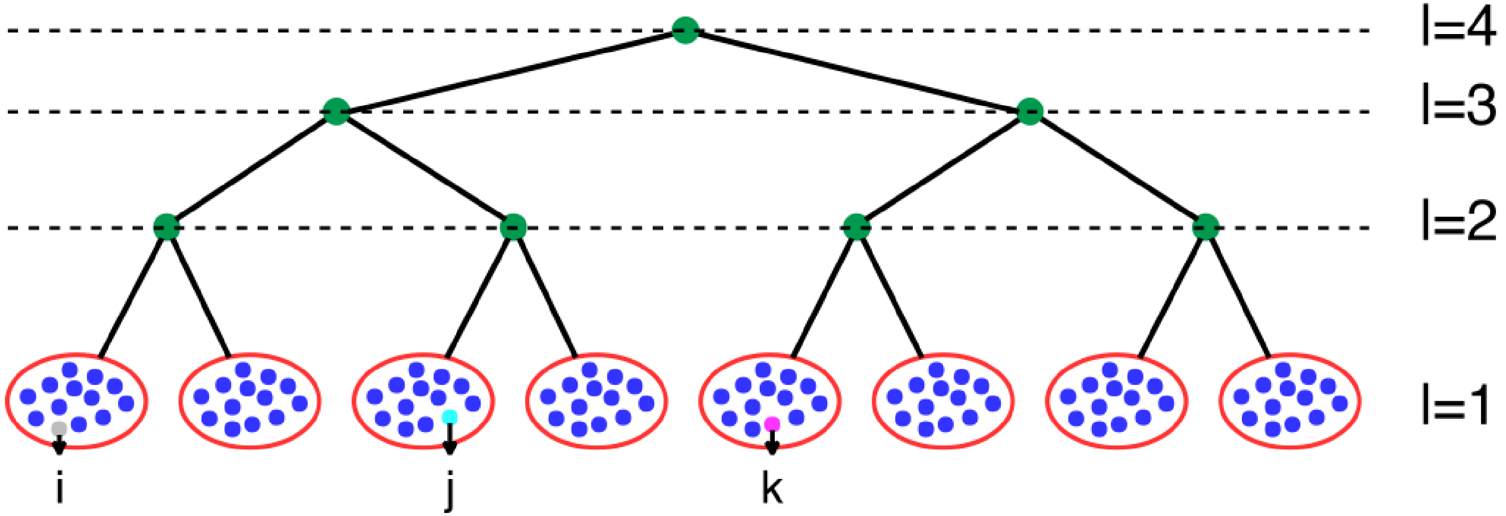
Topical Searchability (3/6)

- **Example:** Consider a topical tree structure designed to classify individuals within a social network
  - At the top would be the global population and the lower groups could represent a geographic subdivision of the population into continents, countries, cities, and neighborhoods
  - Different social attributes (e.g., occupations, hobbies, schools, religions) lead to different hierarchical divisions and trees



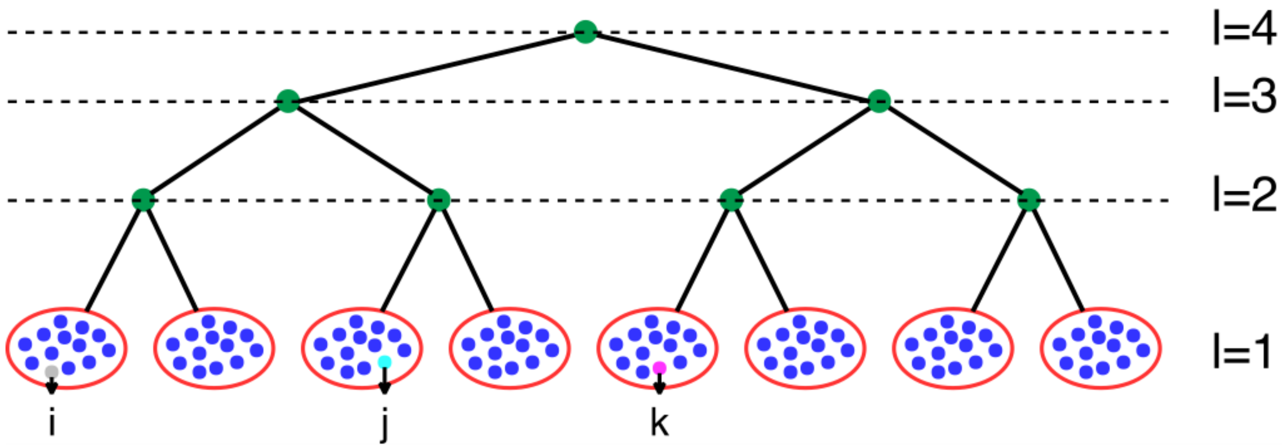
Topical Searchability (4/6)

- A topical distance tree serves as a conceptual framework that enables the estimation of the topical distance between nodes in a network
- If two individuals belong to the same smallest identifiable group within this hierarchy, their topical distance is defined as one
- Example: Consider two professors who are both working in the same academic department at Indiana University in Bloomington



Topical Searchability (5/6)

- If two individuals do not belong to the same smallest identifiable group, their respective groups will eventually merge as one ascends the hierarchical tree
- This convergence occurs when the search reaches their nearest common ancestor category in the tree, which represents the most specific shared attribute between the nodes
- In this scenario, the topical distance is quantitatively defined by the number of levels in the tree, measured from the bottom-most level up to this nearest common ancestor
- Examples
  - The topical distance between node i and node j is 3
  - The topical distance between node j and node k is 4



Topical Searchability: Underlying Assumptions (6/6)

- **Assumption I:** people can estimate their topical distance from anybody
  - Less stringent hypothesis than in the geographic model, where individuals know each other's exact position
- **Assumption II:** the topical distance tree captures the social network's homophily, so that the link probability between two nodes decreases as their topical distance increases, according to a decay function
- By using the greedy search algorithm, i.e., by letting each person forward the message to the neighbor with the shortest topical distance from the target, **there is a special topical decay function that allows for efficient search**
- **Web as a test case:** the probability that two pages are linked decays with their topical distance as expressed by the topical searchability condition **—> the Web is searchable!**

3. Summary

Summary (1/3)

- **Majority Opinion Model**
  - Nodes adopt the majority opinion of their neighbors
  - Different opinions coexist in the final state
  - Consensus reached only on one- and two-dimensional grids
  - Consensus opinion matches the majority opinion in the initial configurations
- **Voter Model**
  - Nodes adopt the opinion of a randomly selected neighbor
  - Dynamics lead to consensus on all networks
  - Consensus probability (on an opinion) matches the fraction of nodes holding that opinion in the initial configurations

Summary (2/3)

- **Bounded-Confidence Models**
  - Opinions affect each other if their difference is smaller than the confidence bound parameter
  - Final opinion clusters depend on confidence bound value and network structure
  - Sufficiently large confidence bound leads from random initial opinions to consensus on any network
- **Coevolution Models**
  - Combine the processes of selection and social influence
    - a model: Nodes either adopt a neighbor's opinion or choose a new neighbor with the same opinion
  - Final state results in segregated homogeneous opinion communities

Summary (3/3)

- **Network Search Algorithms**
  - Breadth-first search: standard approach for *exhaustive* network searches (e.g., Web crawlers)
  - Breadth-first search can be infeasible for large networks, requiring local heuristic search
  - A local search heuristic is to forward queries to nodes with the largest degree to reach hubs quickly
- **Searchable Networks**
  - Short paths can be found connecting source and target nodes
  - Searchability may be due to geographic link distribution or hierarchical node organization
  - Distance estimation between two nodes in the hierarchy, helps identify the neighbor closest to the target

References

[1] Menczer, F., Fortunato, S., & Davis, C. A. (2020). A First Course in Network Science Cambridge: Cambridge University Press.

- Chapter 7.3, 7.4

[2] OLAT course page: https://olat.vcrp.de/url/RepositoryEntry/4669112833