TOWARDS A BETTER UNDERSTANDING OF ADDICTION ON SOCIAL NETWORK: A METRIC COMPARISON

Graph Mining Presentation
(December 2018)

Lucia Nasti

lucia.nasti@di.unipi.it

UNIVERSITÀ DI PISA

INTRODUCTION

- **Aim:** to study the problem of Internet addiction and its spread through interaction on social network

Our proposal:

- a Hybrid Automata model of the Dopaminergic System, used to simulate the activity of a user type on a virtual social network
- To test different network topologies

THE MODEL OF DOPAMINERGIC SYSTEM

- **Dopamine concentration.** The equation describes the dynamics of variable *D* representing the dopamine concentration in the prefrontal cortex:

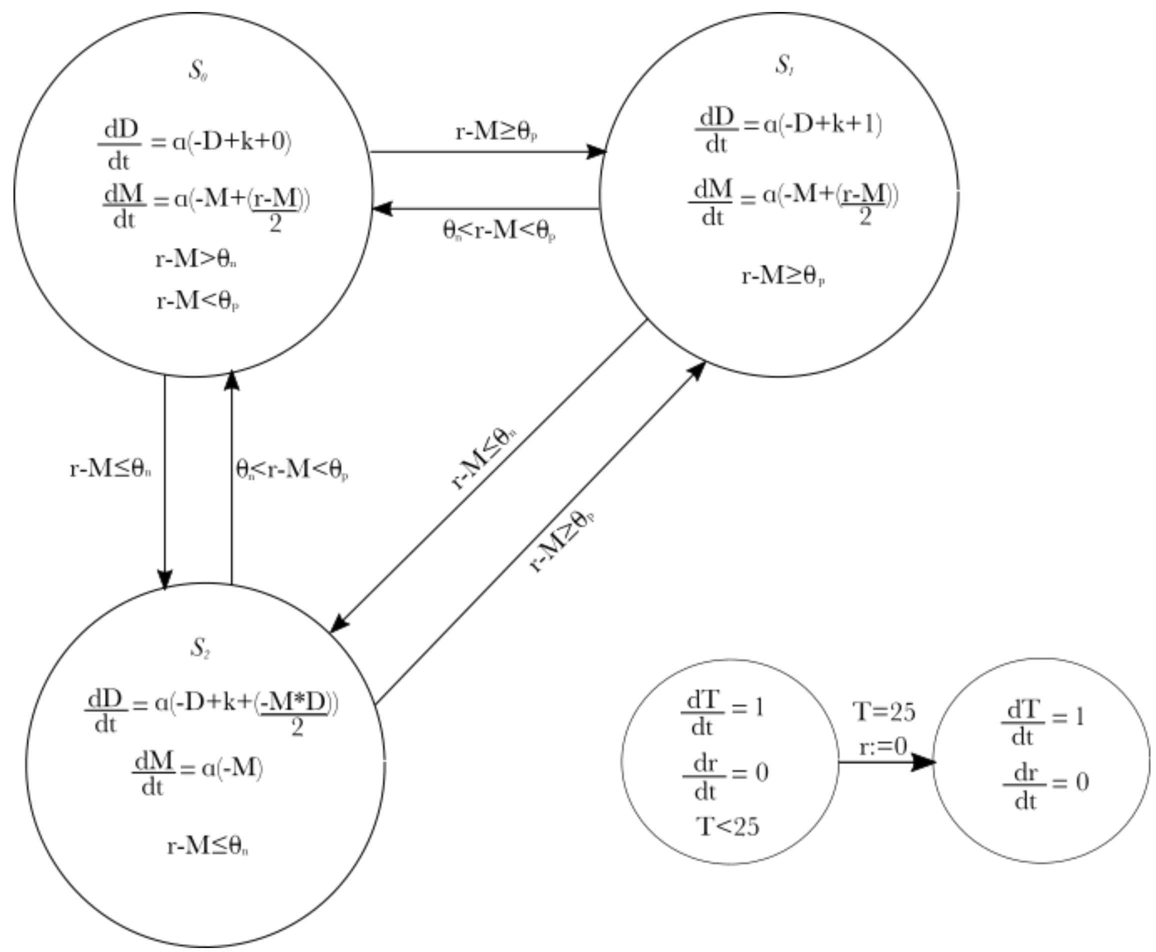
$$\frac{dD}{dt} = \alpha \left(-D + k + \begin{cases} 1, & if\ r - M \geq \theta_p \\ 0, & if\ \theta_n \leq r - M \leq \theta_p \\ -\frac{D*M}{2}, & if\ r - M \leq \theta_n \end{cases} \right)$$

- **Memory.** The equation describes the opponent process (a contrary emotional reaction to a previous stimulus) that is modelled as a memorisation process of previous stimuli.

$$\frac{dM}{dt} = \alpha \left(-M + \begin{cases} \frac{r-M}{2}, & if\ r > M \\ 0, & otherwise \end{cases} \right)$$

HYBRID AUTOMATA MODEL

Hybrid Automata are finite state automata in which states are associated to differential equations that describe the dynamics of a set of continuous variables.



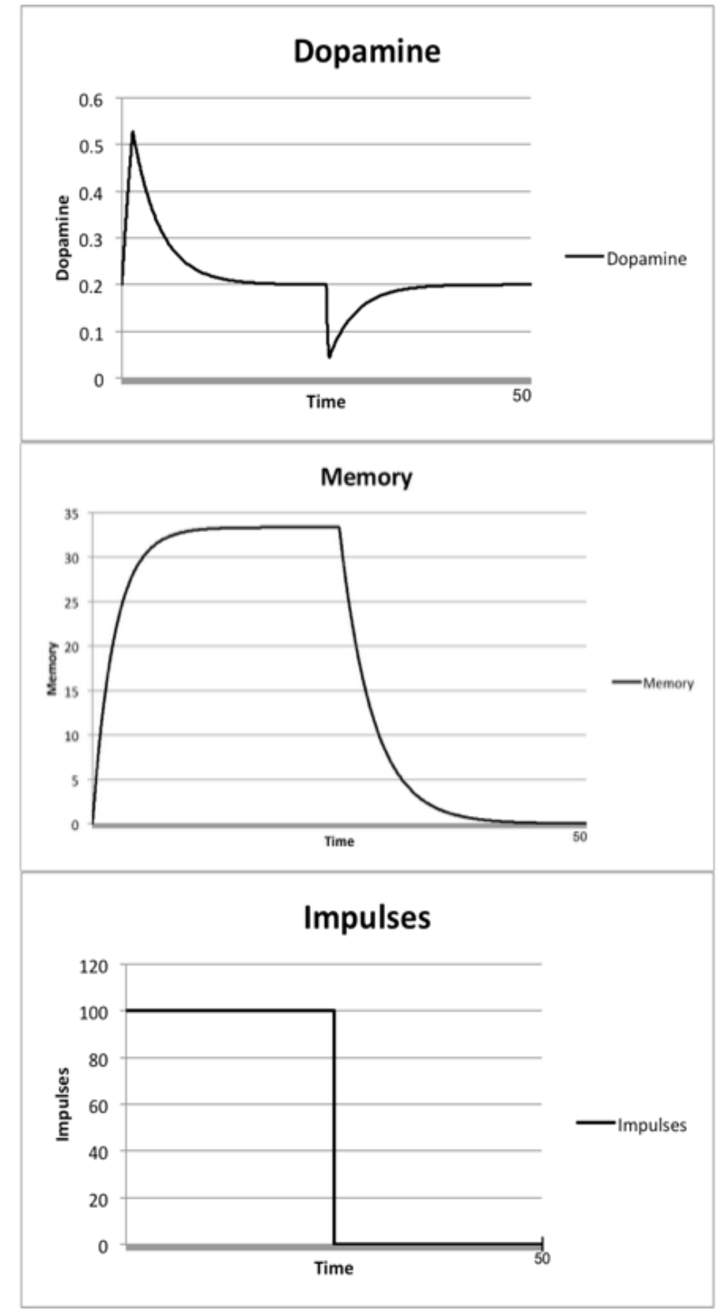
- Compositionality of Hybrid Automata
- Better description of transition

SIMULATION AT CONSTANT PULSE

The **trend of dopamine**, similar to the graph obtained by Gutkin, shows an initial peak which results in a withdrawal symptom, previous to the interruption of the stimulus itself.

The performance of the memory, however, corresponds to the **opponent process**.

To establish if a user became addicted, we consider the memory threshold: $M >= 15$



THE INTERNET ADDICTION

Excessive use of Internet as a mechanism to **escape** from the daily dissatisfaction.

Main expressions:

- Gaming
- Social network
- Surfing

We represent the social network as a **graph**:

- Each node of graph is a **user**
- Each user is characterised by the Dopaminergic System and the **propensity factor**

THE PROPENSITY FACTOR

- It is a parameter, which spans the range [0,1], to model the user's predisposition to communicate

- It influences the probability to send (or reply) messages through the social network

THE NETWORK COMMUNICATION MODEL

- Each user can send at most **one message each day**
- The probability to send a message is proportional to his/her propensity factor. If he/she sends a message, this is received by his/her neighbours
- On the same day, all users that have received one or more messages choose whether to reply or not, again with a probability proportional to their propensity factors.

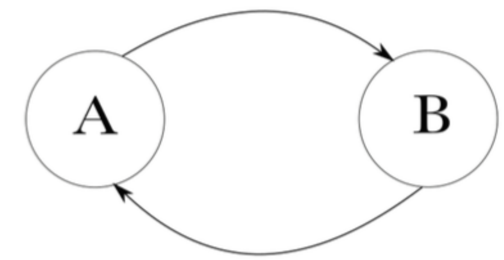
Stimuli:

- **medium-high:** when user receives a message
- **high:** when an addicted* user sends a message (*prediction error phenomenon*)

*Memory$\geq$15

GRAPH WITH ONLY TWO NODES

This graph is used to **examine the role of the propensity factor**. All the possible combinations of users have been tested to study the behaviour of the Dopaminergic System in different situations



User B

User A	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1
0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0
0.1	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=66	A=0; B=97	A=0; B=99
0.2	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=83	A=0; B=100	A=0; B=100	A=0; B=100	A=0; B=100	A=0; B=100	A=0; B=100
0.3	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=0	A=0; B=95	A=1; B=100	A=2; B=100	A=6; B=100	A=6; B=100	A=5; B=100	A=8; B=100
0.4	A=0; B=0	A=0; B=0	A=90; B=0	A=96; B=3	A=95; B=96	A=93; B=100	A=91; B=100	A=96; B=100	A=91; B=100	A=94; B=100	A=93; B=100
0.5	A=0; B=0	A=0; B=0	A=100; B=0	A=100; B=5	A=100; B=94	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100
0.6	A=0; B=0	A=0; B=0	A=100; B=0	A=100; B=6	A=100; B=96	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100
0.7	A=0; B=0	A=0; B=0	A=100; B=0	A=100; B=2	A=100; B=95	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100
0.8	A=0; B=0	A=51; B=0	A=100; B=0	A=100; B=2	A=100; B=99	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100
0.9	A=0; B=0	A=88; B=0	A=100; B=0	A=100; B=5	A=100; B=96	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100
1	A=0; B=0	A=99; B=0	A=100; B=0	A=100; B=2	A=100; B=96	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100	A=100; B=100

None One Both

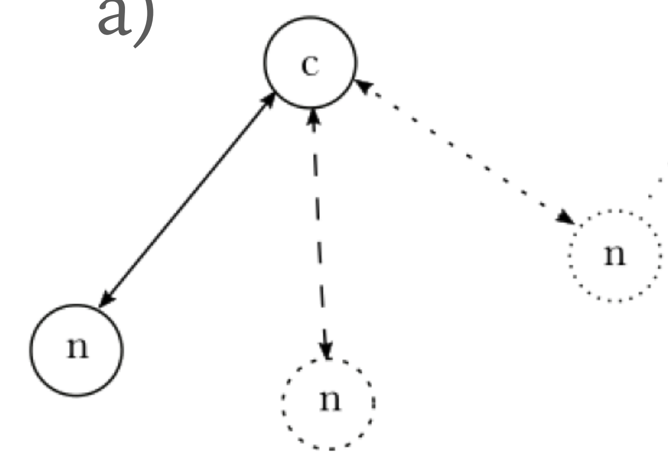
We find three values:

- low propensity: 0.2
- medium propensity: 0.35
- high propensity: 0.9

EXPERIMENTS WITH STAR GRAPH

a) **To study the propagation of addiction** in a graph, we start by considering how many peripheral nodes ***n*** are necessary to cause addiction of the central node ***c***, according to different propensity factor.

a)



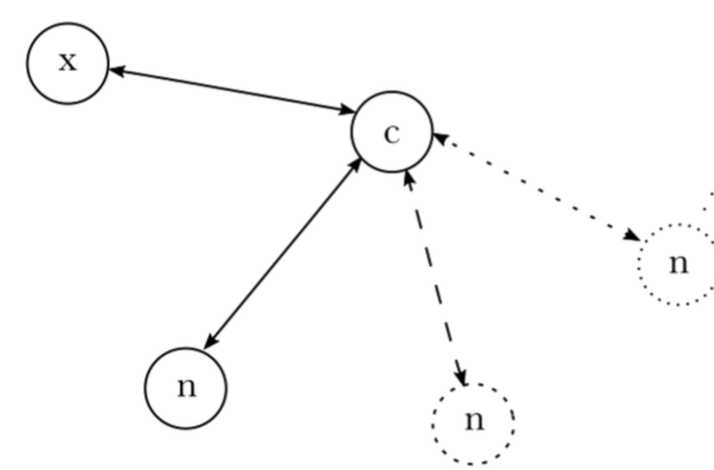
Nodes N

Node C	0.2	0.35	0.9
0.2	9	7	4
0.35	3	1	1
0.9	1	1	1

Few Nodes □ More Nodes □

b) After, we study how one of the peripheral nodes ***x*** whose propensity factor is 0.2 (fixed) can be influenced by the others (both ***c*** and ***n***).

b)



Nodes N

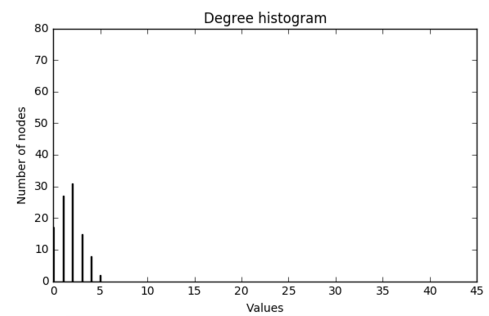
Node C	0.2	0.35	0.9
0.2	44	36	17
0.35	17	13	5
0.9	15	7	4

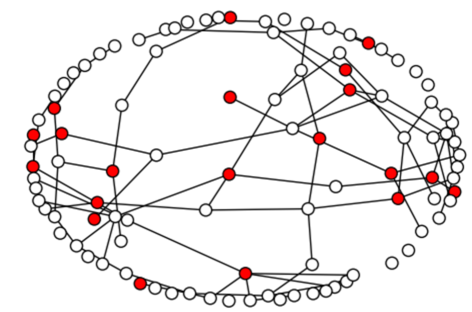
Few Nodes □ More Nodes □

EXPERIMENTS WITH RANDOM AND SCALE-FREE NETWORKS

For our experiments, we used three kinds of networks of 100 nodes, generated in different ways:

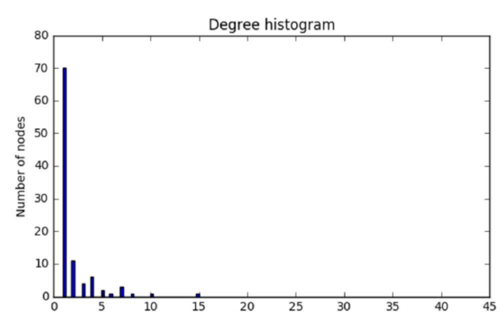
a) with the Erdős-Rényi graph (ER), approach:

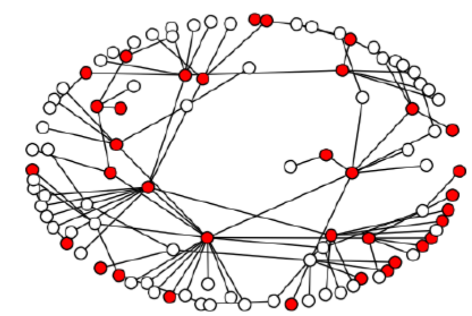




20 nodes addicted in ER

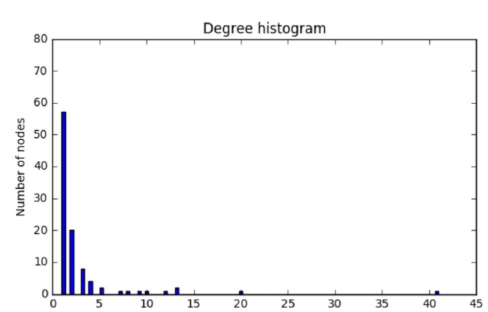
b) with the Barabási-Albert (BA) approach:

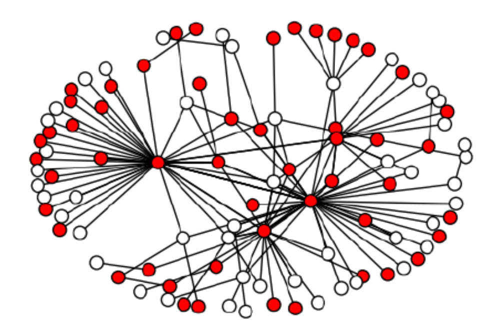




35 nodes addicted in BA

c) with the Bollobás-Riordan (BR) approach:





52 nodes addicted in BR

EXPERIMENTS WITH HUGER GRAPH

Real dataset information:

- This dataset consists of 'circles' (or 'friends lists') from Facebook (by SNAP)
- Number of nodes: 4039
- Number of edges: 88234

A scale free graph (Barabasi):

- Number of nodes: 4039
- Number of edges: 88334

A random graph (Erdős-Rényi):

- Number of nodes: 4039
- Number of edges: 88234

CLOSENESS CENTRALITY

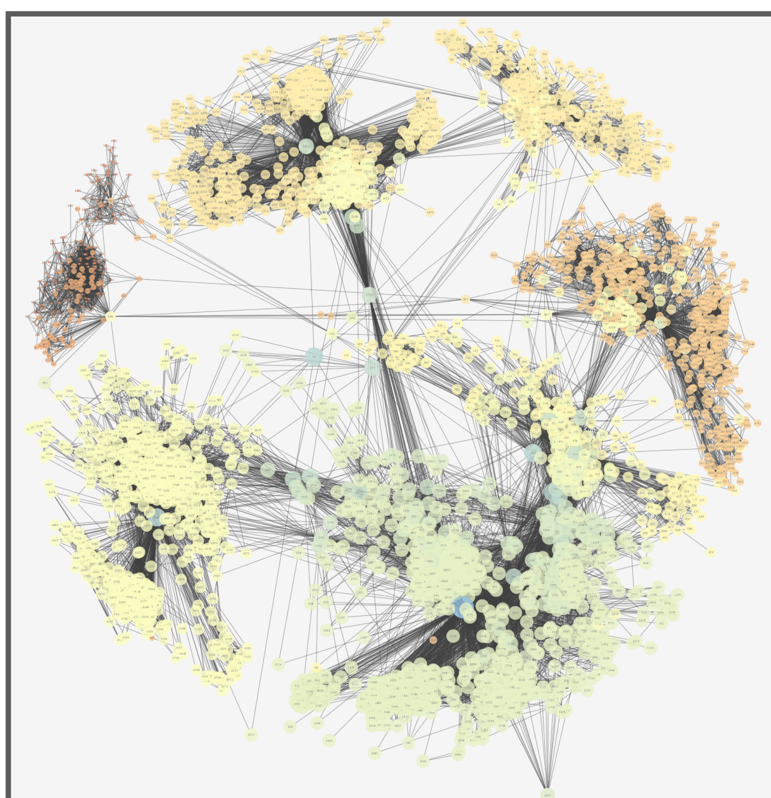
Closeness centrality can be regarded as a measure of how long it will take to spread information from *v* to all other nodes sequentially.

The more central a node is, the lower its total distance to all other nodes.

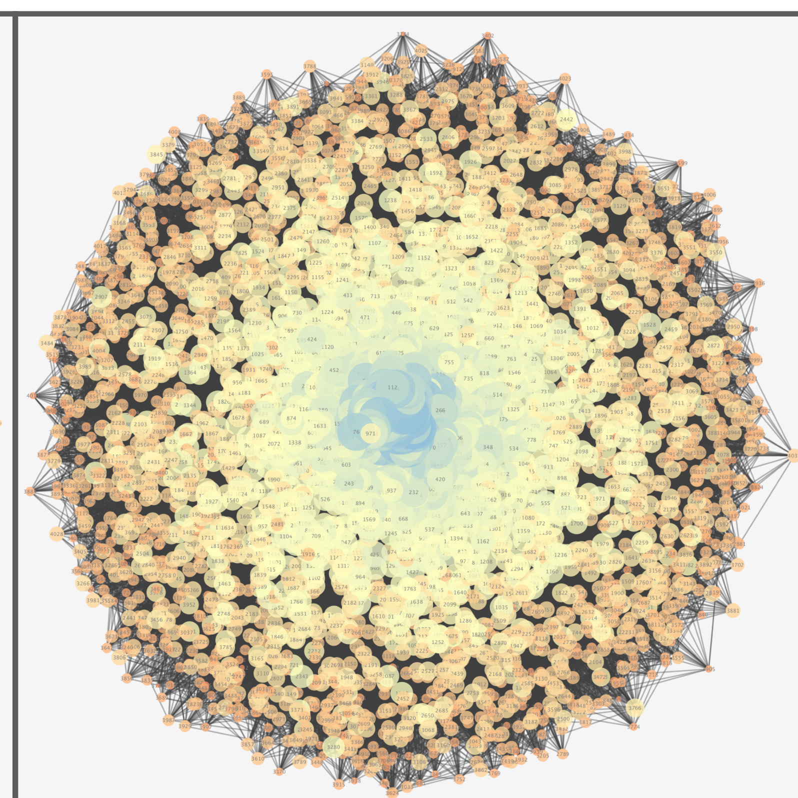
THE MAX VALUE OF CLOSENESS CENTRALITY:

Real Graph	Barabasi Graph	Random Graph
0.45	0.53	0.41

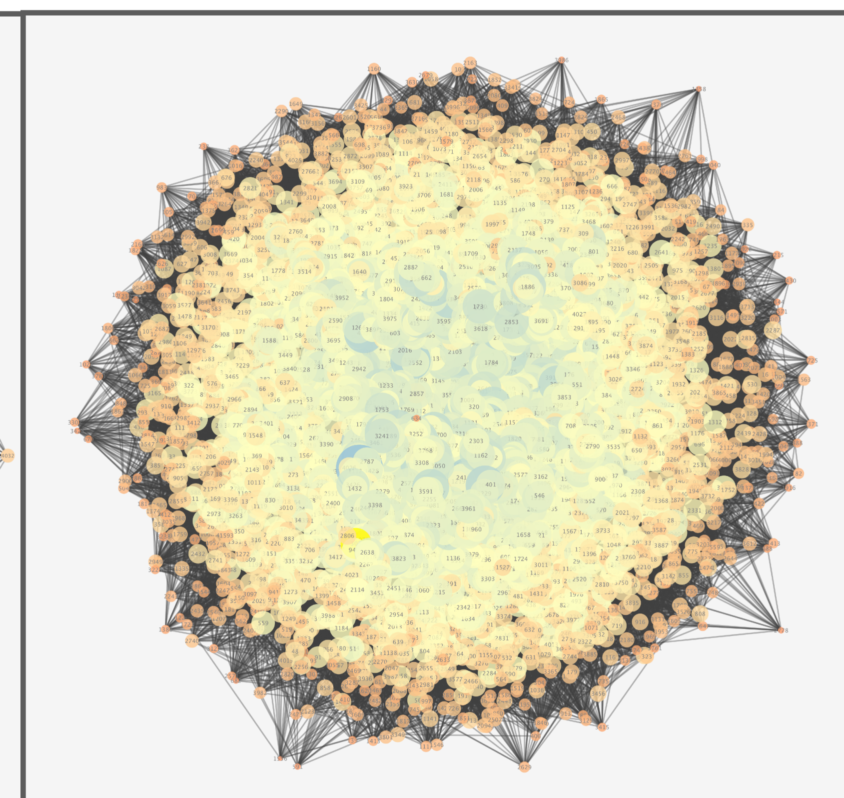
REAL GRAPH

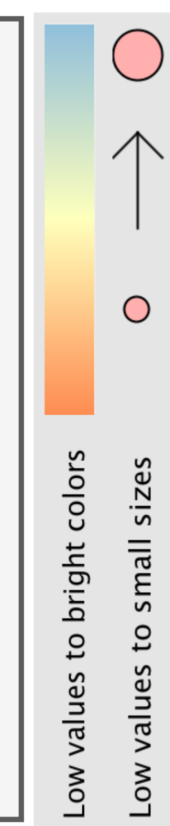


BARABASI GRAPH



RANDOM GRAPH





BETWEENNESS CENTRALITY

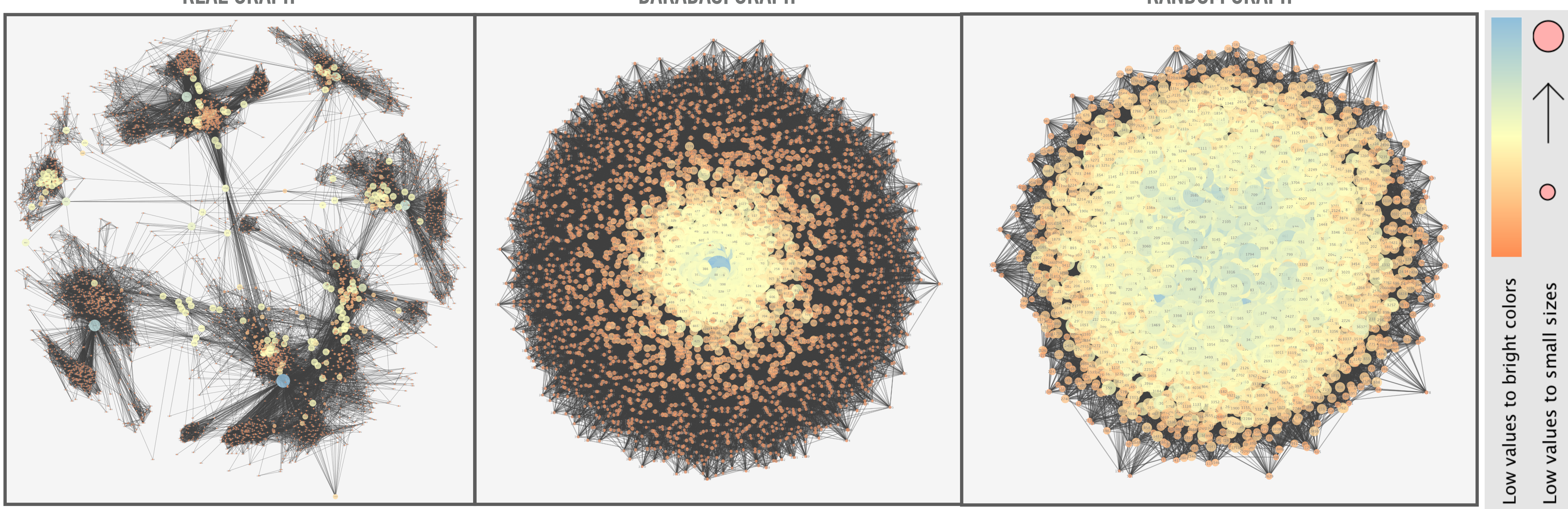
Betweenness centrality quantifies the number of times a node acts as a bridge along the shortest path between two other nodes.

In this conception, vertices that have a high probability to occur on a randomly chosen shortest path between two randomly chosen vertices have a high betweenness.

THE MAX VALUE OF BETWEENNESS CENTRALITY:

Real Graph	Barabasi Graph	Random Graph
0.48	0.021	0.001

REAL GRAPH | BARABASI GRAPH | RANDOM GRAPH



DEGREE CENTRALITY

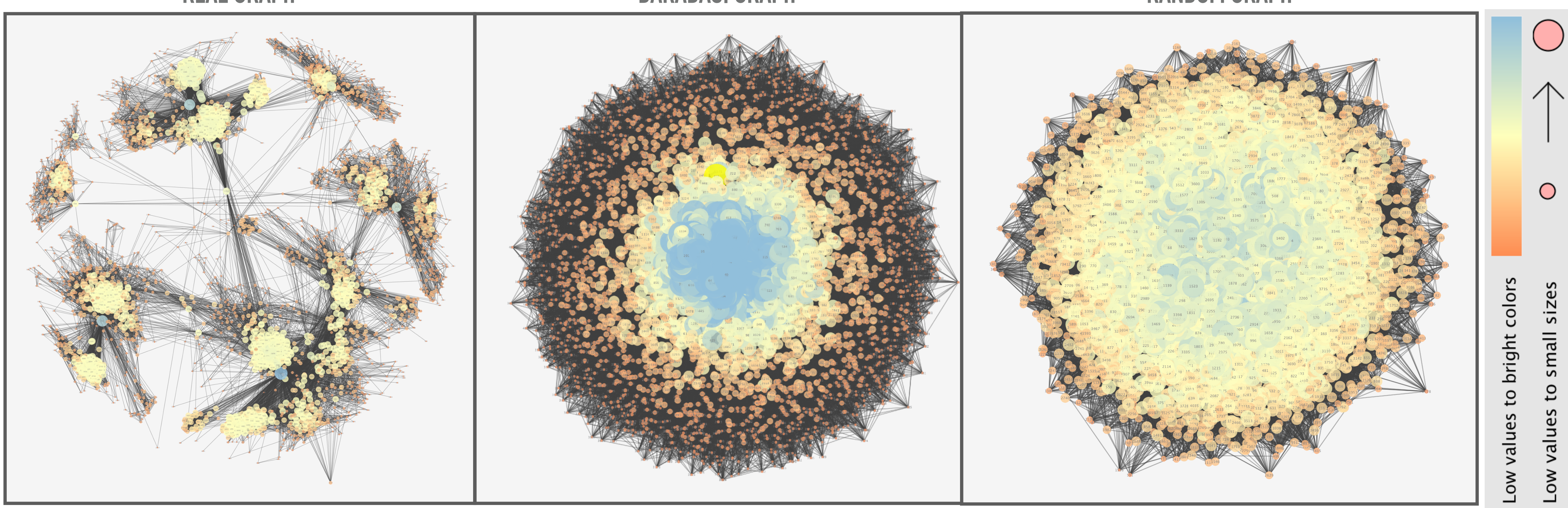
Degree centrality: a highly effective measure of the influence or importance of a node.

In many social network settings people with more connections tend to have more power and more visibility.

THE MAX VALUE OF DEGREE CENTRALITY:

Real Graph	Barabasi Graph	Random Graph
0.25	0.1	0.017

REAL GRAPH BARABASI GRAPH RANDOM GRAPH



CLUSTERING COEFFICIENT

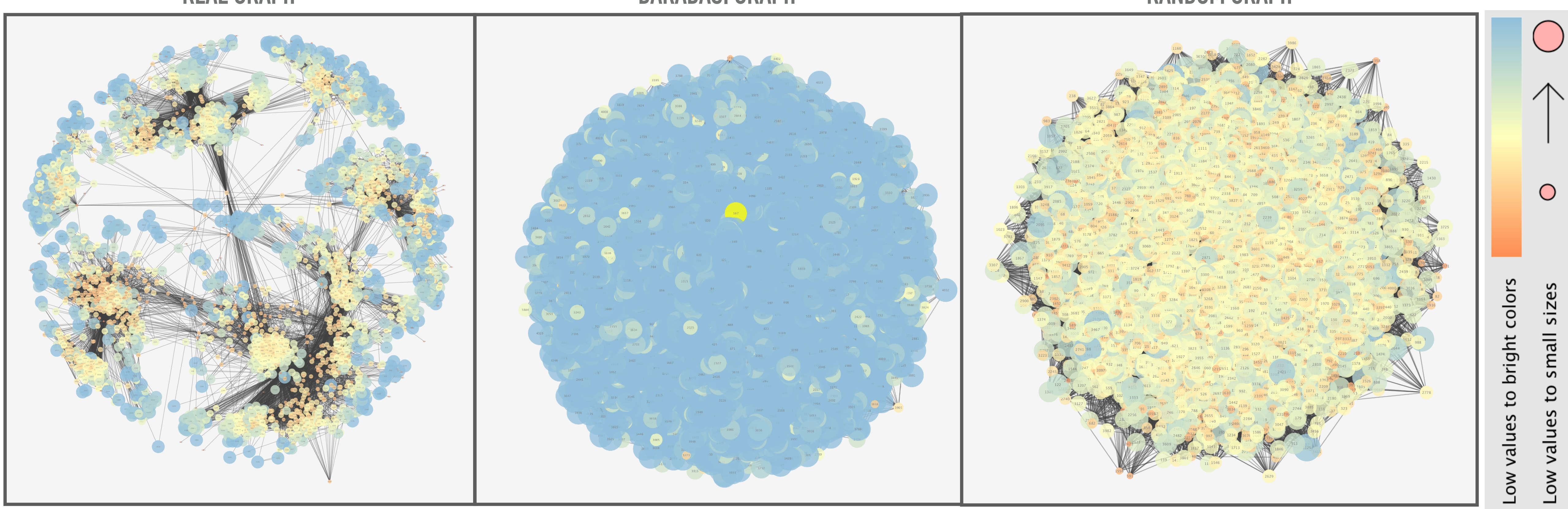
Clustering coefficient: is a measure of the degree to which nodes in a graph tend to cluster together.

Evidence suggests that in most real-world networks, and in particular social networks, nodes tend to create tightly knit groups characterised by a relatively high density of ties.

THE MAX VALUE OF CLUSTERING COEFFICIENT:

Real Graph	Barabasi Graph	Random Graph
0.606	0.038	0.011

REAL GRAPH | BARABASI GRAPH | RANDOM GRAPH



ANALYSIS OF THE SPREAD OF THE ADDICTION

- To test the model, we count the number of addicted nodes (performing 100 simulations for each kind of graph)
- As expected, the topology and not the size of the network is responsible for the largest impact on Internet addiction
- Considering the model of communication, the most influential metrics are closeness and betweenness

THE AVERAGE NUMBER OF ADDICTED NODES (CALCULATING ON 100 SIMULATIONS):

Real Graph	Barabasi Graph	Random Graph
3541.5	3467.2	2990.4

ANALYSIS OF THE SPREAD OF THE ADDICTION

- Considering the model of communication, the most influential metrics are closeness and betweenness
- The nodes characterised with the highest values of closeness and betweenness are the first ones to become addicted

THE NODES WITH THE HIGHEST VALUE OF CLOSENESS AND BETWEENNESS:

	Real Graph	Barabasi Graph	Random Graph
Closeness	User 107	User 25	User 2518
Betweenness	User 107	User 28	User 2518

CONCLUSIONS

- To propose a computational framework for the study of addiction
- To use the theory of Hybrid Automata to develop a modular model of the Dopaminergic System
- To study how addiction is correlated to network topologies
- To study different metrics on different networks

Future work:

- Further validation of our model on a network of the messaging exchange
- To investigate different stimuli

Thank you!