Statistical physics of complex networks

IXXI Lyon July 2008

Marc Barthelemy

CEA, France
EHESS-CAMS, France
References: reviews

• **Statistical mechanics of complex networks**
  Reka Albert, Albert-Laszlo Barabasi
  Reviews of Modern Physics 74, 47 (2002)
  cond-mat/0106096

• **The structure and function of complex networks**
  cond-mat/0303516

• **Evolution of networks**
  cond-mat/0106144
References: books

- Evolution and Structure of the Internet
  - Romualdo Pastor-Satorras
  - Alessandro Vespignani
  - Cambridge

- Scale-Free Networks
  - Guido Caldarelli
  - Oxford University Press
References: books
I. Introduction: Complex networks
   1. Complex systems and networks
   2. Graph theory and characterization of large networks: tools
   3. Characterization of large networks: results
   4. Models

II. Dynamical processes
   1. Resilience and vulnerability
   2. Epidemiology

III. Advanced topics
   1. Global disease spread
   2. Community detection
   3. Evolution and formation of the urban street network in cities
Complex Systems

- No commonly accepted definition

Properties

- Large number of (possibly non-identical) interacting constituents
- Emergent behavior
- Adapt and evolve (resilient to failure)-different from complicated!

Network structure (nodes+links)

- Present everywhere-data recently available
Complex Systems - Methodology

- Theoretical & Empirical analysis
  - Data-driven research
  - Characterization and modeling
  - Dynamical processes (eg. Epidemics)

- Interdisciplinarity
  - Collaboration with scientists from other fields
  - Confront with the communities
What is a network?

Network = set of nodes joined by links

- very abstract representation
- very general
- convenient to describe many different systems
Most networks of interest are:

- Complex
- Very large

Statistical tools needed!

‘Statistical mechanics’ of large networks
### Ubiquity of networks

<table>
<thead>
<tr>
<th>Category</th>
<th>Nodes</th>
<th>Links</th>
</tr>
</thead>
<tbody>
<tr>
<td>Social networks</td>
<td>Individuals</td>
<td>Social relations</td>
</tr>
<tr>
<td>IT: Internet WWW</td>
<td>Routers/AS</td>
<td>Cables</td>
</tr>
<tr>
<td></td>
<td>Webpages</td>
<td>Hyperlinks</td>
</tr>
<tr>
<td>Biology: PIN Ecosystems</td>
<td>Proteins</td>
<td>Hyperlinks</td>
</tr>
<tr>
<td></td>
<td>Species</td>
<td>Trophic relation</td>
</tr>
<tr>
<td>Infrastructures</td>
<td>Hubs</td>
<td>Airlines, roads, …</td>
</tr>
</tbody>
</table>
Example: social networks

Many social networks are the support of some dynamical processes

- (Epidemics)
- Rumor propagation
- Opinion/consensus formation
- Cooperative phenomena
- ...

Scientific collaboration network

Nodes: scientists
Links: co-authored papers

Weights: depending on
  • number of co-authored papers
  • number of authors of each paper
  • number of citations…
Citation network

**Nodes**: papers

**Links**: citations

Science citation index
S. Redner
Actor’s network

**Nodes**: actors

**Links**: cast jointly

N = 212,250 actors

\( \langle k \rangle = 28.78 \)

Character network

**Nodes:** characters

**Links:** co-appearance in a scene

Les Miserables-V. Hugo
Newman & Girvan, PRE (2004)
-> Community detection problem
The web of Human sexual contacts

Liljeros et al., Nature (2001)
Information technology

Importance of Internet and the web

- Congestion
- Virus propagation
- Cooperative/social phenomena (online communities, etc.)
- ...
Internet

- Nodes = routers
- Links = physical connections

different granularities

Router Level

Autonomous System level
Internet mapping

- continuously evolving and growing
- intrinsic heterogeneity
- self-organizing

Largely unknown topology/properties

Many mapping projects (topology and performance):
CAIDA, NLANR, RIPE, …
Internet backbone

Nodes: Computers, routers

Links: physical lines

Large-scale visualization
World Wide Web

Virtual network to find and share informations

Nodes: WWW documents
Links: URL links

Over 1 billion documents

ROBOT: collects all URL's found in a document and follows them recursively
Networks in biology

- **Cellular level**: Extracting useful information from the huge amount of available data (genome, etc)

- **Species level**: Stability of ecosystems, biodiversity
Metabolic Network

**Nodes**: metabolites  
**Links**: chemical reactions

Protein Interactions

**Nodes**: proteins  
**Links**: interactions
Food webs

**Nodes**: species

**Links**: feeds on

N. Martinez
Transportation networks

Transporting energy, goods or individuals

- formation and evolution

- congestion, optimization

- disease spread
Transporting water

**Nodes**: intersections, auxins sources
**Links**: veins

Example of a planar network
Transporting goods

State of Indiana (Bureau of Transportation statistics)
Transportation networks: intra city
Transportation networks: intra city

**TRANSIMS project**

**Nodes**: locations (homes, shops, offices, …)

**Links**: flow of individuals

<table>
<thead>
<tr>
<th>Person ID</th>
<th>Location ID</th>
<th>Location type</th>
<th>Arrival time</th>
<th>Departure time</th>
</tr>
</thead>
<tbody>
<tr>
<td>116</td>
<td>4356</td>
<td>Home</td>
<td>00:00</td>
<td>07:00</td>
</tr>
<tr>
<td>116</td>
<td>98135</td>
<td>Work</td>
<td>08:00</td>
<td>11:00</td>
</tr>
<tr>
<td>116</td>
<td>71457</td>
<td>Work</td>
<td>11:20</td>
<td>13:00</td>
</tr>
<tr>
<td>116</td>
<td>98135</td>
<td>Work</td>
<td>13:20</td>
<td>17:00</td>
</tr>
<tr>
<td>116</td>
<td>4356</td>
<td>Home</td>
<td>18:00</td>
<td>19:15</td>
</tr>
<tr>
<td>116</td>
<td>21343</td>
<td>Social</td>
<td>19:30</td>
<td>21:00</td>
</tr>
<tr>
<td>116</td>
<td>4356</td>
<td>Home</td>
<td>21:00</td>
<td>07:00</td>
</tr>
<tr>
<td>324</td>
<td>12679</td>
<td>Home</td>
<td>00:00</td>
<td>07:00</td>
</tr>
<tr>
<td>324</td>
<td>431</td>
<td>School</td>
<td>08:00</td>
<td>14:00</td>
</tr>
<tr>
<td>324</td>
<td>12679</td>
<td>Home</td>
<td>14:30</td>
<td>19:00</td>
</tr>
</tbody>
</table>

*Nature (2004)*
Transportation networks: inter city

Inter-cities movements

- Sardinian network:
  - Nodes: 375 Cities
  - Link $w_{ji} = w_{ij}$: 
    # of individuals going from i to j (daily and by any means)

Nodes: cities
Links: commuters flow

de Montis & al, Env. Plan. 2007
Transportation networks: global scale

**Nodes**: airports

**Links**: direct flight
Studies on complex networks

• 1. Empirical studies
  Typology- find the general features

• 2. Modeling
  Basic mechanisms/reproducing stylized facts

• 3. Dynamical processes
  Impact of the topology on the properties of dynamical processes: epidemic spread, robustness, …
Empirical studies: Unprecedented amount of data.....

- Transportation infrastructures (eg. BTS)
- Census data (socio-economical data)
- Social networks (eg. online communities)
Empirical studies: sampling issues

- Social networks: various samplings/networks
- Transportation network: reliable data
- Biological networks: incomplete samplings
- Internet: various (incomplete) mapping processes
- WWW: regular crawls
- ...

posibility of introducing biases in the measured network characteristics
Networks characteristics

Networks: of very different origins

Do they have anything in common? Possibility to find common properties?

- The abstract character of the graph representation and graph theory allow to give some answers…
- Important ingredients for the modeling
Modeling complex networks

Microscopical processes

- many interacting elements
- dynamical evolution
- self-organisation

Statistical physics

Properties at the macroscopic level

- Non-trivial structure
- Emergent properties, cooperative phenomena
Model validation

Modeling other attributes: clustering, assortativity, spatial effects…

Comparison with large scale datasets