An Overview of Complex Networks Design and Analyses

Sanket Patil

May 12, 2008

1. Introduction
   - Definition
   - Representation
2. Topological Features
   - Scale-free Nature
   - “Small World” properties
   - Other features
3. Philosophy
   - Emergence
   - Machines vs Societies
   - Self Interest
4. Design
   - Systems Design: Desiderata
   - Complex Systems Design
5. Analyses
6. Conclusions

Complex Networks

- Not “simple networks”
- Not random graphs
- Have non-trivial topological features

Complex Networks

- Not “simple networks”
- Not random graphs
- Have non-trivial topological features

Complex Networks

- Not "simple networks"
- Not random graphs
- Have non-trivial topological features

Examples

- Computer Networks, the Internet, the Web
- Food-webs, protein interaction networks
- Social Networks, Trade Networks
- The brain

Examples

- Computer Networks, the Internet, the Web
- Food-webs, protein interaction networks
- Social Networks, Trade Networks
- The brain

Examples

- Computer Networks, the Internet, the Web
- Food-webs, protein interaction networks
- Social Networks, Trade Networks
- The brain

Examples

- Computer Networks, the Internet, the Web
- Food-webs, protein interaction networks
- Social Networks, Trade Networks
- The brain

Complex Networks

- Designed (for performance)
- Have evolved/emerged over time
- Are teleological / have purpose of life

Complex Networks

- Designed (for performance)
- Have evolved/emerged over time
- Are teleological / have purpose of life

Complex Networks

- Designed (for performance)
- Have evolved/emerged over time
- Are teleological / have purpose of life

Representation

- Networks are modelled as “Graphs”
- A node/vertex/point represents a machine, a human, a cell etc.
- An edge/arc/line represents a *relation* between two nodes
- Edges can be undirected or directed.
- *Weights* are used to convey additional (extra-topological) information

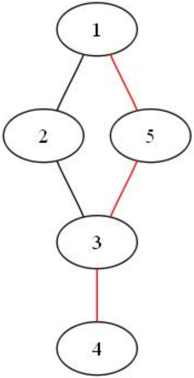
Basic Definitions

- **Degree:** The number of edges *incident* on a node
- **Indegree (outdegree):** Number of incoming (outgoing) edges
- **Degree Distribution:** A probability distribution of node degrees

Basic Definitions

- **Degree:** The number of edges *incident* on a node
- **Indegree (outdegree):** Number of incoming (outgoing) edges
- **Degree Distribution:** A probability distribution of node degrees

- **Path:** A sequence of *adjacent* vertices
- **Pathlength:** The number of edges in a path



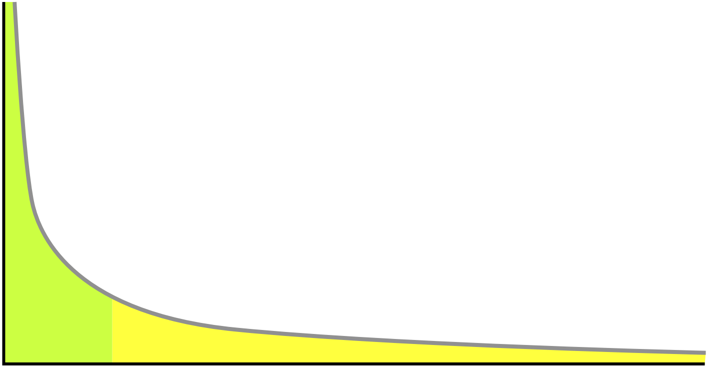
Scale-free Nature

- Network dynamics/behaviour is independent of the size
- Power Law degree distributions

Scale-free Nature

- Network dynamics/behaviour is independent of the size
- Power Law degree distributions

Power Law Distribution



Properties

- $P[X = k] \; \alpha \; k^{-p}$
- Small number of nodes with a very high degree and a large number of nodes with a very low degree
- "heavy tail", "long tail", "80:20"
- Income distributions, page rank, wikipedia contribution

Properties

- $P[X = k] \; \alpha \; k^{-p}$
- Small number of nodes with a very high degree and a large number of nodes with a very low degree
- "heavy tail", "long tail", "80:20"
- Income distributions, page rank, wikipedia contribution

Properties

- $P[X = k] \alpha\ k^{-p}$
- Small number of nodes with a very high degree and a large number of nodes with a very low degree
- "heavy tail", "long tail", "80:20"
- Income distributions, page rank, wikipedia contribution

Properties

- $P[X = k] \; \alpha \; k^{-p}$
- Small number of nodes with a very high degree and a large number of nodes with a very low degree
- “heavy tail”, “long tail”, “80:20”
- Income distributions, page rank, wikipedia contribution

Preferential Attachment

- Nodes prefer to attach to "popular" nodes
- "Rich getting richer"
- Entrenchment

Preferential Attachment

- Nodes prefer to attach to "popular" nodes
- "Rich getting richer"
- Entrenchment

Preferential Attachment

- Nodes prefer to attach to “popular” nodes
- “Rich getting richer”
- Entrenchment

“Small World” Properties

- Stanley Milgram: Small world experiment
- Mean Path Length 6
- Critique: not a comprehensive study
- “Six degrees of separation”
- Watts and Strogatz

"Small World" Properties

- Stanley Milgram: Small world experiment
- Mean Path Length 6
- Critique: not a comprehensive study
- "Six degrees of separation"
- Watts and Strogatz

“Small World” Properties

- Stanley Milgram: Small world experiment
- Mean Path Length 6
- Critique: not a comprehensive study
- “Six degrees of separation”
- Watts and Strogatz

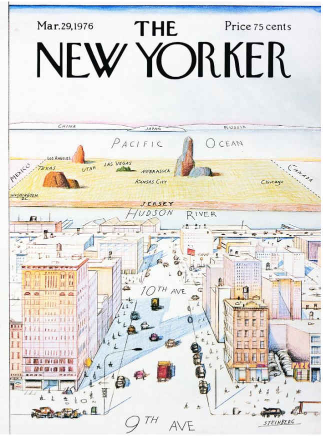
“Small World” Properties

- Stanley Milgram: Small world experiment
- Mean Path Length   6
- Critique: not a comprehensive study
- “Six degrees of separation”
- Watts and Strogatz

"Small World" Properties

- Stanley Milgram: Small world experiment
- Mean Path Length 6
- Critique: not a comprehensive study
- "Six degrees of separation"
- Watts and Strogatz

Saul Steinberg: Ninth Avenue



- Social Commentary
- Small world geometry

- Social Commentary
- Small world geometry

Kleinberg's "small world"

- Ninth Avenue as a powerful analogy
- The "far" is almost as accessible as the "near"
- How are your friends/acquaintances distributed?

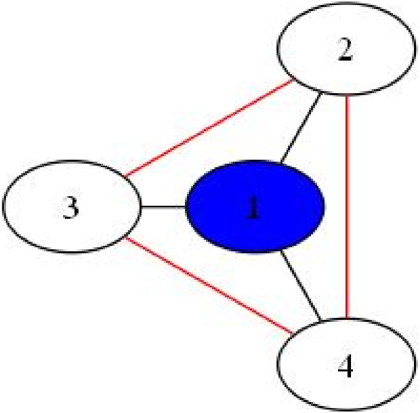
Kleinberg's "small world"

- Ninth Avenue as a powerful analogy
- The "far" is almost as accessible as the "near"
- How are your friends/acquaintances distributed?

Kleinberg's "small world"

- Ninth Avenue as a powerful analogy
- The "far" is almost as accessible as the "near"
- How are your friends/acquaintances distributed?

Clustering



Clustering Coefficient

- **Neighbourhood:** Nodes that are adjacent to a node
- $C_i = \frac{no\ \ of\ \ edges\ \ in\ \ the\ \ neighbourhood}{total\ \ possible\ \ edges}$
- $C_{graph} = \frac{\sum_{i=0}^{n} C_i}{n}$

Clustering Coefficient

- **Neighbourhood:** Nodes that are adjacent to a node
- $C_i = \frac{no\ \ of\ \ edges\ \ in\ \ the\ \ neighbourhood}{total\ \ possible\ \ edges}$
- $C_{graph} = \frac{\sum_{i=0}^{n} C_i}{n}$

Clustering Coefficient

- **Neighbourhood:** Nodes that are adjacent to a node
- $C_i = \frac{no\ \ of\ \ edges\ \ in\ \ the\ \ neighbourhood}{total\ \ possible\ \ edges}$
- $C_{graph} = \frac{\sum_{i=0}^{n} C_i}{n}$

Small Worlds

- High clustering coefficient
- Low average path length

Navigability in small worlds

- Kleinberg's metric space models (circa 2000)
- Short paths do exist
- But can we *find* them using *local* information?

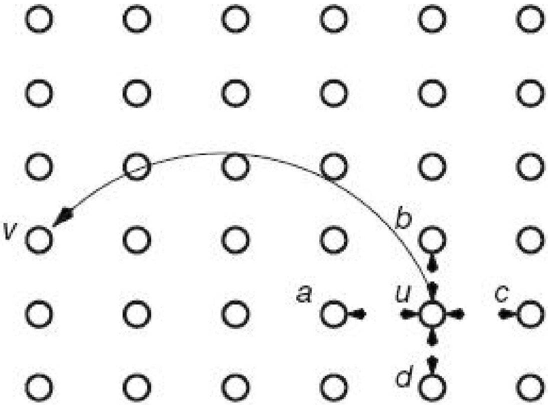
Navigability in small worlds

- Kleinberg’s metric space models (circa 2000)
- Short paths do exist
- But can we *find* them using *local* information?

Navigability in small worlds

- Kleinberg’s metric space models (circa 2000)
- Short paths do exist
- But can we *find* them using *local* information?

Short range and long range connections



Clustering Exponent

- Connections based on distance (r) and clustering exponent α
- For a node *u*, the probability of connecting to *v* is $r^{-\alpha}$
- Highly clustered neighbourhood
- Number of long range links decays with distance
- Only when $\alpha = 2$, a decentralized routing algorithm can be found which has a *log n* bound

Clustering Exponent

- Connections based on distance (r) and clustering exponent α
- For a node *u*, the probability of connecting to *v* is $r^{-\alpha}$
- Highly clustered neighbourhood
- Number of long range links decays with distance
- Only when $\alpha = 2$, a decentralized routing algorithm can be found which has a *log n* bound

Clustering Exponent

- Connections based on distance (r) and clustering exponent α
- For a node *u*, the probability of connecting to *v* is $r^{-\alpha}$
- Highly clustered neighbourhood
- Number of long range links decays with distance
- Only when $\alpha = 2$, a decentralized routing algorithm can be found which has a *log n* bound

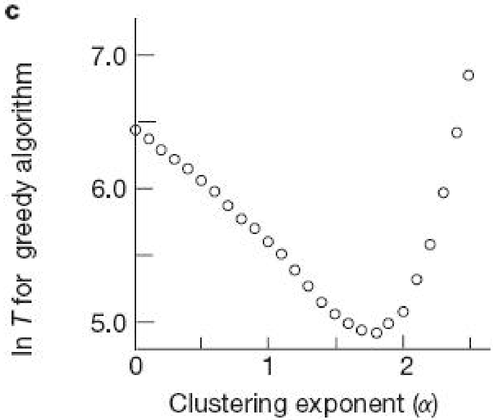
Clustering Exponent

- Connections based on distance (r) and clustering exponent α
- For a node *u*, the probability of connecting to *v* is $r^{-\alpha}$
- Highly clustered neighbourhood
- Number of long range links decays with distance
- Only when $\alpha = 2$, a decentralized routing algorithm can be found which has a *log n* bound

Clustering Exponent

- Connections based on distance (r) and clustering exponent α
- For a node u, the probability of connecting to v is $r^{-\alpha}$
- Highly clustered neighbourhood
- Number of long range links decays with distance
- Only when $\alpha = 2$, a decentralized routing algorithm can be found which has a $log\ n$ bound

Finding short paths



Other Features

- Communities
- Hierarchical Structures

Other Features

- Communities
- Hierarchical Structures

Complex Systems: Philosophy

- Systems Theory
- Cybernetics
- Non-linear Dynamics
- Multiagent Systems and AI
- Network Science, Connectionism, Cognitive Psychology

Complex Systems: Philosophy

- Systems Theory
- Cybernetics
- Non-linear Dynamics
- Multiagent Systems and AI
- Network Science, Connectionism, Cognitive Psychology

Complex Systems: Philosophy

- Systems Theory
- Cybernetics
- Non-linear Dynamics
- Multiagent Systems and AI
- Network Science, Connectionism, Cognitive Psychology

Complex Systems: Philosophy

- Systems Theory
- Cybernetics
- Non-linear Dynamics
- Multiagent Systems and AI
- Network Science, Connectionism, Cognitive Psychology

Complex Systems: Philosophy

- Systems Theory
- Cybernetics
- Non-linear Dynamics
- Multiagent Systems and AI
- Network Science, Connectionism, Cognitive Psychology

Strudy of Complex Networks

- Structure or Topology
- How structure governs the function
- How design can be analysed teleologically
- Can we use this knowledge to build better systems?

Strudy of Complex Networks

- Structure or Topology
- How structure governs the function
- How design can be analysed teleologically
- Can we use this knowledge to build better systems?

Strudy of Complex Networks

- Structure or Topology
- How structure governs the function
- How design can be analysed teleologically
- Can we use this knowledge to build better systems?

Strudy of Complex Networks

- Structure or Topology
- How structure governs the function
- How design can be analysed teleologically
- Can we use this knowledge to build better systems?

Emergence

- The connections between the components is as important as the components
- Emergent behaviour of complex networks
- Case Study: The Human Brain

Emergence

- The connections between the components is as important as the components
- Emergent behaviour of complex networks
- Case Study: The Human Brain

Emergence

- The connections between the components is as important as the components
- Emergent behaviour of complex networks
- Case Study: The Human Brain

Machines vs Societies

- Systems Thinking: parts and whole
- Machines funcion based on norms
- Societies are declarative
- Societies are more robust
- Case study: The Heart

Machines vs Societies

- Systems Thinking: parts and whole
- Machines funcion based on norms
- Societies are declarative
- Societies are more robust
- Case study: The Heart

Machines vs Societies

- Systems Thinking: parts and whole
- Machines funcion based on norms
- Societies are declarative
- Societies are more robust
- Case study: The Heart

Machines vs Societies

- Systems Thinking: parts and whole
- Machines funcion based on norms
- Societies are declarative
- Societies are more robust
- Case study: The Heart

Machines vs Societies

- Systems Thinking: parts and whole
- Machines funcion based on norms
- Societies are declarative
- Societies are more robust
- Case study: The Heart

Self Interest

- Autonomous agents
- Local constraints and self interest
- Global or environmental dampeners

Self Interest

- Autonomous agents
- Local constraints and self interest
- Global or environmental dampeners

Self Interest

- Autonomous agents
- Local constraints and self interest
- Global or environmental dampeners

Systems Design: Desiderata

- Safety
- Resilience
- Fairness
- Livenss

Systems Design: Desiderata

- Safety
- Resilience
- Fairness
- Livenss

Systems Design: Desiderata

- Safety
- Resilience
- Fairness
- Livenss

Systems Design: Desiderata

- Safety
- Resilience
- Fairness
- Livenss

Complex Systems Design

- Efficiency
- Robustness
- Cost

Complex Systems Design

- Efficiency
- Robustness
- Cost

Complex Systems Design

- Efficiency
- Robustness
- Cost

An Optimization Problem

- Designing complex systems is an optmization process
- Search in a multidimensional space

An Optimization Problem

- Designing complex systems is an optmization process
- Search in a multidimensional space

Approaches

- Mathematical Programming
- Ant Colony Optimization
- Swarm Intelligence
- Simulated Annealing
- Genetic Algorithms

Graph Theoretic Analyses

- Graph thoretic properties used in design and analyses
- Constraints and objectives are defined in terms of graph properties
- Design is usually evolution of optimal graphs under the given constraints

Graph Theoretic Analyses

- Graph thoretic properties used in design and analyses
- Constraints and objectives are defined in terms of graph properties
- Design is usually evolution of optimal graphs under the given constraints

Graph Theoretic Analyses

- Graph thoretic properties used in design and analyses
- Constraints and objectives are defined in terms of graph properties
- Design is usually evolution of optimal graphs under the given constraints

Efficiency

- **APL:** Average of all pairs shortest paths
- **Diameter:** The longest shortest path. An upper bound on efficiency.

Efficiency

- **APL:** Average of all pairs shortest paths
- **Diameter:** The longest shortest path. An upper bound on efficiency.

Local Objectives

- **Eccentricity:** The longest shortest path *for a node*
- The greatest separation a node suffers
- Average eccentricity or eccentricity distribution are also good indicators of efficiency
- **Radius:** Smallest eccentricity. “Central” node
- Number of “central” nodes can be another measure

Local Objectives

- **Eccentricity:** The longest shortest path *for a node*
- The greatest separation a node suffers
- Average eccentricity or eccentricity distribution are also good indicators of efficiency
- **Radius:** Smallest eccentricity. “Central” node
- Number of “central” nodes can be another measure

Local Objectives

- **Eccentricity:** The longest shortest path *for a node*
- The greatest separation a node suffers
- Average eccentricity or eccentricity distribution are also good indicators of efficiency
- **Radius:** Smallest eccentricity. “Central” node
- Number of “central” nodes can be another measure

Local Objectives

- **Eccentricity:** The longest shortest path *for a node*
- The greatest separation a node suffers
- Average eccentricity or eccentricity distribution are also good indicators of efficiency
- **Radius:** Smallest eccentricity. “Central” node
- Number of “central” nodes can be another measure

Local Objectives

- **Eccentricity:** The longest shortest path *for a node*
- The greatest separation a node suffers
- Average eccentricity or eccentricity distribution are also good indicators of efficiency
- **Radius:** Smallest eccentricity. “Central” node
- Number of “central” nodes can be another measure

Cost

- **Density(connectance): no of edges/no of possible edges**
- Edges per node
- Weights, costs and other environment dependent measures

Cost

- **Density(connectance): no of edges/no of possible edges**
- Edges per node
- Weights, costs and other environment dependent measures

Cost

- **Density(connectance): no of edges/no of possible edges**
- Edges per node
- Weights, costs and other environment dependent measures

Robustness

- Centrality Measures
- Connectivity

Robustness

- Centrality Measures
- Connectivity

Centrality

- **Degree Centrality:** degree distribution
- **Betweenness:** importance of nodes based on the no. of paths passing through them
- **Closeness:** *per node* average path length
- **Eigenvector Centrality:** importance of nodes based not just on *how many* are endorsing, but also *who* is connected.
- When a centrality distribution is uniform, the network is most robust

Centrality

- **Degree Centrality:** degree distribution
- **Betweenness:** importance of nodes based on the no. of paths passing through them
- **Closeness:** *per node* average path length
- **Eigenvector Centrality:** importance of nodes based not just on *how many* are endorsing, but also *who* is connected.
- When a centrality distribution is uniform, the network is most robust

Centrality

- **Degree Centrality:** degree distribution
- **Betweenness:** importance of nodes based on the no. of paths passing through them
- **Closeness:** *per node* average path length
- **Eigenvector Centrality:** importance of nodes based not just on *how many* are endorsing, but also *who* is connected.
- When a centrality distribution is uniform, the network is most robust

Centrality

- **Degree Centrality:** degree distribution
- **Betweenness:** importance of nodes based on the no. of paths passing through them
- **Closeness:** *per node* average path length
- **Eigenvector Centrality:** importance of nodes based not just on *how many* are endorsing, but also *who* is connected.
- When a centrality distribution is uniform, the network is most robust

Centrality

- **Degree Centrality:** degree distribution
- **Betweenness:** importance of nodes based on the no. of paths passing through them
- **Closeness:** *per node* average path length
- **Eigenvector Centrality:** importance of nodes based not just on *how many* are endorsing, but also *who* is connected.
- When a centrality distribution is uniform, the network is most robust

Connectivity

- Connected, strongly connected and weakly connected
- *Vertex Cut:* The smallest number of vertices whose removal renders the network disconnected
- *Edge Cut:* The smallest number od edges whose removal renders the network disconnected
- A graph is $k-connected$ if the size of its *vertex cut* is k
- Higher the connectivity, more robust the network

Connectivity

- Connected, strongly connected and weakly connected
- *Vertex Cut:* The smallest number of vertices whose removal renders the network disconnected
- *Edge Cut:* The smallest number od edges whose removal renders the network disconnected
- A graph is $k - \textit{connected}$ if the size of its *vertex cut* is k
- Higher the connectivity, more robust the network

Connectivity

- Connected, strongly connected and weakly connected
- *Vertex Cut:* The smallest number of vertices whose removal renders the network disconnected
- *Edge Cut:* The smallest number od edges whose removal renders the network disconnected
- A graph is $k - connected$ if the size of its *vertex cut* is k
- Higher the connectivity, more robust the network

Connectivity

- Connected, strongly connected and weakly connected
- *Vertex Cut:* The smallest number of vertices whose removal renders the network disconnected
- *Edge Cut:* The smallest number od edges whose removal renders the network disconnected
- A graph is $k-connected$ if the size of its *vertex cut* is k
- Higher the connectivity, more robust the network

Connectivity

- Connected, strongly connected and weakly connected
- *Vertex Cut:* The smallest number of vertices whose removal renders the network disconnected
- *Edge Cut:* The smallest number od edges whose removal renders the network disconnected
- A graph is $k - connected$ if the size of its *vertex cut* is k
- Higher the connectivity, more robust the network

Independent/Disjoint Paths

- Two paths are **vertex independent** if they have no common vertices (except the terminals)
- Two paths are **edge independent** if they have no common edges
- **Menger's Theorem:** *Max-flow min-cut*
- More independent paths implies higher robustness

Independent/Disjoint Paths

- Two paths are **vertex independent** if they have no common vertices (except the terminals)
- Two paths are **edge independent** if they have no common edges
- **Menger's Theorem:** *Max-flow min-cut*
- More independent paths implies higher robustness

Independent/Disjoint Paths

- Two paths are **vertex independent** if they have no common vertices (except the terminals)
- Two paths are **edge independent** if they have no common edges
- **Menger's Theorem:** *Max-flow min-cut*
- More independent paths implies higher robustness

Independent/Disjoint Paths

- Two paths are **vertex independent** if they have no common vertices (except the terminals)
- Two paths are **edge independent** if they have no common edges
- **Menger's Theorem:** *Max-flow min-cut*
- More independent paths implies higher robustness

Number of connected components

- Robustness can be measured in terms of the *number* of components that result due to failures
- *Size* of the connected components also matters in many cases
- **Toughness:** A toughness of t means, for a given number $k(>1)$, at least $t * k$ nodes need to be removed to fragment the graph into k connected components
- **Graph Toughness:** Biggest value of t

Number of connected components

- Robustness can be measured in terms of the *number* of components that result due to failures
- *Size* of the connected components also matters in many cases
- **Toughness:** A toughness of t means, for a given number $k(> 1)$, at least $t * k$ nodes need to be removed to fragment the graph into k connected components
- **Graph Toughness:** Biggest value of t

Number of connected components

- Robustness can be measured in terms of the *number* of components that result due to failures
- *Size* of the connected components also matters in many cases
- **Toughness:** A toughness of t means, for a given number $k(> 1)$, at least $t * k$ nodes need to be removed to fragment the graph into k connected components
- **Graph Toughness:** Biggest value of t

Number of connected components

- Robustness can be measured in terms of the *number* of components that result due to failures
- *Size* of the connected components also matters in many cases
- **Toughness:** A toughness of t means, for a given number $k(> 1)$, at least $t * k$ nodes need to be removed to fragment the graph into k connected components
- **Graph Toughness:** Biggest value of t

Conclusions

- Complex Networks design have non-trivial topological properties
- Structure governs function
- Designed for efficiency under constraints of robustness and cost
- Graph thoretic measures are useful for design and analyses

Conclusions

- Complex Networks design have non-trivial topological properties
- Structure governs function
- Designed for efficiency under constraints of robustness and cost
- Graph thoretic measures are useful for design and analyses

Conclusions

- Complex Networks design have non-trivial topological properties
- Structure governs function
- Designed for efficiency under constraints of robustness and cost
- Graph thoretic measures are useful for design and analyses

Conclusions

- Complex Networks design have non-trivial topological properties
- Structure governs function
- Designed for efficiency under constraints of robustness and cost
- Graph thoretic measures are useful for design and analyses