Small World

Networks, Dynamics and the Small World Phenomonen
Duncan J. Watts

Presentation by:
Henrik Thostrup Jensen and Morten Goodwin Olsen
for S2_4 Mining the Social Web Mining for web2.0 PhD Course at Aalborg University

Small World Phenomonen – Informal Definition

- A and B (complete strangers) for the first time.
- After some conversation, they find out that A and B have a mutual acquantance.
- Famous: six degrees of separation.

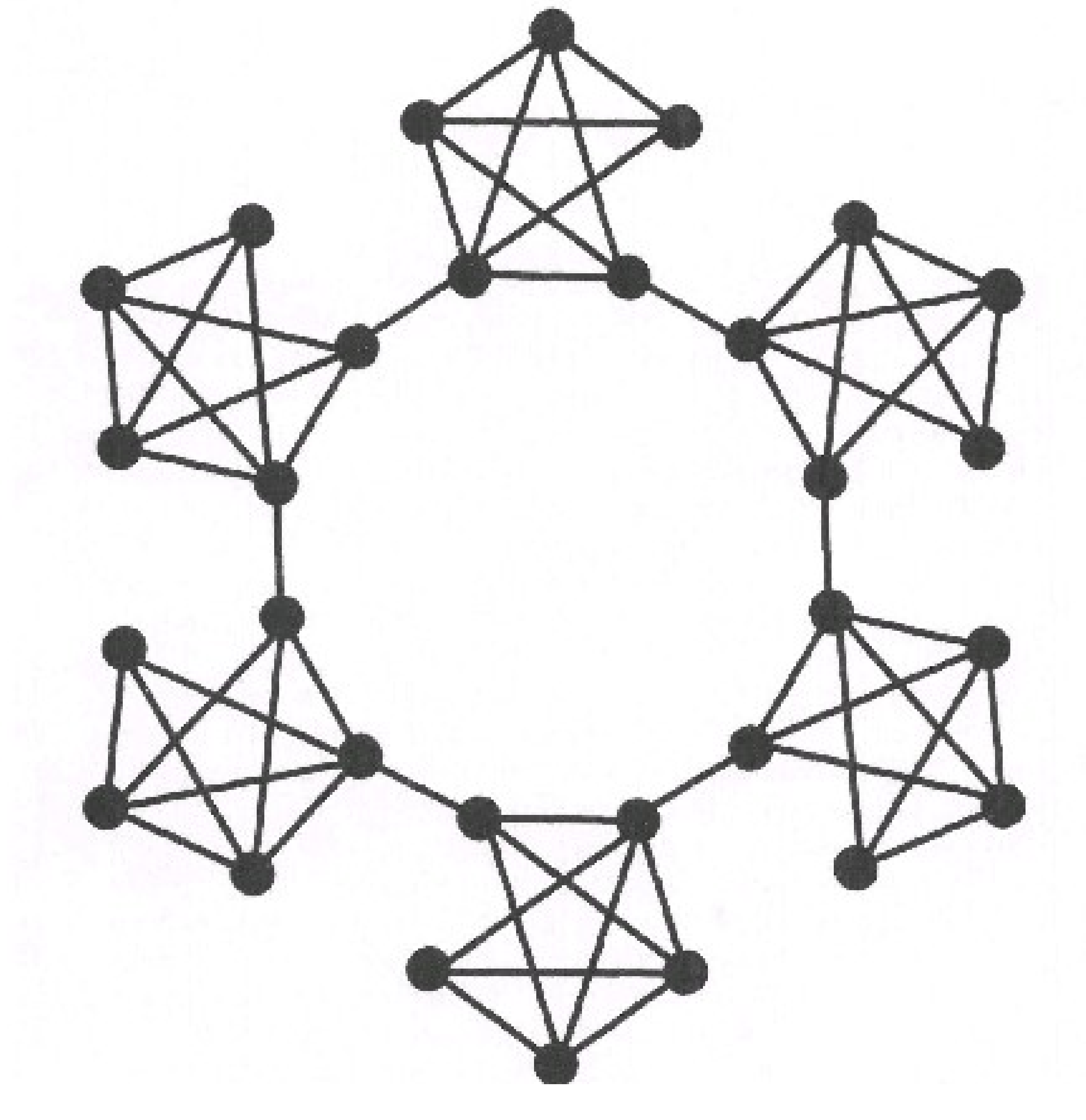
Properties

- The network is:
  - Numerically large, n >> 1. (real world: n billions)
  - Sparse, each person knows k people (hundred thousand times smaller than n)
  - Decentralised, no dominant central vertex. $k_{max} << n$
  - Higly clustered

Formal Definition

- Population n is fixed.
- Average degree k of vertices are
  - sparse: $k_{max} << n$,
  - sufficiently dense: $k_{max} >> 1$.
- Connected: Any vertix can be reached from any vertix.

Most Clustered Sparse Graph - Caveman



Most Clustered Sparse Graph - Caveman

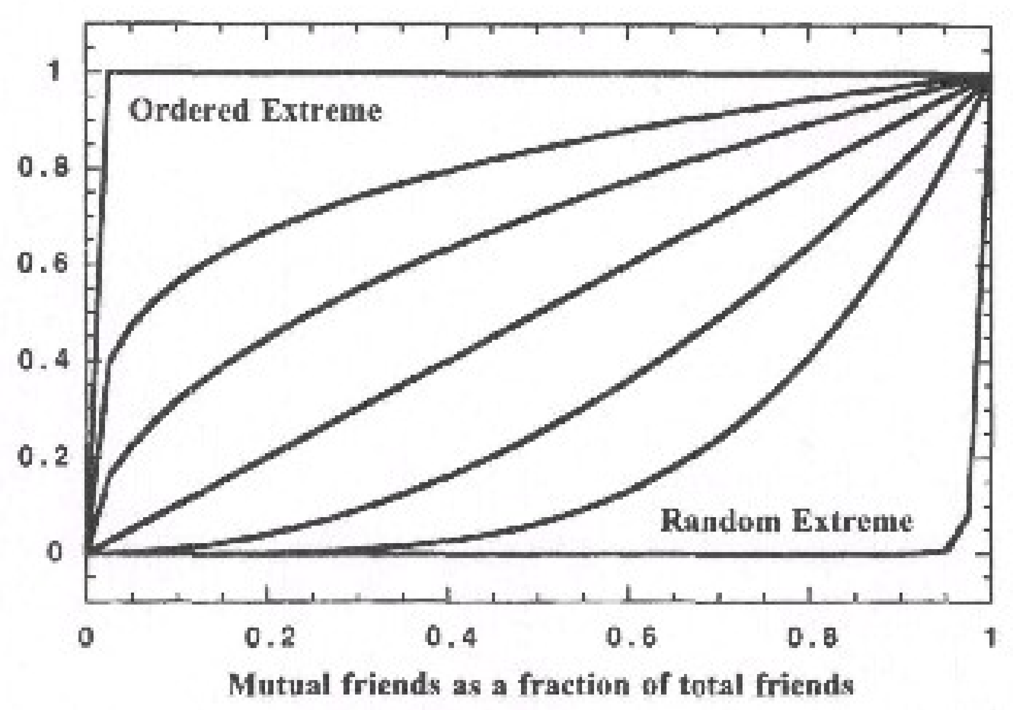
- People become friends only by aqcuaintance
- Clustering Coefficient $C = 1-6/(k^2-1)$
  - Goes towards 1 as k increases.
- Average Path length $L = n/(2(k+1))$
  - When n>>k, L is large.
- Highly cluster / locally ordered

Random Graph

- People become friends by complete random
- Clustering Coefficient $C = k/n$
- Average Path Length $L = ln(n)/ln(k)$
- $L_{random}<<L_{caveman}$
- Note:
  - Sparse Condition implies $k << n$.
  - Graphs with long characteristics, short path length
- clustering vanishes for large n

Real World

- In between Complete Ordered (Caveman) and Random
- Potential Graphs – alpha model



alpha model

- Abstract model
- Each alpha, enourmos finite number of graphs.
  - Small alpha => isolated graphs, L = infinite
- To deal with this
  - Minimal structure
    - No special vertices (no star, chain etc.)
  - Minimally connected
    - No more edges than necessary

Definitions – Small world

- large-n,sparse connected, decentralised $(n \gg k_{max} \gg 1)$. $L = L_{random}$. $C \gg C_{random}$
- Range, r= second shortest path.
- Shortcut, edge $r>2$
- Phi = Fraction of all edges that are shortcuts.
- Contraction = Pair of vertices that share one and only one common neighbor.
- **Adding contractions e.g. caveman makes a graph closer to small world.**

Evidence of small worlds

- Milgrams experiements suggest it
  - The letter experiments explained yesterday
  - We cannot always find the shortest chain
- Mapping the excact networks is impossible
  - There might not be an excact network
  - Global information not available locally
- The graph does not have to be a social network

Evindence of small worlds

- Movie actor collaboration
  - Vertex: Movie actor, Edge: Two actors appearing in same movie
  - n = 226000, k=61
- Power Grid
  - Well connected grid -> efficiancy and robustness
  - n = 4941, k=2.94
- Neural network of C. elegans
  - Roundworm :-)
  - n = 282, k = 14

Results from studies

CHARACTERISTIC PATH LENGTH (L) AND CLUSTERING COEFFICIENT (C) FOR THREE REAL NETWORKS

	L_{Actual}	L_{Random}	C_{Actual}	C_{Random}
Movie actors	3.65	2.99	.79	.00027
Power grid	18.7	12.4	.080	.005
C. elegans	2.65	2.25	.28	.05

- Length close to average for random
- Clustering significantly higher
- The three networks exhibits the same charactaristica
  - though order of magnitude difference in size
- Small worlds are very real

Dynamical systems on small worlds

- Small local changes in the graph can have dramatic impact on global properties
- Can this happen to dynamical systems connected in a small-world manner?
  - Or: Does modelling dynamical systems as small-worlds reveal anything new
- Dynamical systems are a bit different

Disease spreading 1/2

- Modelling disease spreading as a graph
- Previous work has modelled population as a uniform mix and with simple structured patterns
  - Structure heavily impacts disease and parasites
  - If the population structure is small-world like, it could be used for better predictions
- Sat up a model for this
  - p = infectiones degree
  - $p < 0.11$: disease die quickly

Disease spreading 2/2

- Sat up a model for this (cont.)
  - 0.11 <= p <= 0.5, depends on structure
  - p > 0.5, disease take over population
- Number of infected depends on topology as well as infections degree
  - Not exactly news
  - New: The required changes to the structure are subtle
- Small world spread disease faster
  - almost as fast as random, significantly faster than caveman graph

Small worlds

- Class of graphs which:
  - n >> 1
  - sparsely connected
  - high clustering
  - low pair length
  - stucture is somewhere between random and ordered
- model came from social network, but found several other places
  - especially in biology and technology