Small worlds

An introduction to the world of degrees.

Konrad Diwold, Uni Osnabrück, Germany

{ Alife Mutants Hackingsession on Systems and Organisms, Bielefeld 2004 }

Content

The small world phenomenon

The small-world phenomenon formalizes the anecdotal notion that “you are only ever six 'degrees of separation' away from anybody else on the planet.”

An example for a small world:
The Erdös number
Every mathematician with a finite Erdös number has a number of less than 8 -- only about 2% are higher, and none is more than 15.

The small world phenomenon

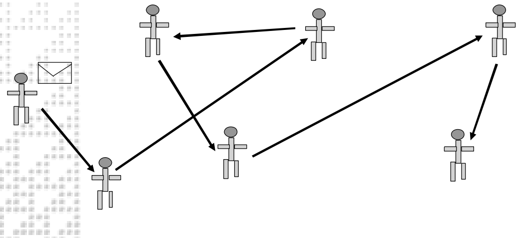
The first who studied this phenomenon was Stanley Milgram (1967):

He tracked the route of letters between 2 random people in different US states that didn't know each other.

A person was only allowed to send the letter to a person it knew the first name of if it didn't know the intended receiver.

The small world phenomenon

Results of Milgrams experiment:
A median chain length about six -> notion of “six degrees of seperation”



The small world phenomenon

A bit more mathematically:
although *empirical social networks* have much greater local density, than *random graphs*, the **mean distance** between two randomly chosen nodes in *empirical social networks* is **not much greater** than that in *random graphs* of the same network density.

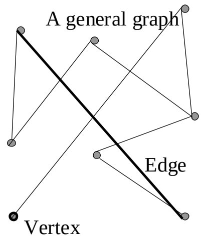
So indirect communication between people through other people is nearly as efficient as in random graphs.

The small world phenomenon

Therefor it is important to identify
the nonrandom property of networks that
yields the small-world phenomenon.

Small world & graphs

Definition: A graph G
consists of a nonempty set of elements
called vertices and a list of unordered
pairs of these elements called edges.



<u>Given a graph G:</u>

V(G) = *vertex set of G*
(i.e. **Order of the graph** ***(n)***)

E(G) = *edge list of G*
(i.e. **Size of the graph** ***(M)***)

Small world & graphs

The graphs we will look at conform to
the following
restrictions:

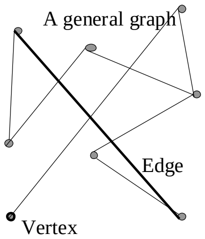
--Undirected

--Unweighted

--Simple

--Sparse

--Connected



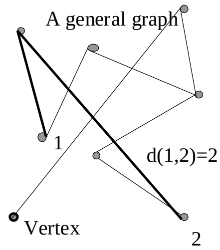
Small world & graphs

Characteristic path length L(G):

is the typical distance d(i,j) between every
vertex and every other vertex.

(it is the median of the means
of the shortest path lengths
connecting each vertex
v ∈ V(G) to all other vertices)

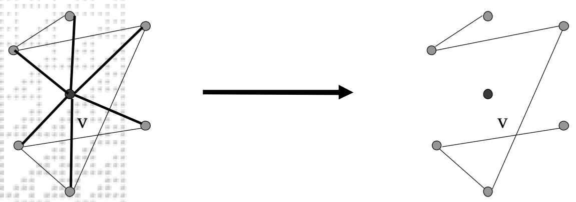
Distance:= how many edges does
it take to get from vertex a
to vertex b.



Small world & graphs

The neighbourhood of a vertex v Γ(v):

is the subgraph that consists of the vertices adjacent to v (not
including v itself)



Neighborhood of vertex v including v

Only the edges between vertices in Γ(v)

Small world & graphs

The clustering coefficient (γ):

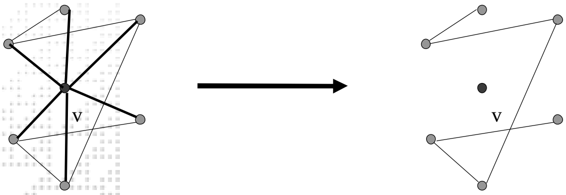
The clustering coefficient γ_v of Γ(v) characterises the extent to
which vertices adjacent to any vertex v are adjacent to each
other.

$\gamma_v = |E(\Gamma_v)| / (k_v \text{ over } 2)$

$|E(\Gamma_v)|$ = the number of edges in the neighbourhood of v

(k_v over 2) = the total number of possible edges

Small world & graphs



Neighborhood of vertex v including v

Only the edges between vertices in $\Gamma(v)$

(k_v over 2) = 21
$|E(\Gamma_v)| = 5$

-> $\gamma_v = 5/21$

Small world & graphs

A d-lattice graph:

A d-lattice is a labeled, unweighted undirected simple graph that is similar to a Euclidean cubic lattice of dimension d in that any vertex v is joined to its lattice neighbours, u_i and w_i, as specified by

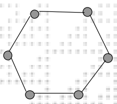
$u_i = [(v-i^{d})+n] \pmod n$,

$w_i = (v+i^{d'}) \pmod n$,

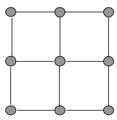
where $1 \leq i \leq k/2$, $1 \leq d' \leq d$, and it is generally assumed that $k > 2d$.

Small world & graphs

Hence a 1-lattice with k = 2 is a ring



A two lattice with k = 4 (boundaries are periodic



Small world & graphs

Random Graphs:
A random graph of order n is nothing more than
a vertex set, consisting of n vertices, and an edge set that is generated in some random fashion.

There are **2 possibilities to create** random graphs:
G(n,M) = random graph with M randomly chosen edges.

G(n,p) = every possible edge exists with the probability $0 < p < 1$ independent of other
edges

Small world & graphs

Random Graphs:
On striking feature of random graphs is, that most monotone properties appear suddenly.

That means there exists a **threshold function M*(n)** that determines whether or not a graph is **either very unlikely** or **likely** to have a **propertie Q**.

Small world & graphs

Random Graphs:
-> One could imagine a random graph as a graph process or dynamic organisms that evolve in time, starting with vertices and no edges and connecting two vertices each step.

The threshold function is the “critical time”:
-before it the property is unlikely
-after it the property is likely

Models of Graphs

2 types of construction-algos can be used to construct grap models:

- relational graphs

New edges are created as a function of preexisting edges.

- spatial graphs

Edge creation depends on a external metric

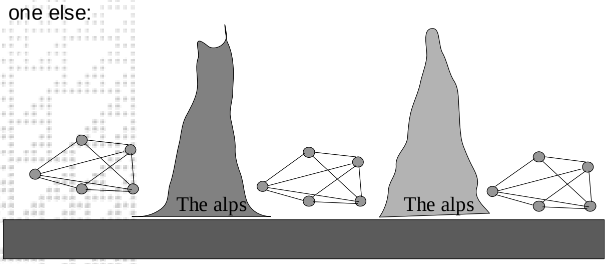
Models of Graphs: relational graphs

<u>The α-Modell:</u>
is designed to represent to construction of a network similar to a real social network, i.e. The new edges are formed due to a function of the currently existing network.
(*the current friends of a person determine to some extent the new friends of a person*)

Models of Graphs: relational graphs

<u>The α-Modell some possible worlds:</u>

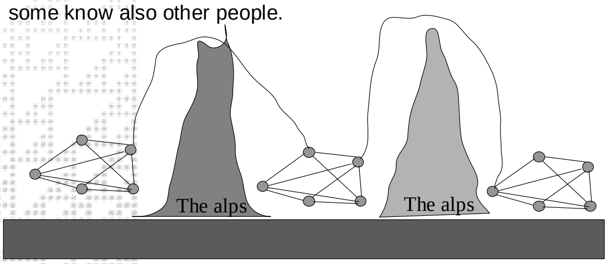
the Caveman world:
Everybody you know knows everybody else you know but no one else:



Models of Graphs: relational graphs

<u>The α-Modell some possible worlds:</u>

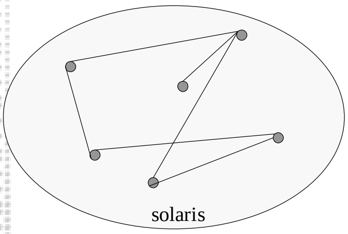
the connected-Caveman world:
Everybody you know knows everybody else you know and some know also other people.



Models of Graphs: relational graphs

<u>The α-Modell some possible worlds:</u>

the Solaris world:
The influence of existing friends on new friendships is indistinguishable from random chance.

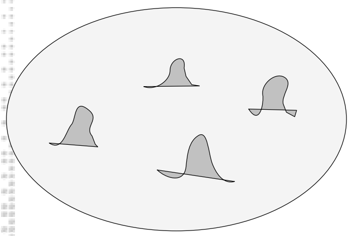


Models of Graphs: relational graphs

<u>The α-Modell some possible worlds:</u>

The truth lies somewhere in between.

A small world world could look like this :-)



Models of Graphs: relational graphs

The α-Modell a construction formula:

$$R_{i,j} = \begin{cases} 1, & m_{i,j} \geq k \\ [\, m_{i,j} / k \,] \text{ ^ alpha } * (1 - p) + p, & k > m_{i,j} > 0 \\ p, & m_{i,j} = 0 \end{cases}$$

$R_{i,j}$ = a measure of vertex i's propensity to connect to vertex j (zero if they are already connected)
$m_{i,j}$ = the number of vertices which are adjacent both to i and j
k = the average degree of the graph
p = a baseline, random probability of an edge (i,j) existing (p << (n over 2)^-1)^2

alpha = a tunable parameter, $0 \leq \alpha \leq \infty$

Models of Graphs: relational graphs

The algorithm proceeds as follows:

1. Fix a vertex i.
2. For every other vertex j, compute $R_{i,j}$ according to the algo with the additional constraint that $R_{i,j} = 0$ if a connection between i and j already exists
3. sum $R_{i,j}$ over all j and normalise each to obtain variables $P_{ij} = R_{i,j} / \sum_{l \neq i} R_{i,l}$. Then since $\sum_j P_{i,j} = 1$, $P_{i,j}$ can be interpreted as the probability that i will connect to j. Furthermore $P_{i,j}$ can be interpreted geometrically as follows: divide the unit intervall (0,1) into n -1 half open subintervals with length $P_{i,j}$, $\forall j \neq i$.
4. A uniform random variable is then generated on (0,1). It must fall into one of the subintervals, say (j*)
5. Connect i to j*.

Models of Graphs: relational graphs

The question concerning the algorithm:

How do the statistical properties of characteristic path length (L) and clustering coefficient (γ) depend upon α for fixed n and k.

Models of Graphs: relational graphs

The α = 0 Case:**:**

When α = 0 new connections are determined almost exclusively by the arrangement of existing connections.
If $m_{i,j} = 0$, then $R_{i,j} = p$ which is very small.
If $m_{i,j} > 0$, then $R_{i,j} = 1$

-> the interval into which all the $R_{i,j}$ are normalized is occupied almost entirely by subintervals that represent those vertices with which the vertex i already has mutal connections.
-> new friendships are likely to be added that increasingly interconnect the subgraph without extending it.

Models of Graphs: relational graphs

The α increases Case:**:**

As α increases, the cluster gradually connect via randomly assigned edges, and the characteristic component length increases to a maximum, at which the entire graph is connected.
After this the characteristic path length drops rapidly to its asymtotic random graph limit as the edge assignment becomes less and less correlated with previously constructed edges.

Models of Graphs: relational graphs

The need for a structure:

Question: What initial condition should be used?

If there are no existing edges ->
edges will form randomly until some connected subgraph forms – which then provides the basis for the future edges.

Side effects of it: for small α, randomly created clusters will remain disconnected from one another

for a large k the cluster will get larger and for sufficiently large k the clusters will connect.

Models of Graphs: relational graphs

The need for a structure:

Question: Which connected substrate could be used.

Constraints for the substrate:

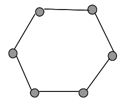
1) It must exhibit minimal structure:
   i.e. No special structures like stars or trees

2) It must be minimal connected:
   it contains no more edges than necessary to connect the graph in a "minimal structure" way.

Models of Graphs: relational graphs

The need for a structure:

The only structure which satisfies both conditions is a topological ring.



One could argue that the ring contains a considerable degree of structure itself, and should not be used due to it. Indeed this would be true if the ring would be a finished product.

But it is only the initial condition.
And allows the greatest possible range of outcomes in terms of
CPL

Models of Graphs: relational graphs

Nevertheless the small world properties can also be established in other Substrates.
(i.e. Lattice Substrates, Tree Substrates, Minimally Connected Random Substrates or No Substrates at all)
The thing needed for it is a high α

Models of Graphs: relational graphs

The β Model:

In this model, there is no talk of mutal mutual friends or clusters, but simply a perfect ring structure that, by virtue of a single parameter, metamorphoses into a random graph.

The algorithm for the β-model starts with a perfect 1-lattice in which each vertex has precisely k neighbors (k/2 on either side) and then randomly rewires the edges of the lattice, with a probability β.

Models of Graphs: relational graphs

The β Model:

Algorithm:

1. Each vertex i is chosen in turn, along with the edge that connects it to its nearest neighbor in a clockwise sense(i,i+1).
2. A uniform random deviate r is generated. If r ≥ β, then the edge (i,i+1) is unaltered. If r < β, then (i,i+1) is deleted and rewired such that i is connected to another vertex J, which is chosen uniformly at random from the entire graph (excluding self connections and repeated connections).
3. After all vertices have been considered once the procedure is repeated for the next near neighbor (i.e. i+2)

Models of Graphs: relational graphs

The β Model:

When β = 0 the resulting graph remains precisely a 1-tattice, when β = 1, all edges are rewired randomly -> close to a random graph.
So the interesting cases are these with intermediate β.

Models of Graphs: relational graphs

Both models encountered jet exhibit similar limiting properties, and both also display sudden and rapid transitions between these limits.

Question: Is there a general mechanism on which the small world feature depends?

Models of Graphs: relational graphs

Range: The range of an edge R(i,j) is the length of the shortest path between i and j in the absence of that edge.

A edge (i,j) is called an r-edge if it has a range R(i,j) = r

An r-edge R(i,j) with r > 2 is called shortcut
(i.e. if r < 2 i and j connect to the same neighborhood)

If r > 2 it is likely that, since the edges i,j are not in the same neighborhood, i and j connect to vertices anywhere in that graph

Given a graph of M = (k*n)/2 edges, the fraction of those edges that are shortcuts is denoted by Φ.

Models of Graphs: relational graphs

If two vertices u and w are both elements of the same neighborhood Γ(v), and the shortest path length between them that does not involve any edges adjacent with v is >2, then v is said to contract u and w, and the pair (u,w) is said to be a contraction .

Ψ = The fraction of all pairs of vertices that are not connected and have only one common neighbour

Ψ is a analogous parameter to Φ.

Models of Graphs: relational graphs

Both α and β effect the possibility of the creation shortcuts and contracts, which seem to be crucial for small world networks since they decrease the characteristic path length

Reference

Ducan J. Watts 2003:
Small worlds The Dynamics of Networks between order and randomness
Princeton University Press
ISBN 0-691-0054109