

Mapping change in large networks

Martin Rosvall & Carl Bergstrom

December 5, 2008

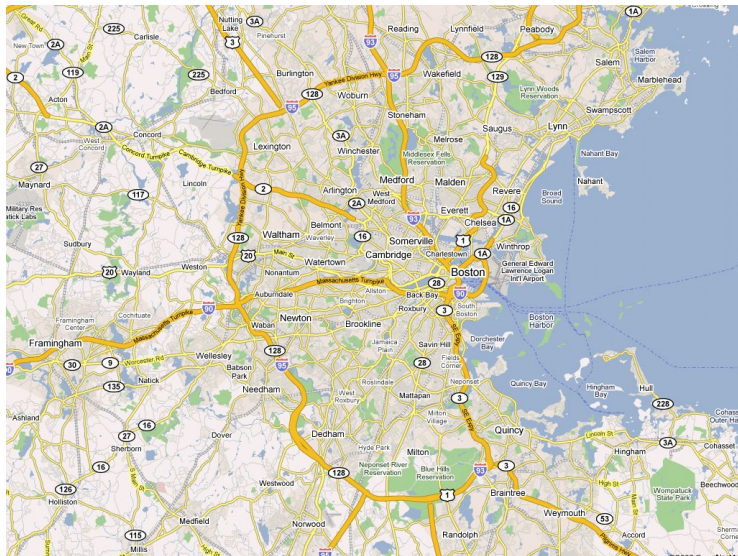
MAPPING

Pictures are pretty



<http://maps.google.com>

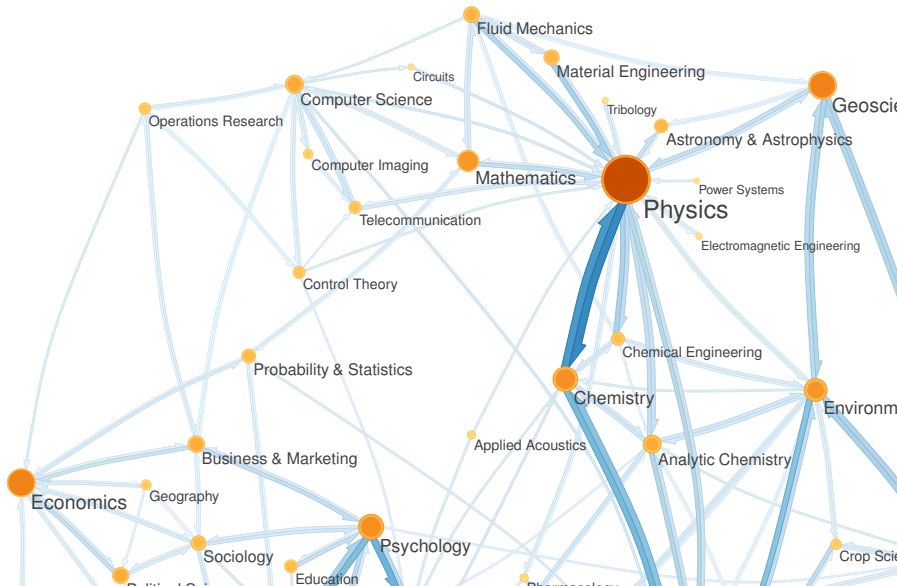
Maps simplify and highlight



Maps are tools

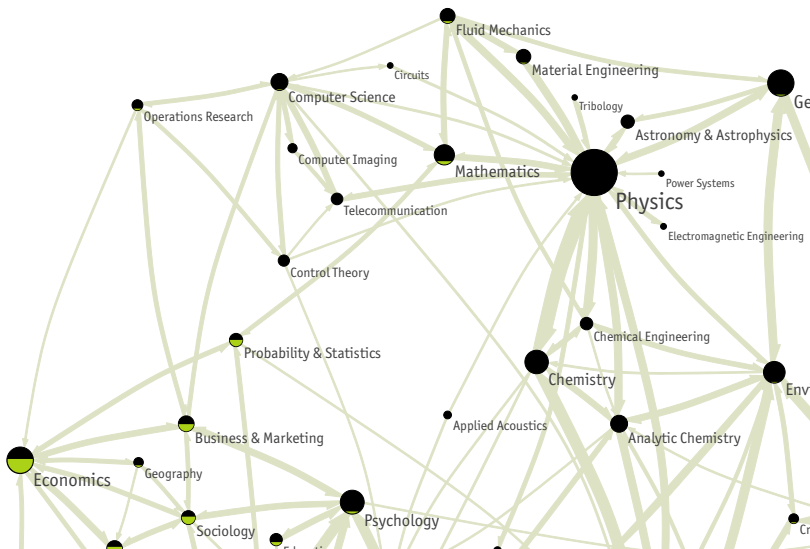


Maps simplify and highlight



Maps are tools

JSTOR coverage



The best maps are also good compressions

34

The Mathematical Theory of Communication

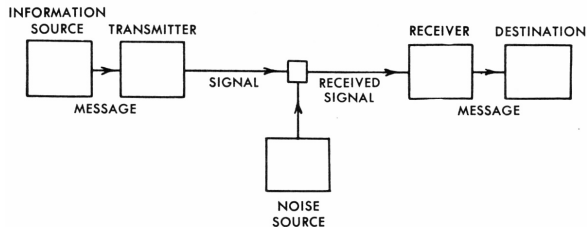
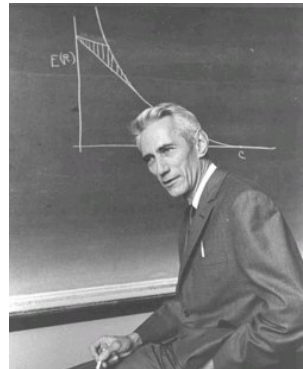
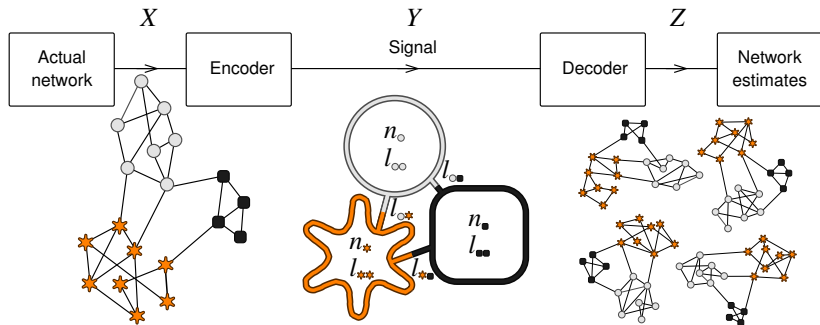


Fig. 1. — Schematic diagram of a general communication system.

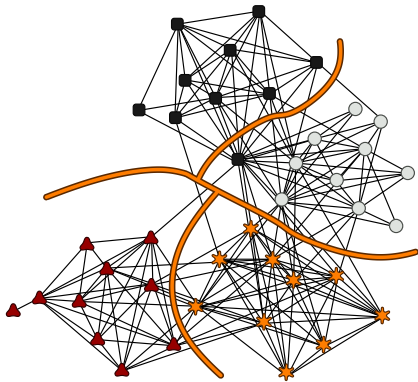
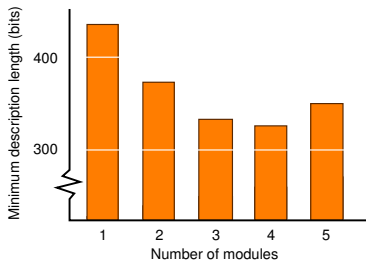


Community detection is dual to optimal coding



$$L = L(Y) + L(X|Y) = n \log m + \frac{1}{2} m(m+1) \log l + \log \left[\prod_{i=1}^m \binom{n_i(n_i-1)/2}{l_{ii}} \prod_{i>j} \binom{n_i n_j}{l_{ij}} \right]$$

Minimum description length — shortest description of model and uncertainty



Social and biological systems are

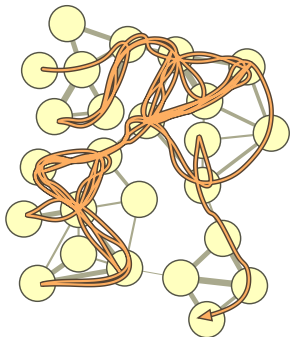
differentiated

multipartite

integrated

dynamic

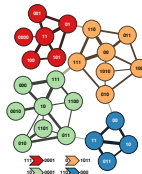
Maps by compressing flow of information on networks



```

1111100 1100 0110 11011 10000 11011 0110 0011 10111 1001
0011 1001 0100 0111 10001 1110 0111 10001 0111 1110 0001
1110 10001 0111 1110 0111 1110 1111010 1110 0000 10100 0001
1110 10001 1111 0110 10110 10110 11011 1001 0010 1000 10111
1001 0100 1001 0100 0011 1000 0011 0110 11011 0110 0011 0100
1001 10111 0111 0100 0111 10001 1110 10001 0111 1010 10110
111111 10110 10011 11110 00011

```



```

111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 1
0011 10 111 000 10 111 000 111 10 011 10 000 111 10 111
0010 10 011 010 011 10 001 111 0001 0 111 010 100 011 00 1
00 010 00 111 00 011 110 101 110 101 111 01 101 01 0001 0 1
111 00 011 110 111 1011 10 111 000 10 000 111 0001 0 111 0
0010 010 1011 100 00 10 011

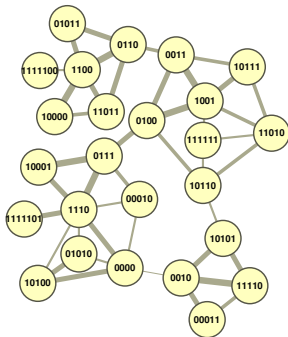
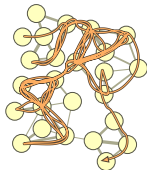
```



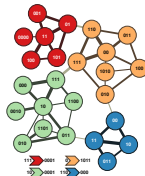
111 0000 11 01 101 100 101 01 0001 110 011 00 110 00 11
000 111 000 10 111 0000 111 10 1 10 0000 111 10 111 11
0010 10 011 010 101 100 111 0000 111 010 100 011 010 11
00 011 00 111 00 111 110 111 110 1011 111 01 101 01 0001 11
111 00 011 110 111 1011 111 011 0000 111 0001 11 11 11
1100 000 1011 110 01 11

Describing flow is a coding game

— objects deserve unique names



1111100 1100 0110 11011 10000 11011 0110 0011 10111 1001
0011 1001 0100 0111 10001 1110 0111 10001 0111 1110 0000
1110 10001 0111 1110 0111 1110 1111101 1110 0000 10100 0000
1110 10001 0111 0100 10110 11010 10111 1001 0100 1001 10111
1001 0100 1001 0100 0011 0100 0011 0110 11011 0110 0011 0100
1001 10111 0011 0100 0111 10001 1110 10001 0111 0100 10110
111111 10110 10101 11110 00011



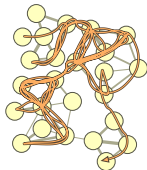
111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111
1001 10 111 000 10 111 000 111 10 011 10 000 111 10 111 10
0010 10 011 010 011 10 000 111 0001 0 111 010 100 011 00 111
00 011 00 111 00 111 110 111 110 0011 110 00 101 00 0001 0 110
111 00 011 110 111 1011 10 111 000 111 0001 0 111 010
1010 010 1011 110 00 10 011



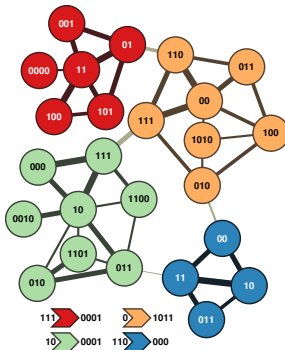
111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111
1001 10 111 000 10 111 000 111 10 011 10 000 111 10 111 10
0010 10 011 010 011 10 000 111 0001 0 111 010 100 011 00 111
00 011 00 111 00 111 110 111 110 0011 110 00 101 00 0001 0 110
111 00 011 110 111 1011 10 111 000 111 0001 0 111 010
1010 010 1011 110 00 10 011

Simplify and highlight

— important objects deserve unique names



111100 1000 0110 1101 1000 1101 0110 0011 1011 1001
0011 1000 0110 1101 1000 1101 0110 0011 1011 1000
1110 1000 0111 1110 0111 1110 1110 1110 1000 1000 0000
1110 1000 0111 0110 1010 1110 1010 1010 1000 1000 1001
1001 0100 1001 0100 0011 0100 0011 0110 1101 0110 0011 0100
1001 1011 0011 0100 0111 1000 1110 1000 0111 0100 1010
10111 10110 1000 10110 0001



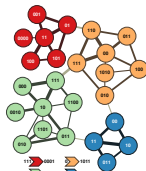
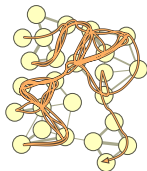
111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111
1011 10 111 000 10 111 000 111 10 011 10 000 111 10 111 10
0010 10 011 010 011 10 000 111 0001 0 111 010 100 011 00 111
00 011 00 111 00 111 110 111 110 1011 111 01 101 01 0001 0 110
111 00 011 110 111 1011 10 111 000 10 000 111 0001 0 111 010
1010 010 1011 110 00 10 011



111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111
1011 10 111 000 10 111 000 111 10 011 10 000 111 10 111 10
0010 10 011 010 011 10 000 111 0001 0 111 010 100 011 00 111
00 011 00 111 00 111 110 111 110 1011 111 01 101 01 0001 0 110
111 00 011 110 111 1011 10 111 000 10 000 111 0001 0 111 010
1010 010 1011 110 00 10 011

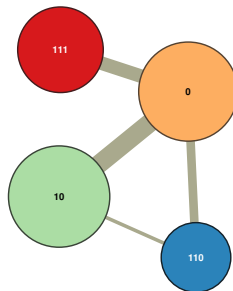
Simplify and highlight

— highlight important objects and filter away details



111100 1100 0110 11011 10000 11001 0110 0011 10111 1001
0011 1000 0100 0111 10001 1110 0111 10000 1110 1110 0000
1110 10001 0111 1110 0111 1110 11101 1110 0000 10100 0000
1110 10001 0111 0100 10110 11010 10111 1000 0100 1001 10111
1001 0100 1001 0100 0111 0100 0111 0110 11011 0110 0011 0100
1001 10111 0011 0100 0111 10001 1110 10000 0111 0100 10110
111111 10110 10100 1110 1110 00011

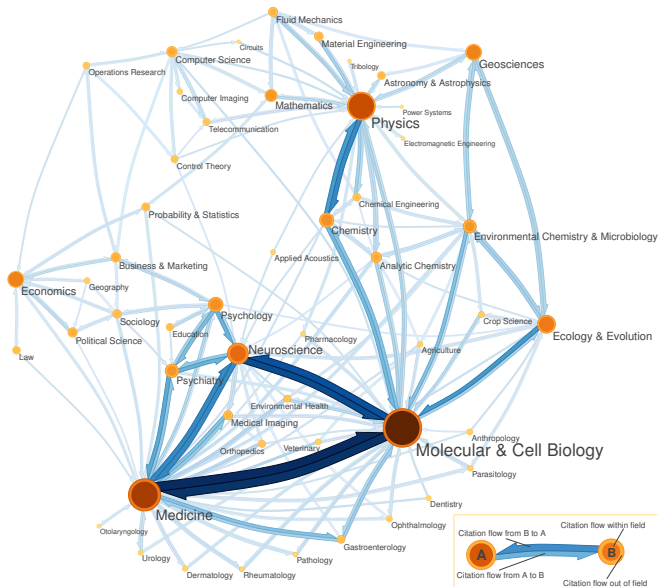
111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111
1011 10 111 000 10 110 000 111 10 011 10 000 111 10 111 10
0010 10 011 010 011 10 000 111 0001 0 111 010 100 011 00 111
00 011 00 111 110 110 110 1011 011 011 011 0001 0 110
111 00 011 110 111 1011 10 111 000 10 000 111 0001 0 111 010
1010 010 1011 110 00 10 011

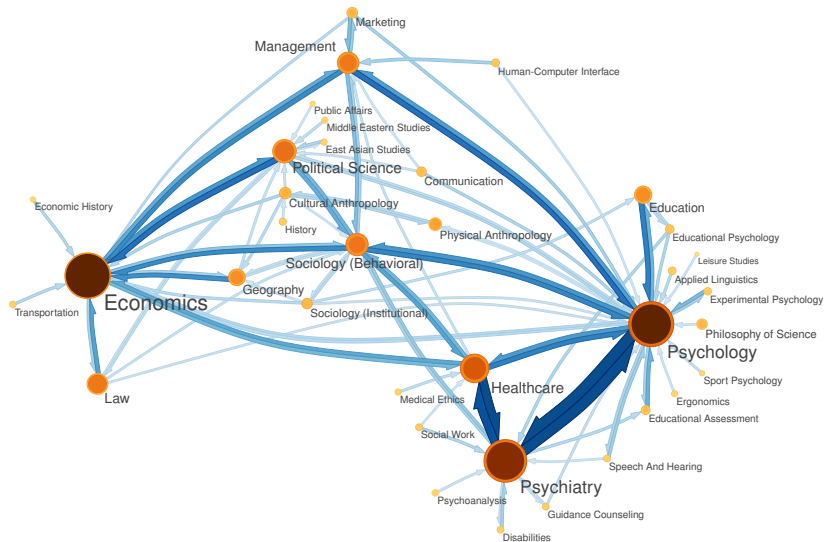


111 0000 11 01 101 100 101 01 0001 0 110 011 00 110 00 111
1011 10 111 000 10 111 000 111 10 011 10 000 111 10 111 10
0010 10 011 010 011 10 000 111 0001 0 111 010 100 011 00 111
00 011 00 111 00 111 110 111 110 1011 111 01 101 01 0001 0 110
111 00 011 110 111 1011 10 111 000 10 000 111 0001 0 111 010
1010 010 1011 110 00 10 011

The map equation

$$L(M) = q_{\curvearrowright} H(\mathcal{Q}) + \sum_{i=1}^m p_{\circlearrowleft}^i H(\mathcal{P}^i)$$

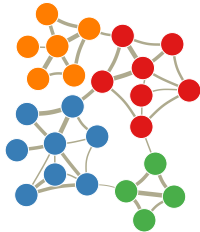
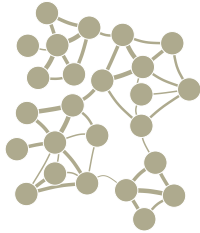




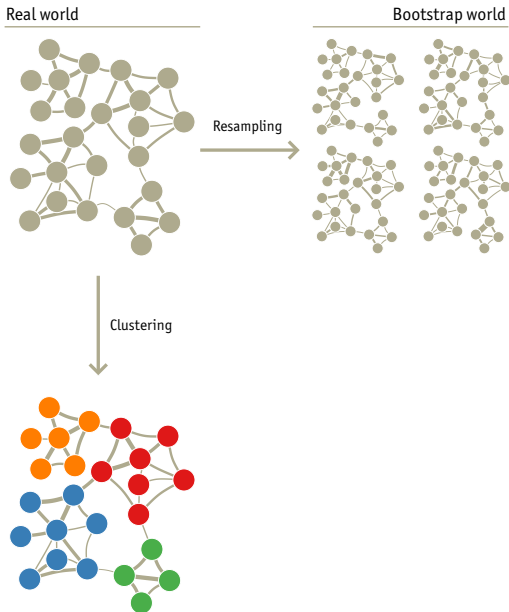
CHANGE

What is real **change** and what is mere statistical **noise**?

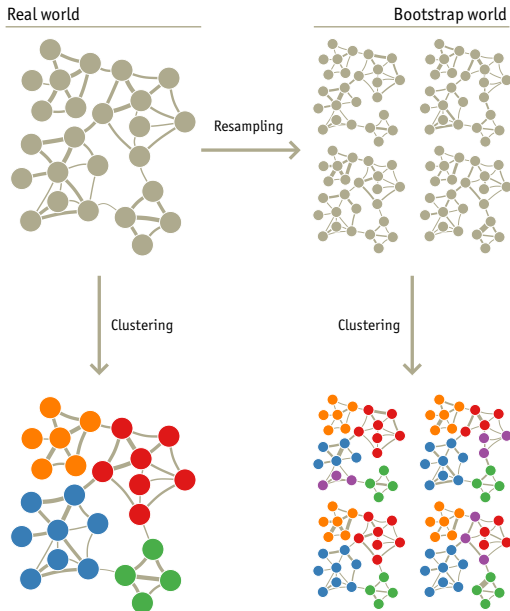
Significance clustering



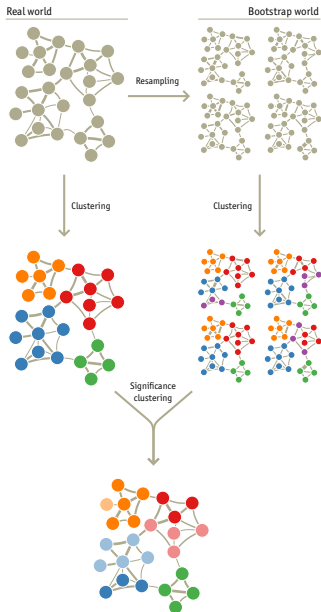
Significance clustering



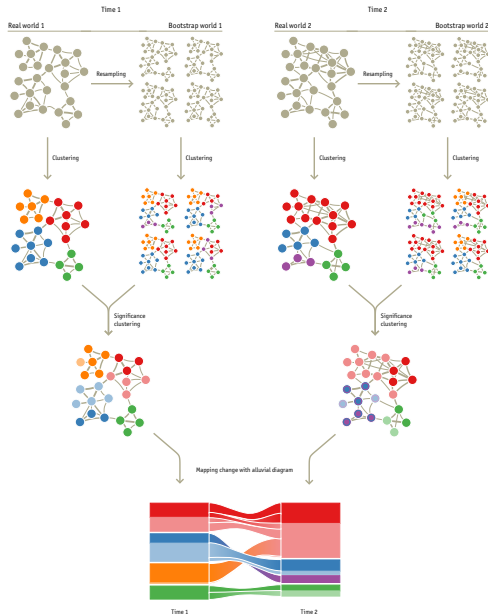
Significance clustering



Significance clustering



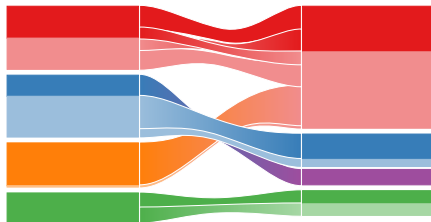
Mapping change



Mapping change



Mapping change with alluvial diagram



Time 1

Time 2

Mapping change in science

