Information propagation in social networks

Maksym Gabielkov, Ashwin Rao, Arnaud Legout

EPI DIANA, Sophia Antipolis

{maksym.gabielkov, arnaud.legout}@inria.fr

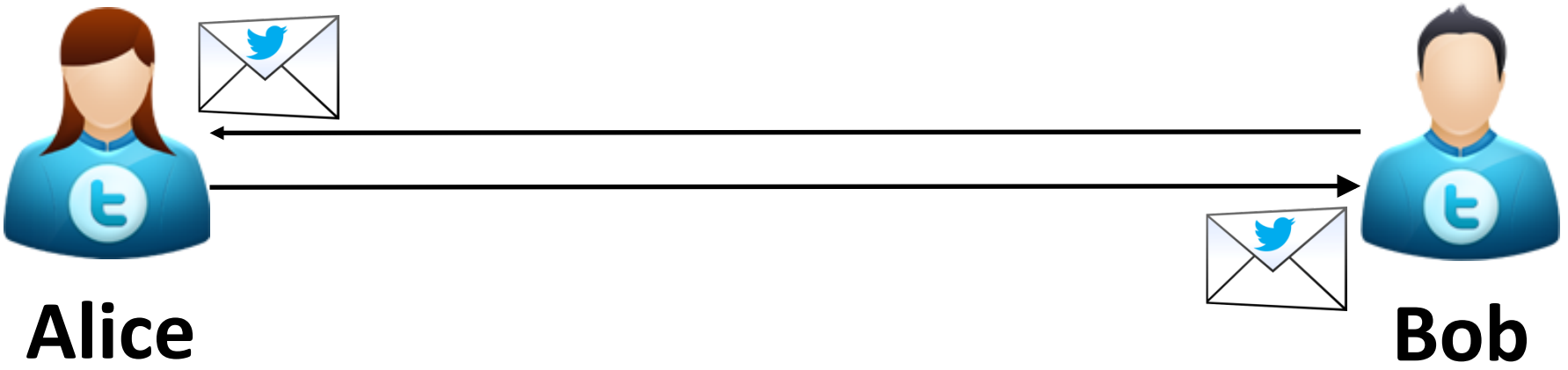
Friends

TF1
Producer
Consumers

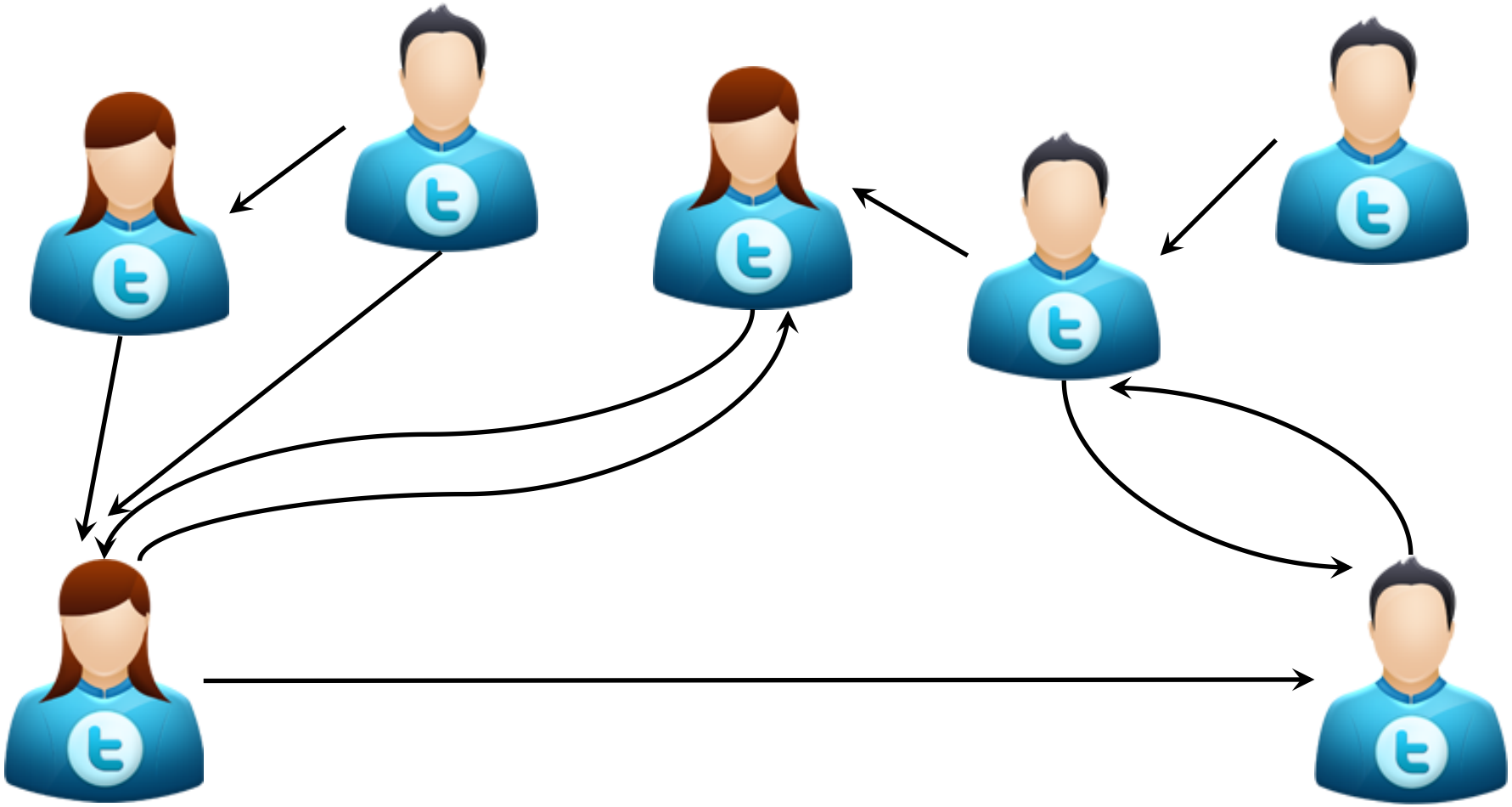
Follow Relationship in Twitter

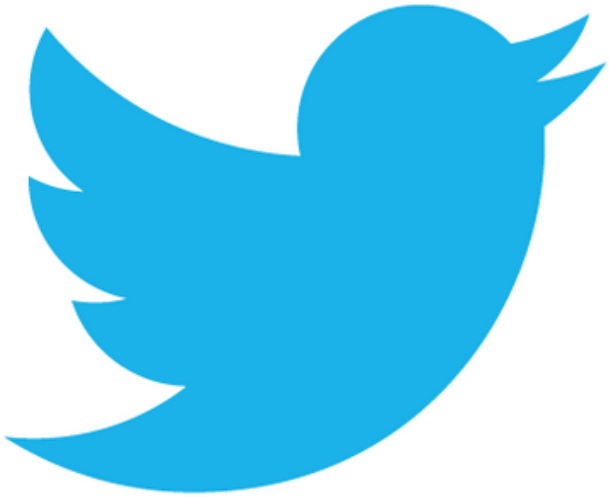
Bob follows Alice

Alice follows Bob



The Twitter Social Graph

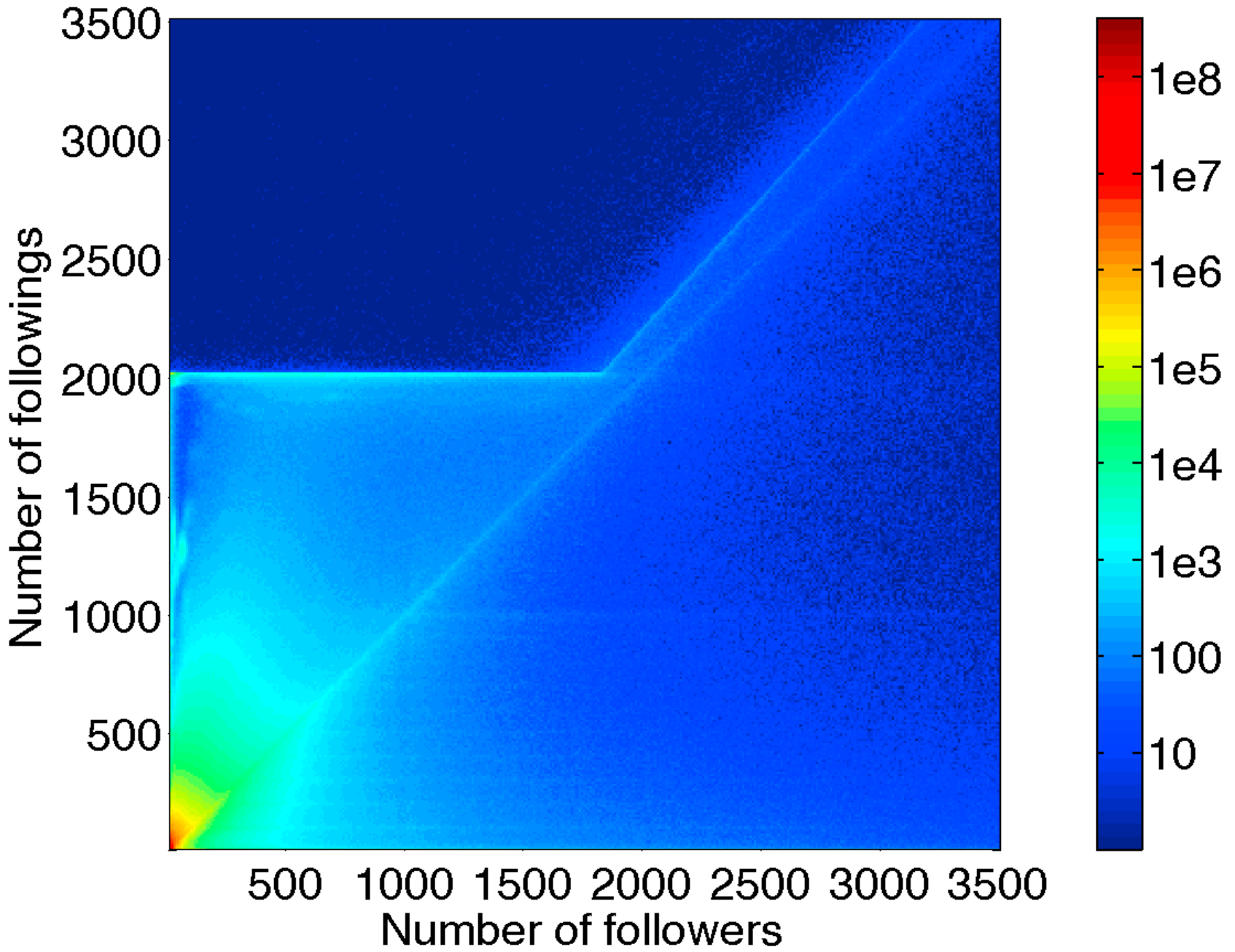




+500 million nodes
+24 billion edges

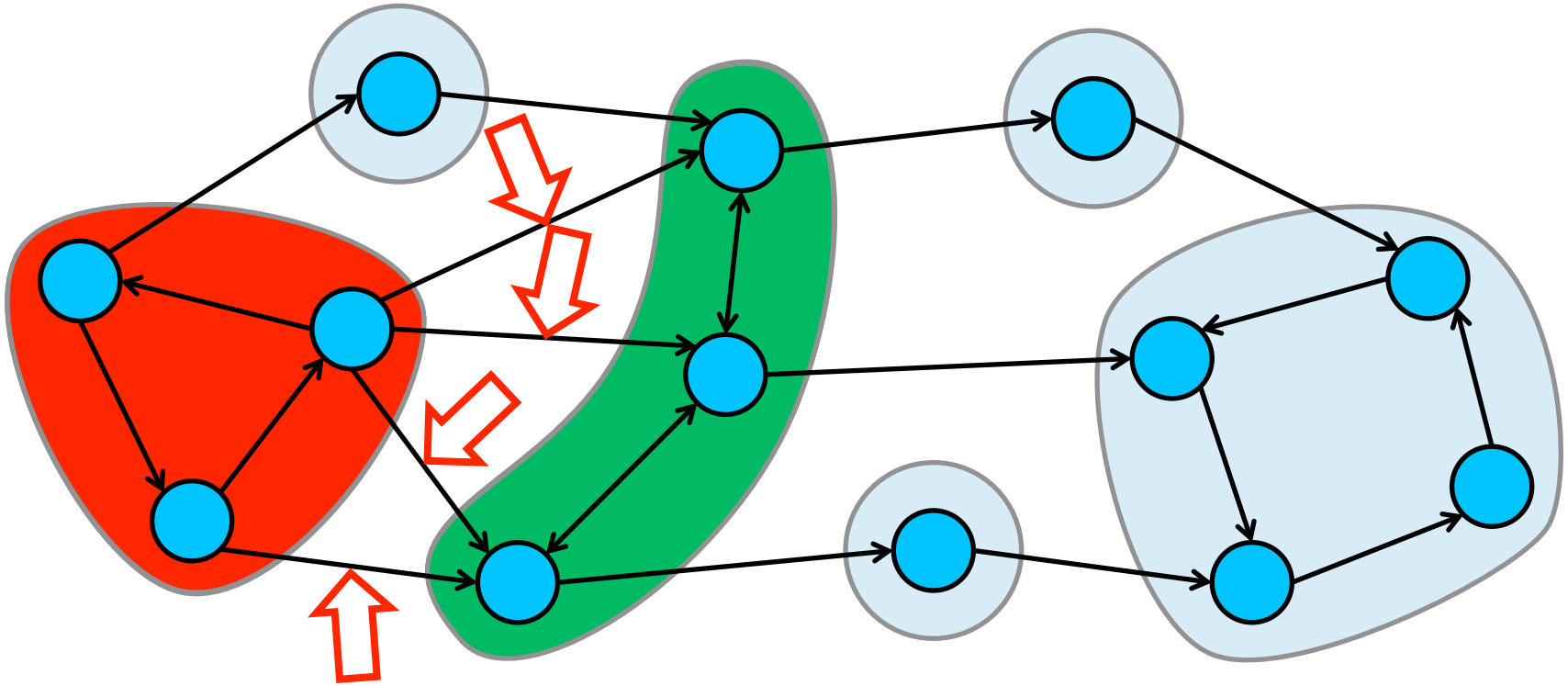
Challenges

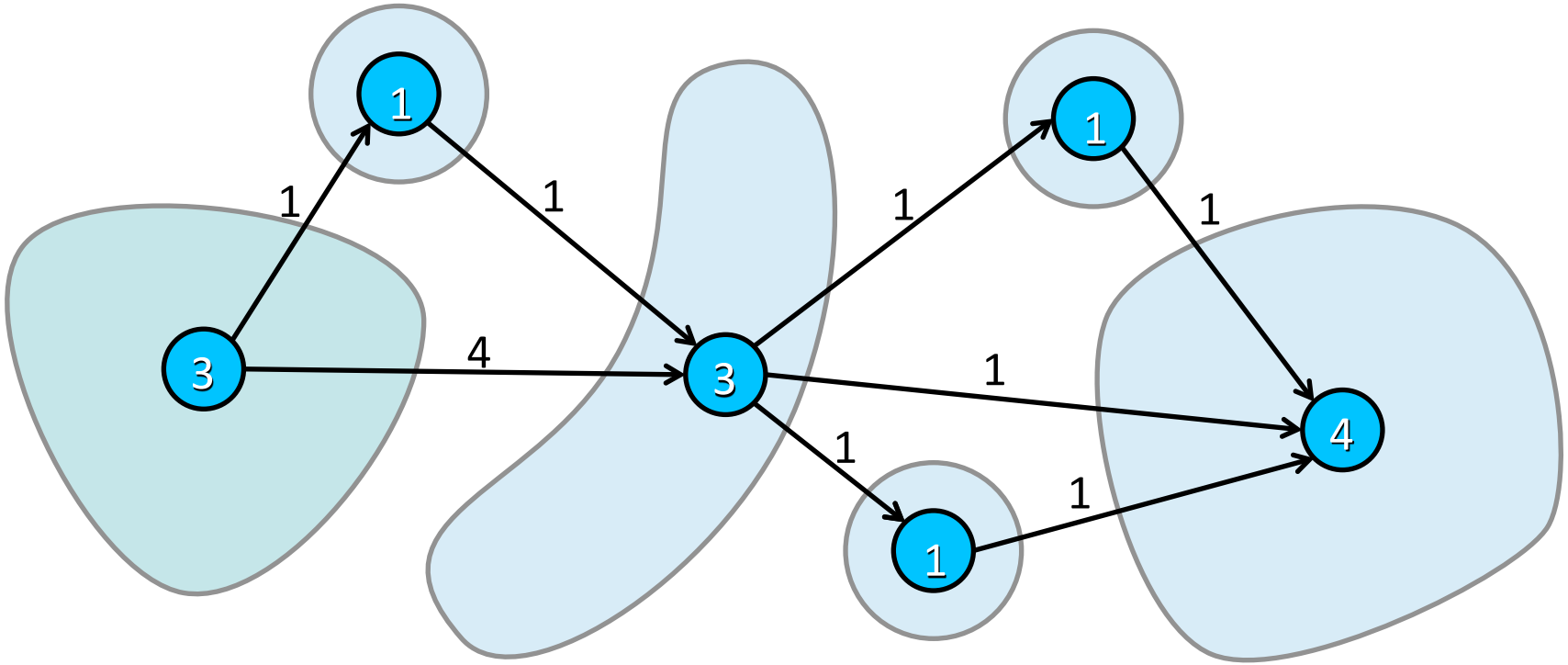
1. Collect the graph
2. Decompose the graph
3. Give a physical meaning to the decomposition

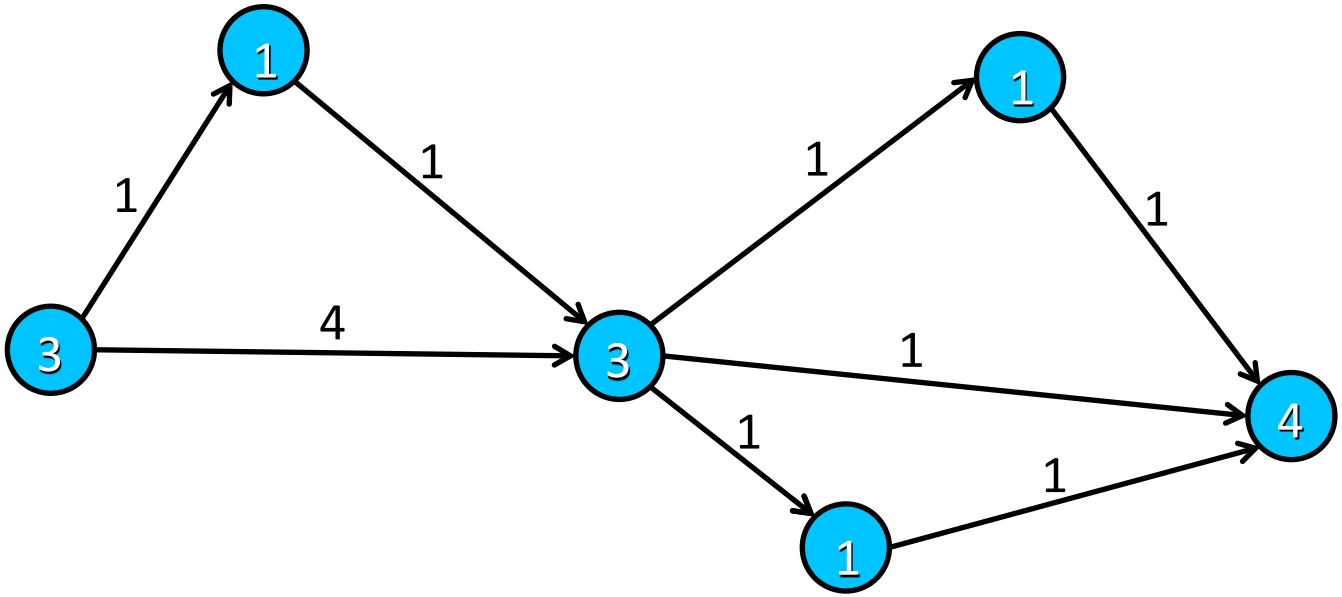

3500
3000
2500
2000
1500
1000
500
Number of followings
500 1000 1500 2000 2500 3000 3500
Number of followers
1e8
1e7
1e6
1e5
1e4
1e3
100
10

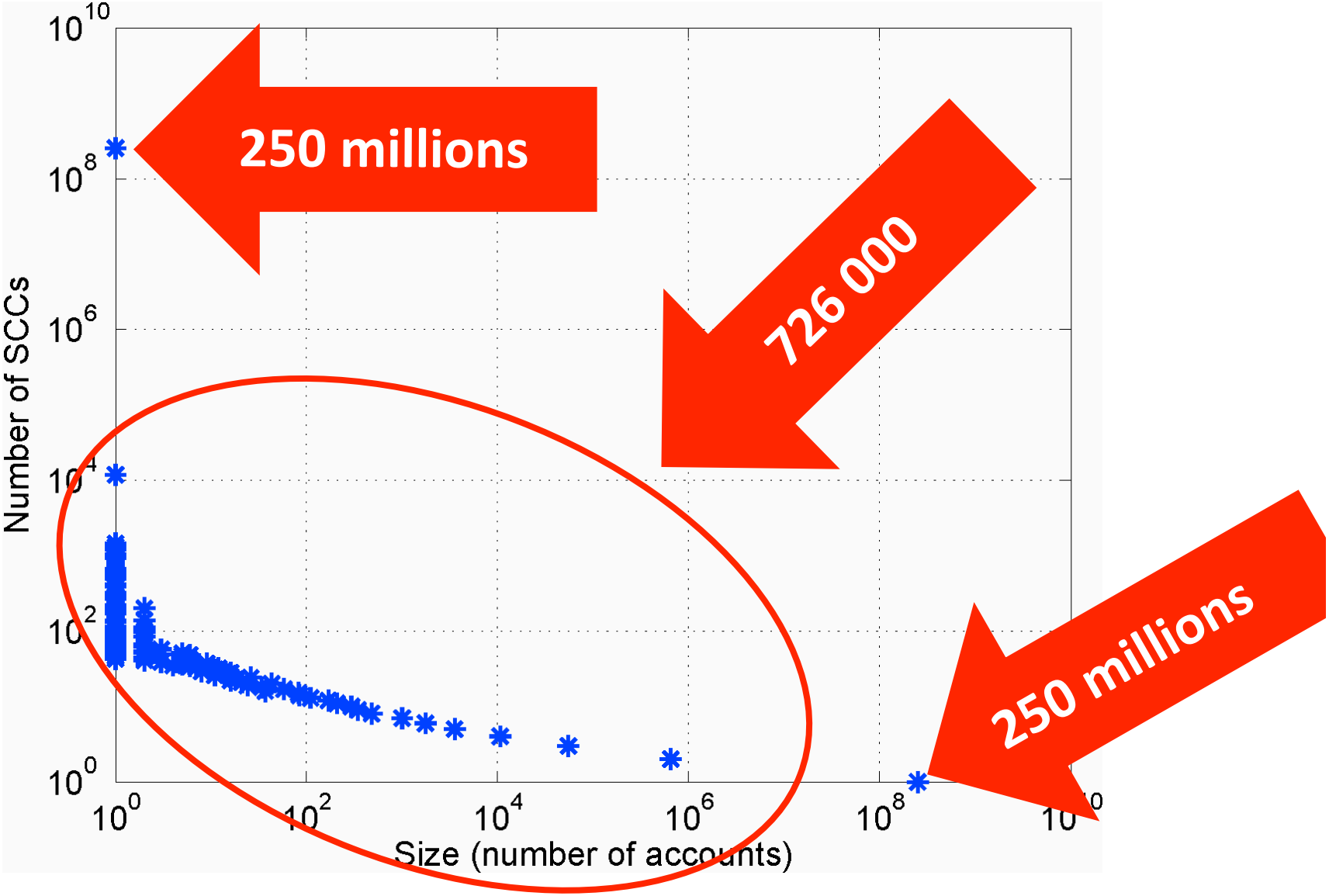
How is constraint information propagation?

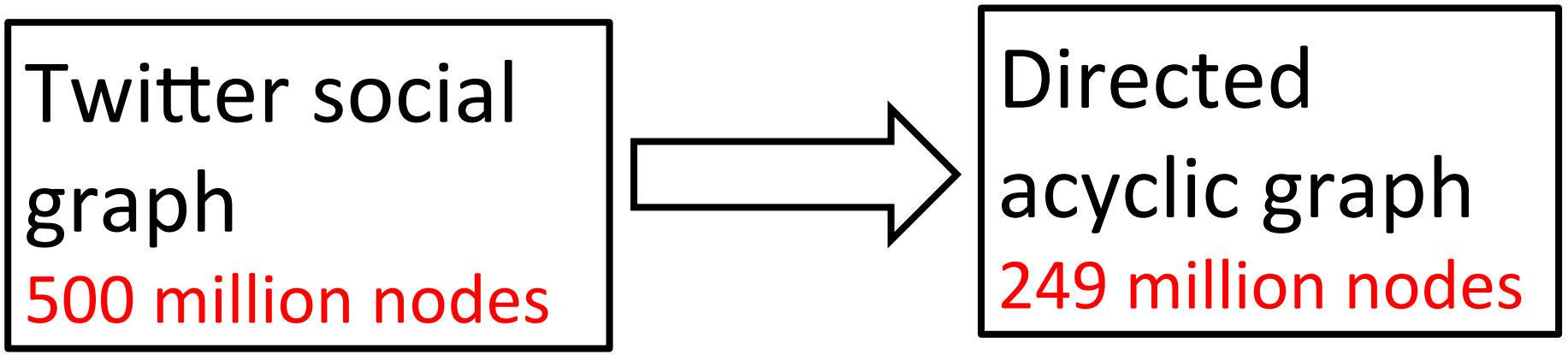
Identify the highways

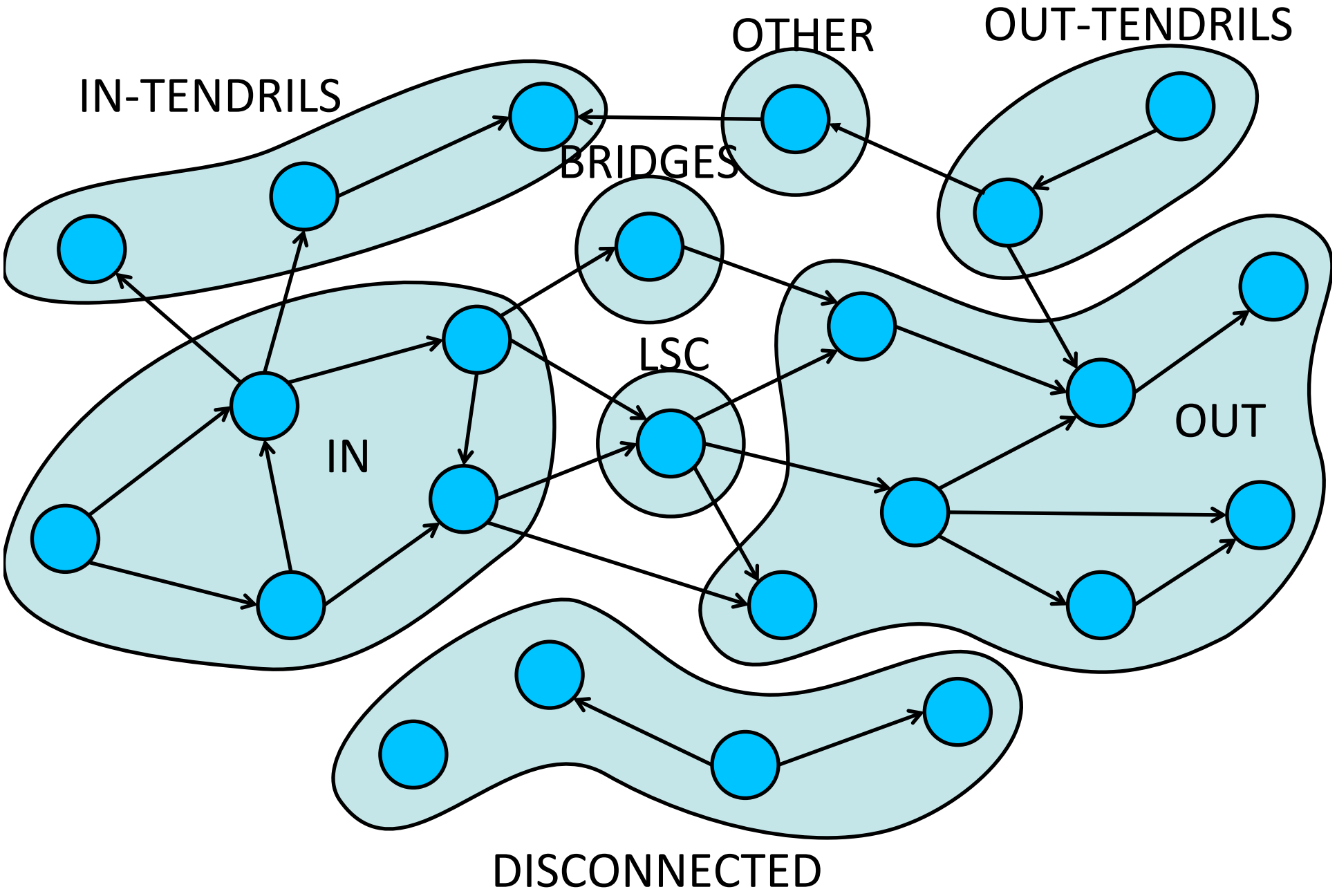


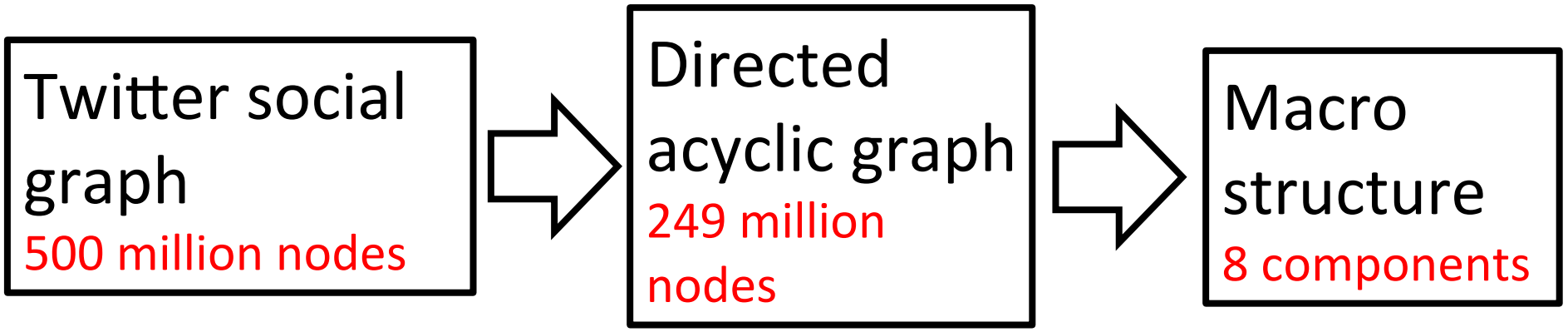

1
1
1
1
1
1
3
4
3
1
4
1
1


1
1
1
1
1
1
3
4
3
1
4
1
1

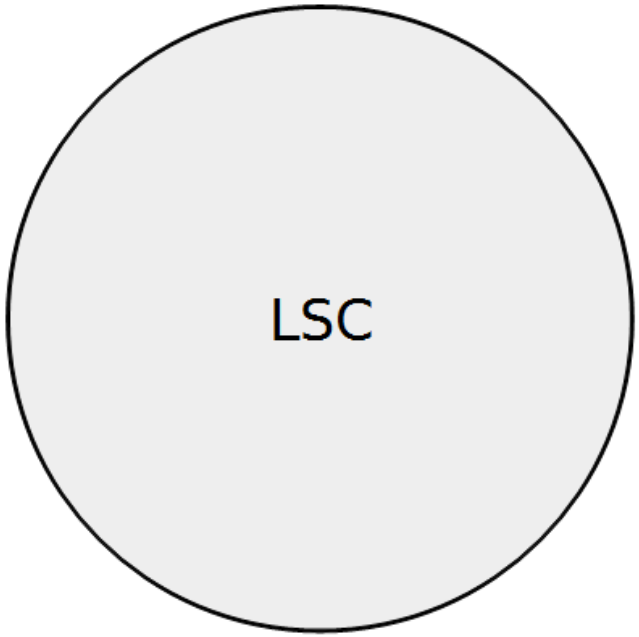

250 millions
726 000
250 millions
Number of SCCs
Size (number of accounts)
10^{10}
10^{8}
10^{6}
10^{4}
10^{2}
10^{0}
10^{0}
10^{2}
10^{4}
10^{6}
10^{8}

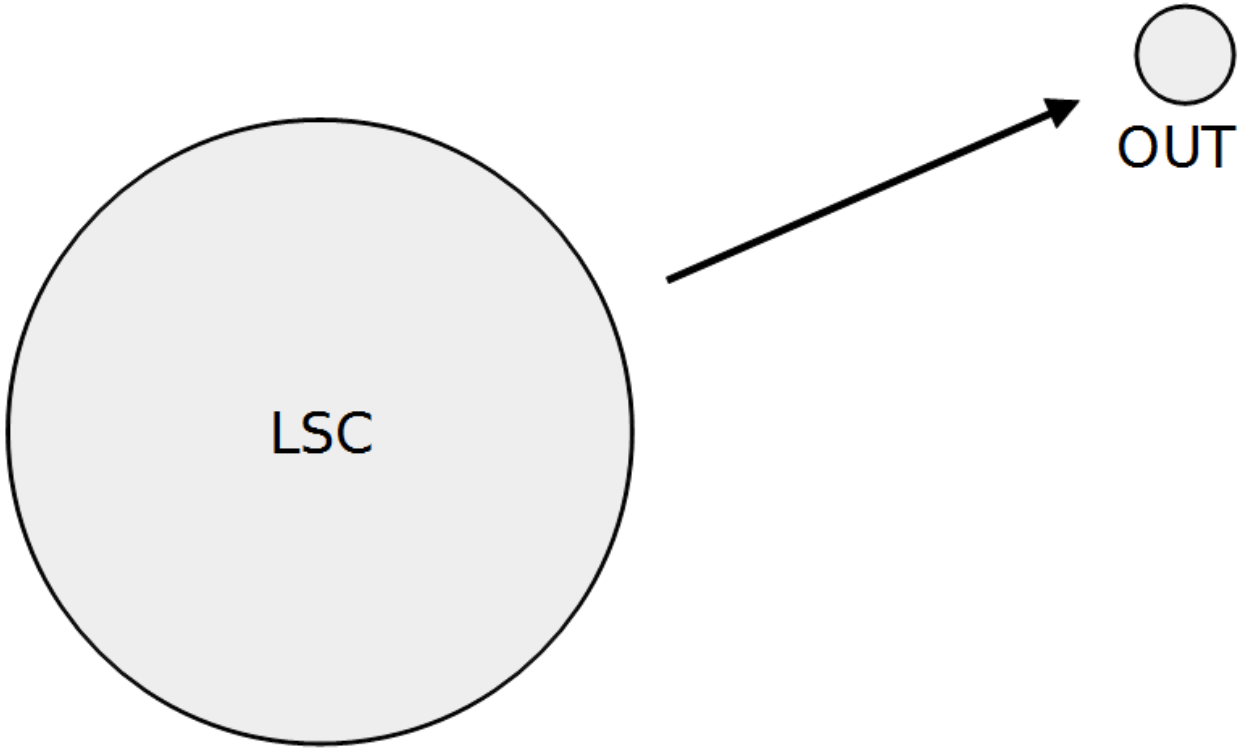

Twitter social graph
500 million nodes
Directed acyclic graph
249 million nodes

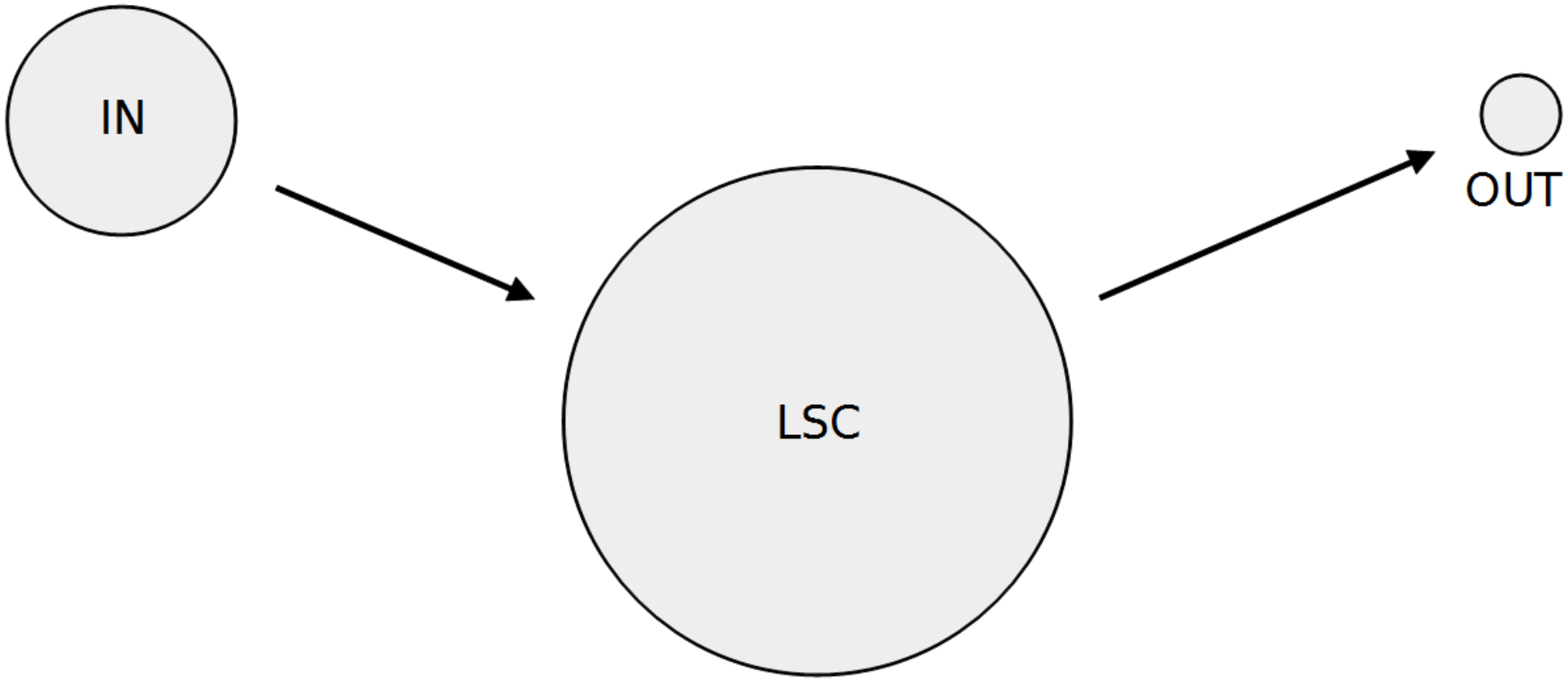

OTHER
OUT-TENDRILS
IN-TENDRILS
BRIDGES
LSC
IN
OUT
DISCONNECTED

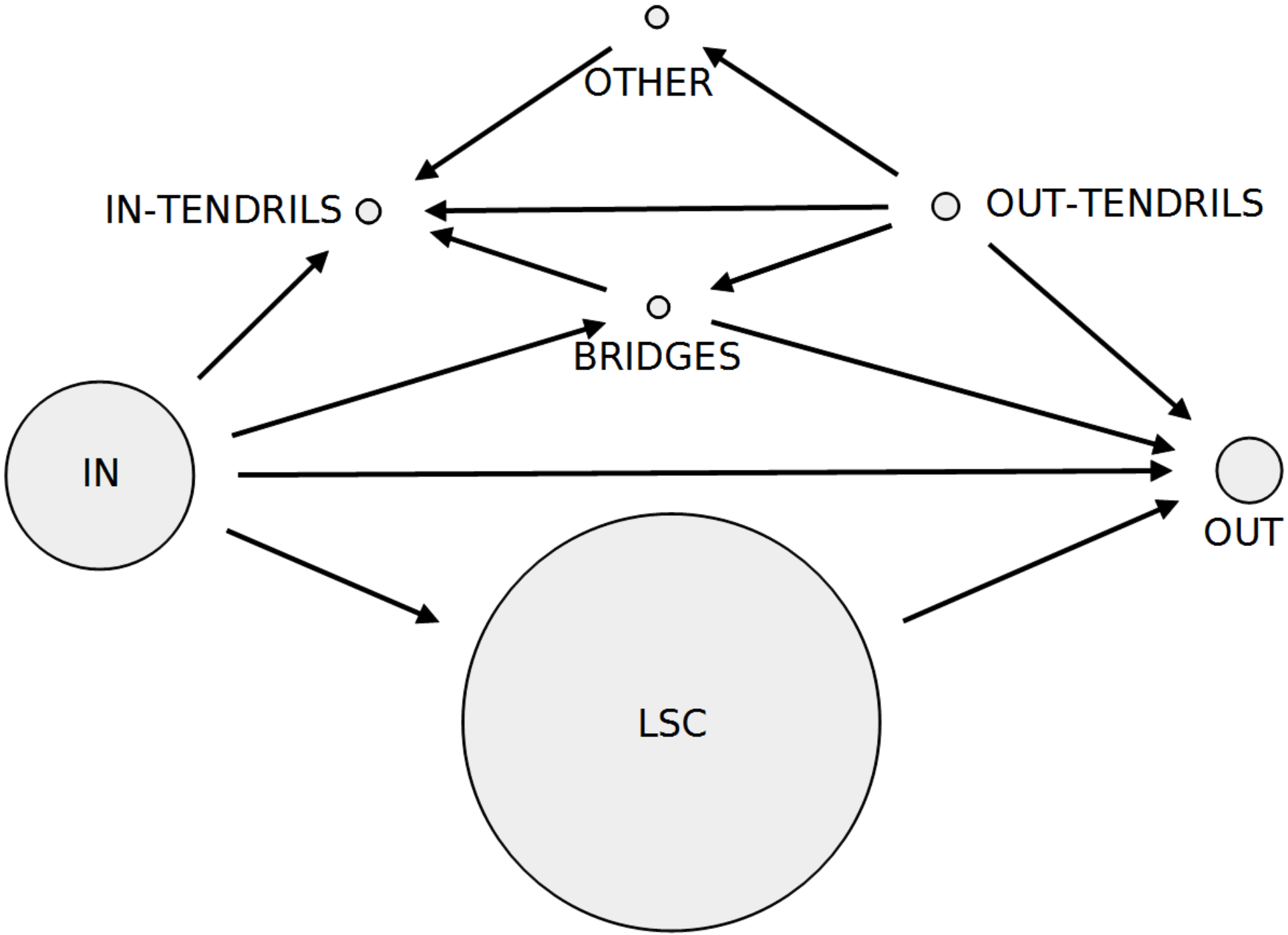

Twitter social graph
500 million nodes
Directed acyclic graph
249 million nodes
Macro structure
8 components

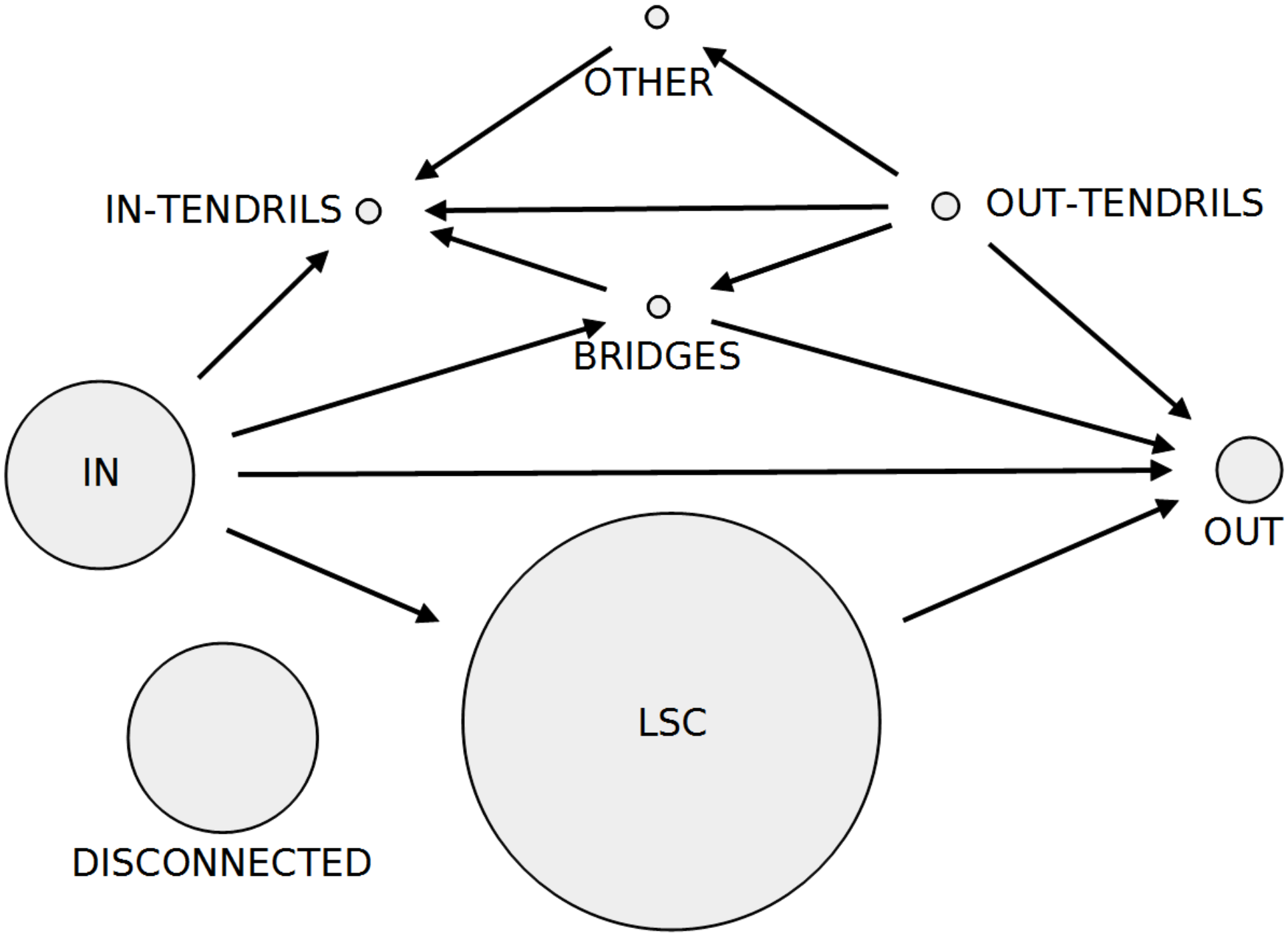
What is the physical meaning of decomposition?

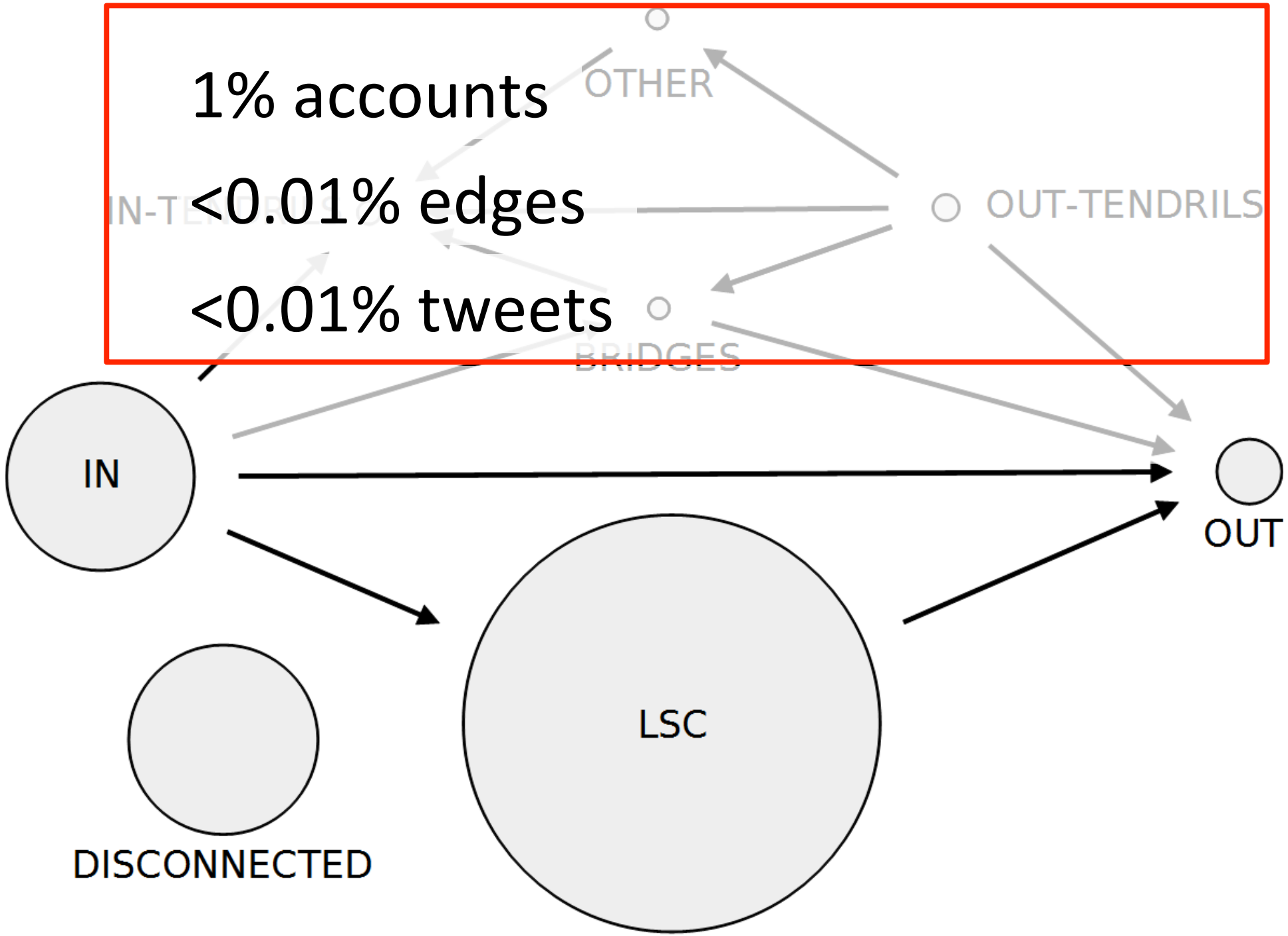

LSC


LSC
OUT


IN
LSC
OUT


OTHER
IN-TENDRILS
OUT-TENDRILS
BRIDGES
IN
OUT
LSC

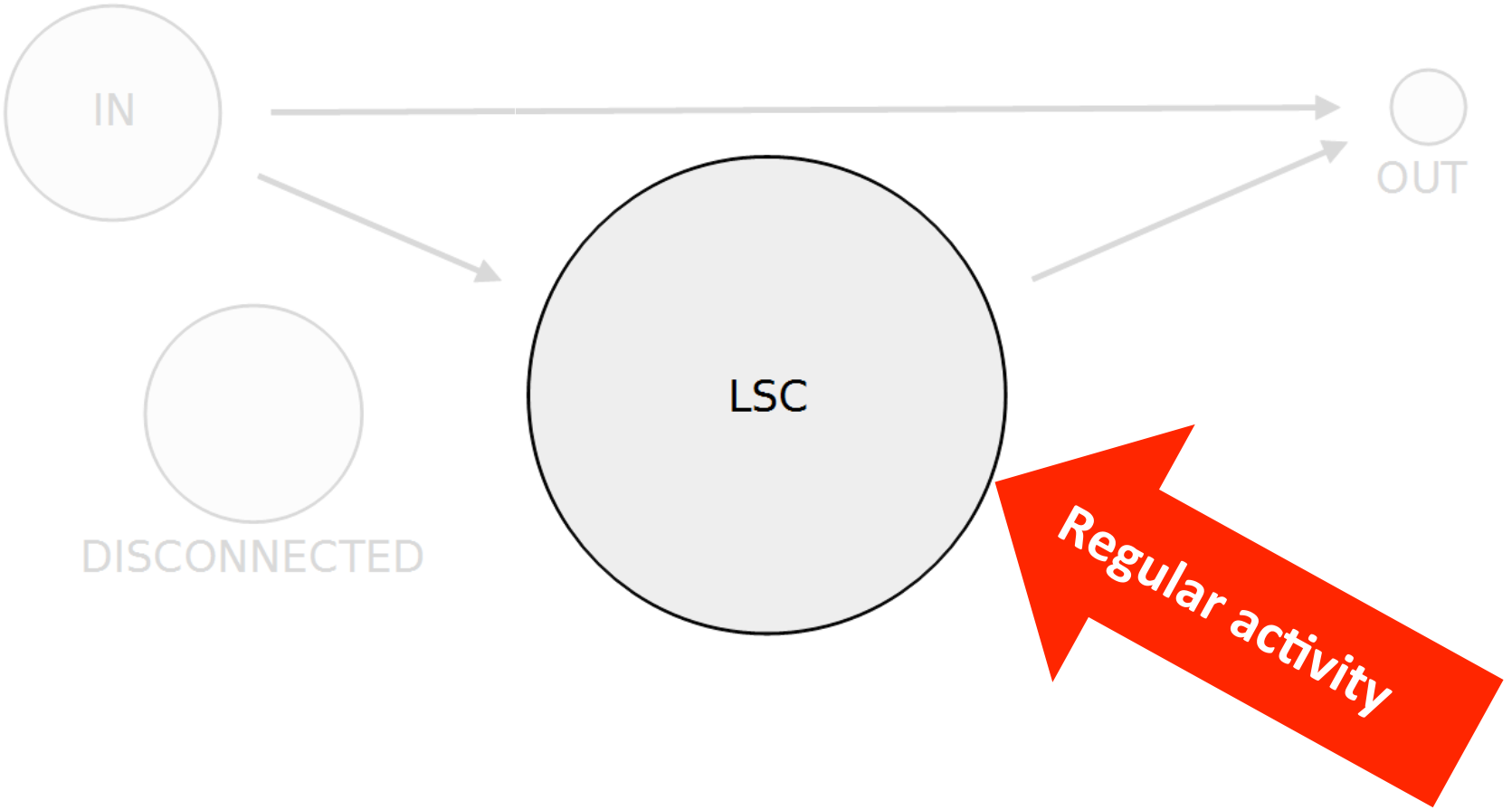

OTHER
IN-TENDRILS
OUT-TENDRILS
BRIDGES
IN
OUT
DISCONNECTED
LSC


1% accounts
<0.01% edges
<0.01% tweets
OTHER
OUT-TENDRILS
BRIDGES
IN
OUT
LSC
DISCONNECTED

98% of the tweets

98% of the edges

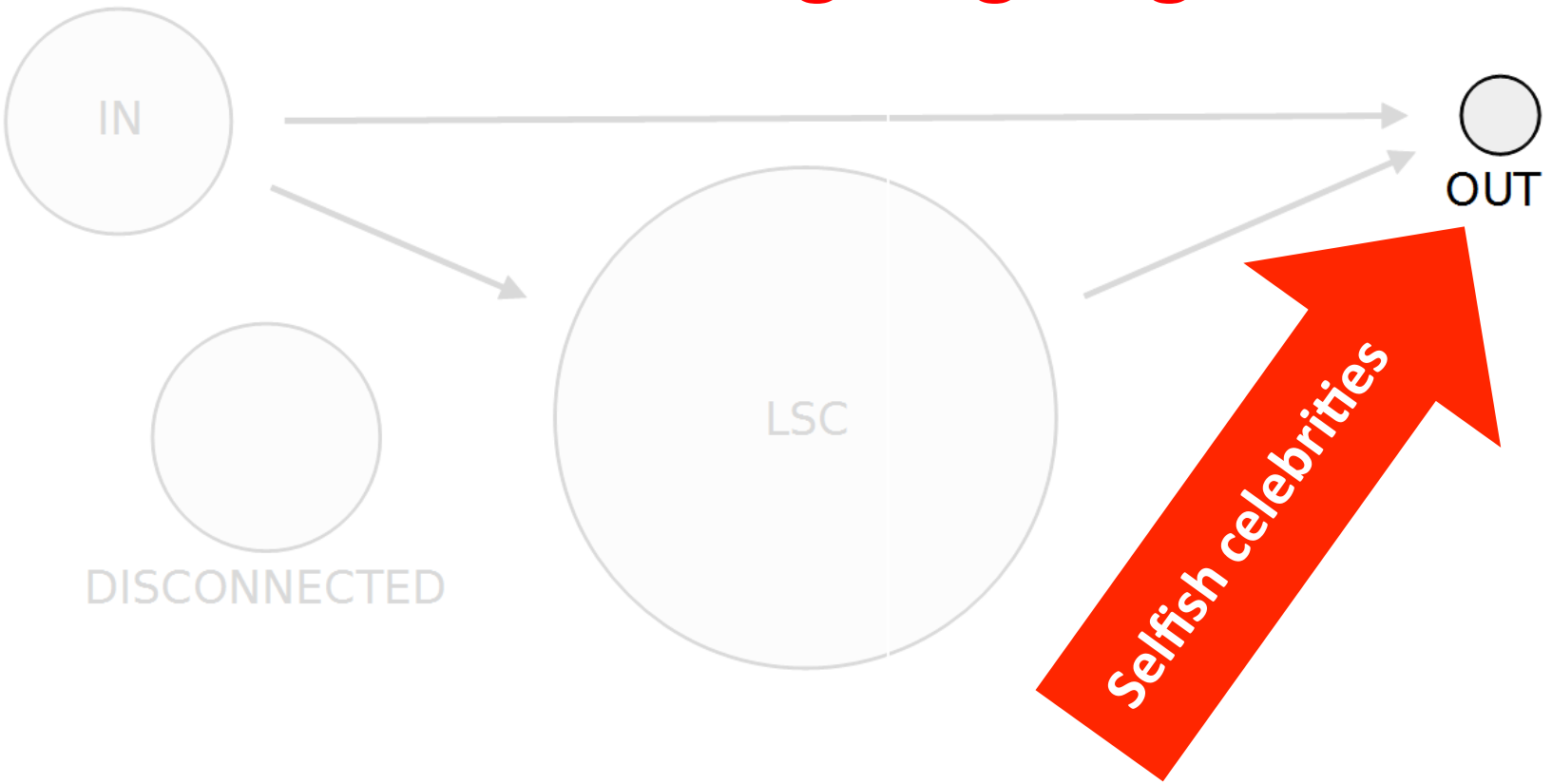
50% of the accounts



1,5% of the tweets

5,3% of the accounts

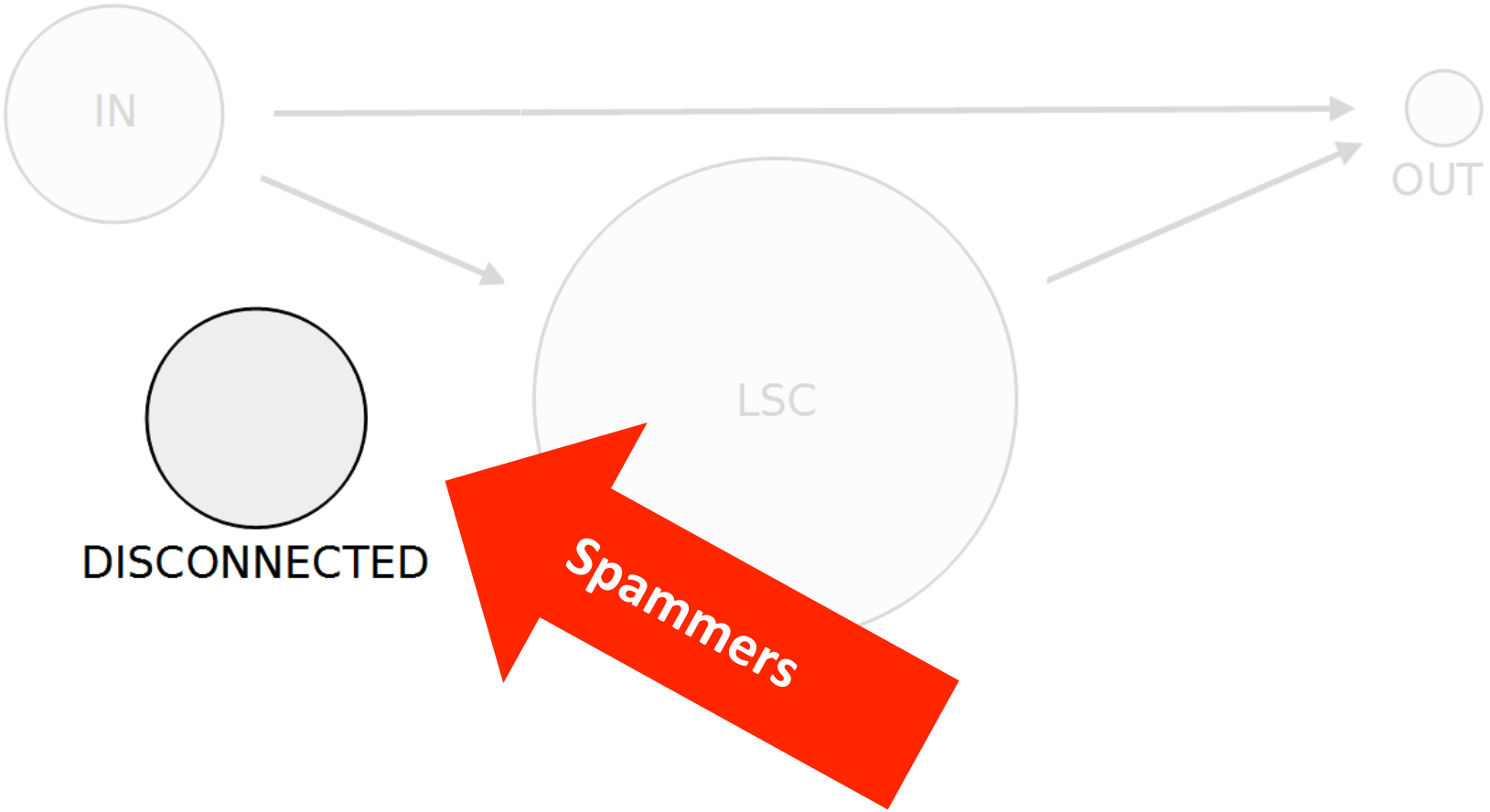
0% outgoing edges



21,4% of the accounts

0,25% of the tweets

21,6% of the accounts

99% no edge

80% no tweet