COMPUTATIONAL SOCIAL SCIENCES

"From Online Social Networks and Human Behavior to Game Theory"

Prof. Yamir Moreno

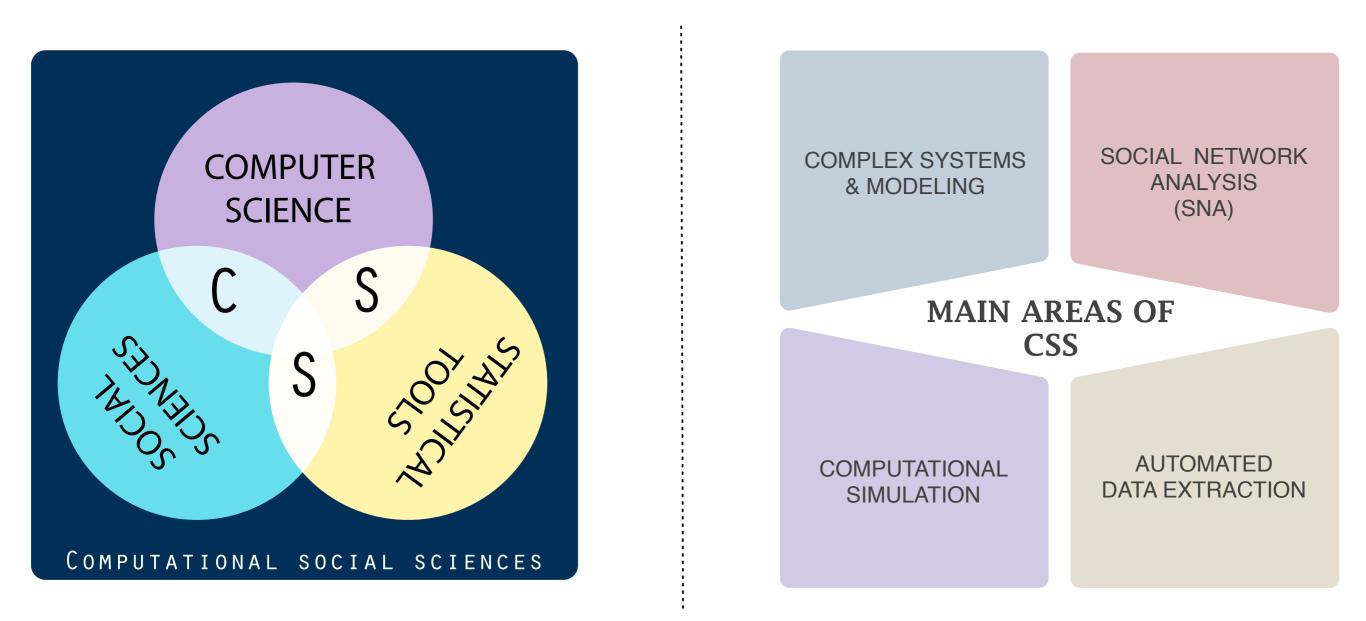
Complex Systems & Networks Lab (**COSNET**), Institute for Biocomputation and Physics of Complex Systems (BIFI) &, Department of Theoretical Physics, University of Zaragoza



Institute for Biocomputation & Physics of Complex Systems Research Institute University of Zaragoza

Universidad Zaragoza

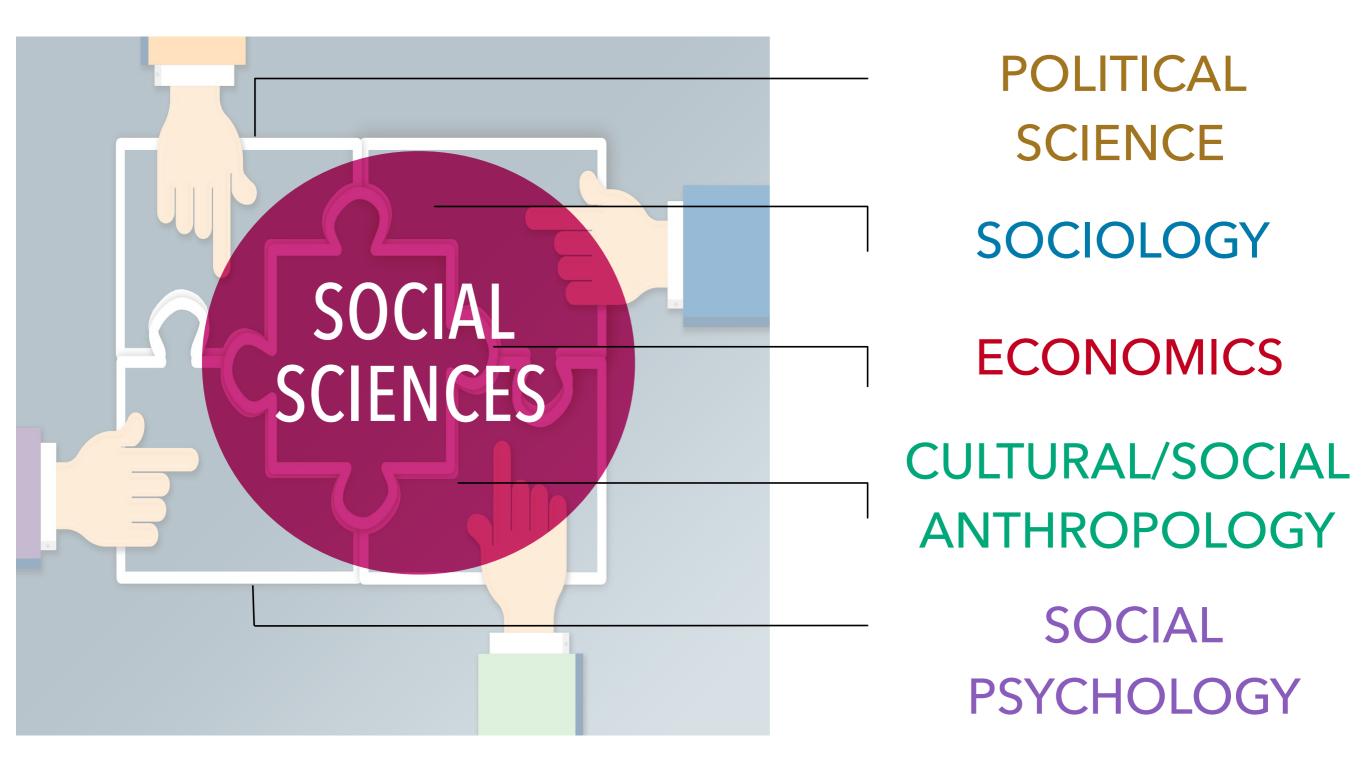
COMPUTATIONAL SOCIAL SCIENCES



The study of social phenomena by means of computing and statistical data processing. CSS has revolutionated the Scientific Method.

SOCIAL SCIENCES — Fields & Scope

Generally include:





WHAT WE HAVE...

Large Datasets (Increasing amount of data)

Big Data era - New data sources: GIS data, sensor data, economic data, etc...

New Technologies and Computing Devices (Tablets, Smartphones,..)

Social Media Platforms (Facebook, Twitter, MySpace...)

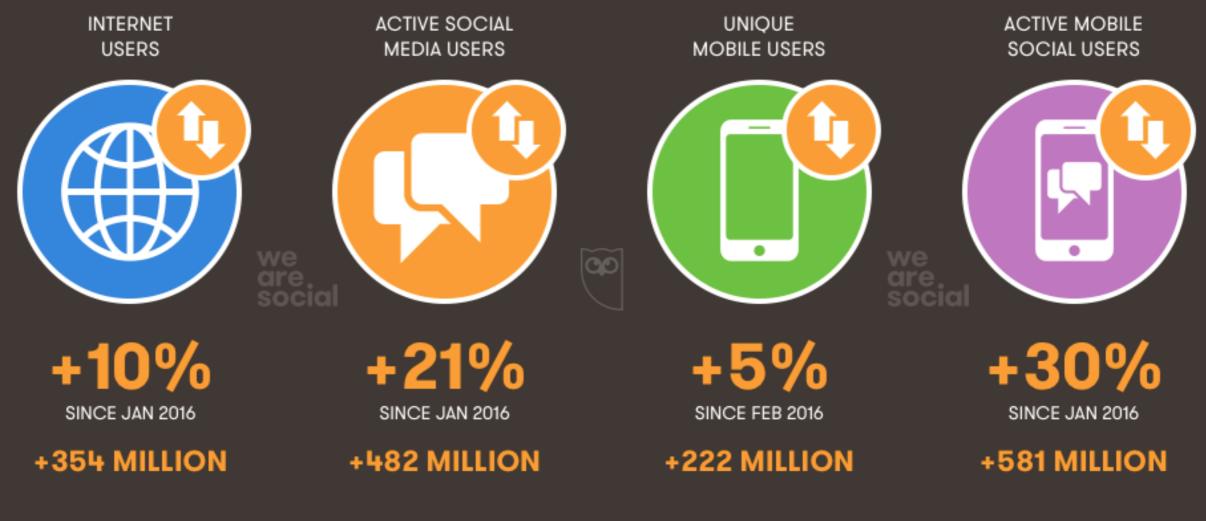
Advances in High Performance Computing (HPC)

DIGITAL REVOLUTION

JAN 2017

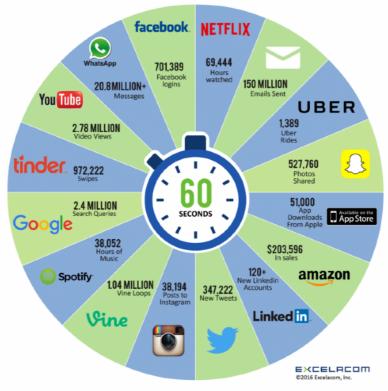
8

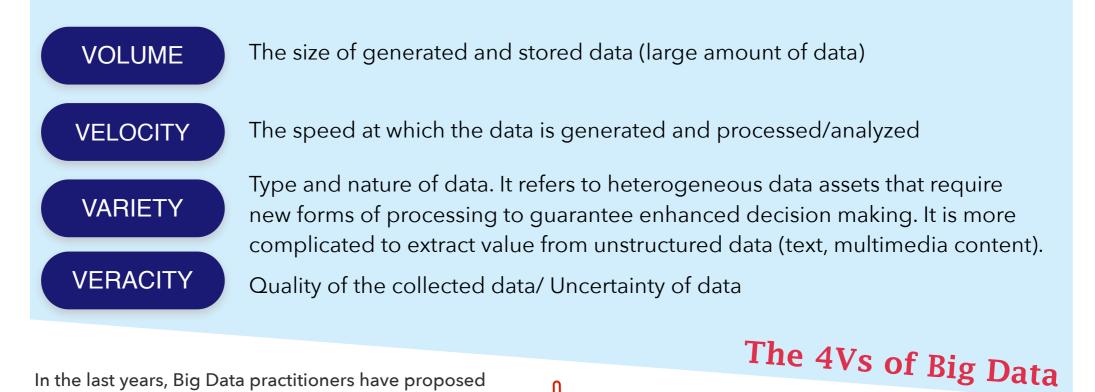
ANNUAL GROWTH YEAR-ON-YEAR CHANGE IN KEY STATISTICAL INDICATORS



SOURCES: POPULATION: UNITED NATIONS; U.S. CENSUS BUREAU; INTERNET: INTERNET: WORLDSTATS; ITU; INTERNETLIVESTATS; CIA WORLD FACTBOOK; FACEBOOK; NATIONAL REGULATORY AUTHORITIES; SOCIAL MEDIA AND MOBILE SOCIAL MEDIA: FACEBOOK; TENCENT; VKONTAKTE; LIVEINTERNET.RU; KAKAO; NAVER; NIKI AGHAEI; CAFEBAZAAR.IR; SIMILARWEB; DING; EXTRAPOLATION OF TNS DATA; MOBILE: GSMA INTELLIGENCE; EXTRAPOLATION OF EMARKETER AND ERICSSON DATA. COMPARISONS TO WE ARE SOCIAL'S "DIGITAL IN 2016" REPORT. Hootsuite dre social

2016 INTERNET MINUTE?





In the last years, Big Data practitioners have proposed additional Vs: VALUE, VARIABILITY, VISUALIZATION

We generate a huge amount of data, but what is really important is that the toughest challenge begins now

DATA ARE REALLY BIG, May Jun BUT CHALLENGES TOO

2017/18

Sep

Oct

2017/18 2016/17

2016/17

Apr

sectem

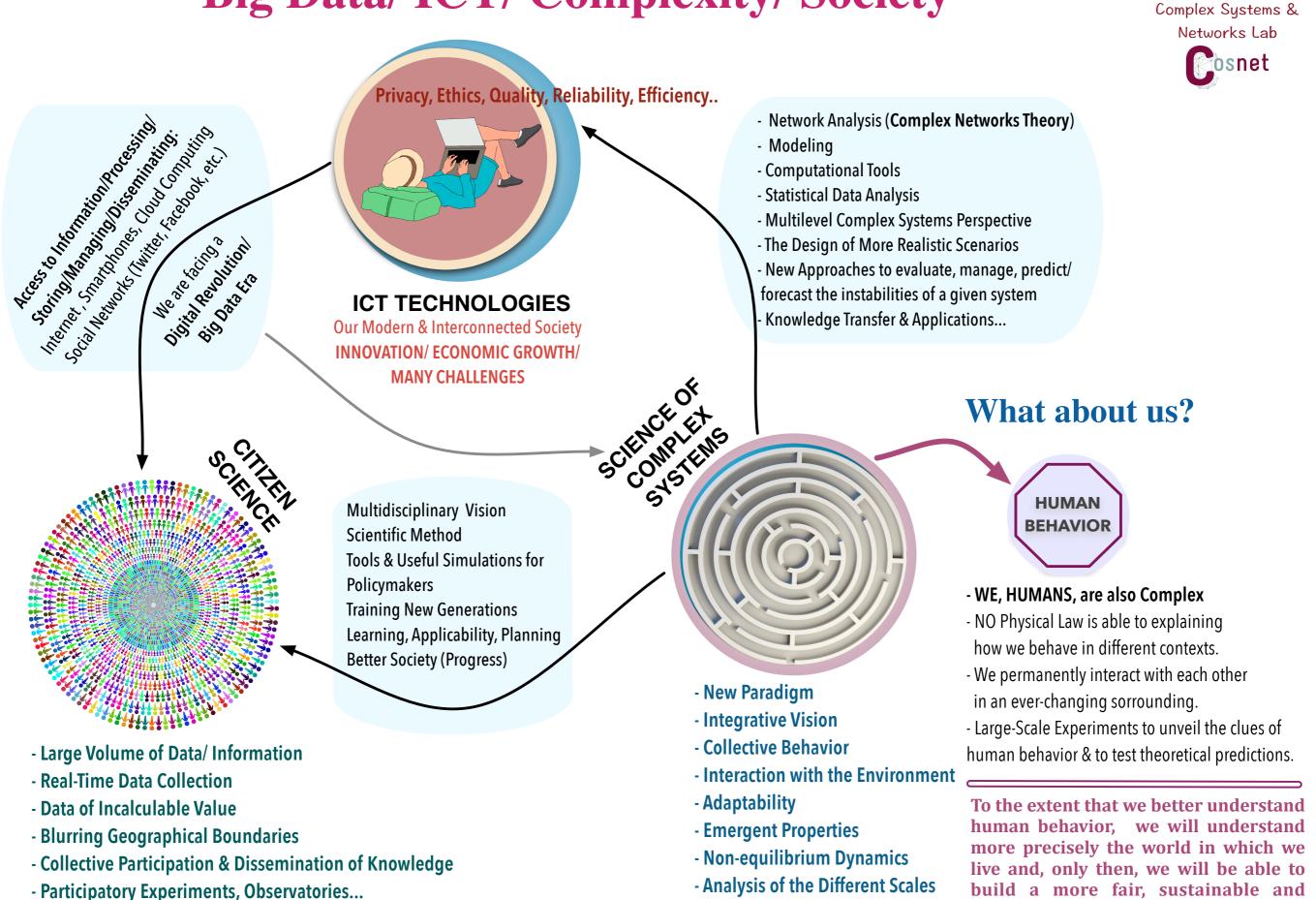
Statistic 2

Large sample size High Dimensionality

We need to overcome the current limitations

COLLECTED ANALYSED AGGREGATED TRANSMITTED SHARED VISUALIZED...

Big Data/ ICT/ Complexity/ Society



of the System...

Infographic: Cosnet Lab

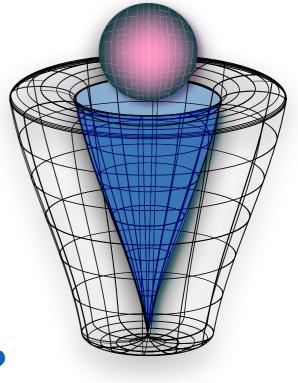
advanced Society.

APPROACH:

Integrative/ Multidisciplinary Approach (COMPLEXITY SCIENCE)

COMPLEXITY SCIENCE: What does it allow us?

- To understand the general principles that govern the behavior of a wide set of real systems in order to be able to predict, model and control their functioning.
- To develop new models and algorithms to analyze the emerging properties of social and technological systems and evaluate the potential failures of them.
- To progress enough in the basic theoretical aspects and in the application of the generated knowledge to achieve a more precise characterization of real systems.
- To develop powerful tools to deal with new forms of interaction between individuals.
 Methodological transformation of current modeling paradigms.
- · To foster interdisciplinarity.



NEW PARADIGM

COMPLEXITY SCIENCE

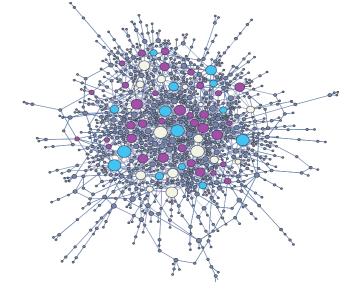
Understanding, Evaluating, Managing, Predicting/Forecasting the Behaviour of Complex Systems

Interconnected & Interdependent elements Emergent Properties & Behavior Non-Linearity Self-Organization Networked Hierarchical Connectivities Multiple Levels of Organization Sensitivity to Initial Conditions Systems change over time Limited predictability Adaptability...

Systems are treated as a whole with a focus beyond traditional boundaries THE WHOLE IS MORE THAN THE SIMPLE SUM OF ITS PARTS VS REDUCTIONIST APPROACH (Detrimental Influence in many areas, reduce phenomena to simple terms, NO emergent properties...) RELATIONSHIPS BETWEEN ACTORS, KEY ACTORS, ACTOR'S LOCATION IN THE NETWORK, NETWORK STRUCTURE, CENTRALITY MEASURES, etc.

SOCIAL NETWORK ANALYSIS - NETWORK THEORY:

- The simultaneous characterization of the interactions and dynamics at a local scale and the study of their integration into a global and coherent dynamics at a system-wide scale.
- New methods, tools and techniques for mapping and measuring the relationships among people and organizations.
- Social network analysis (SNA) offers a useful mathematical and visual analysis of human relationship.
- An importan role in influencing learning and in the diffusion of behaviors, technology, etc.
- Extract and intelligently exploit knowledge from data collected, promote innovation processes and diversify opportunities in all fields.
- Different applications: viral marketing, organizational dynamics, law enforcement, etc
- To foster new opportunities for collaborative research.

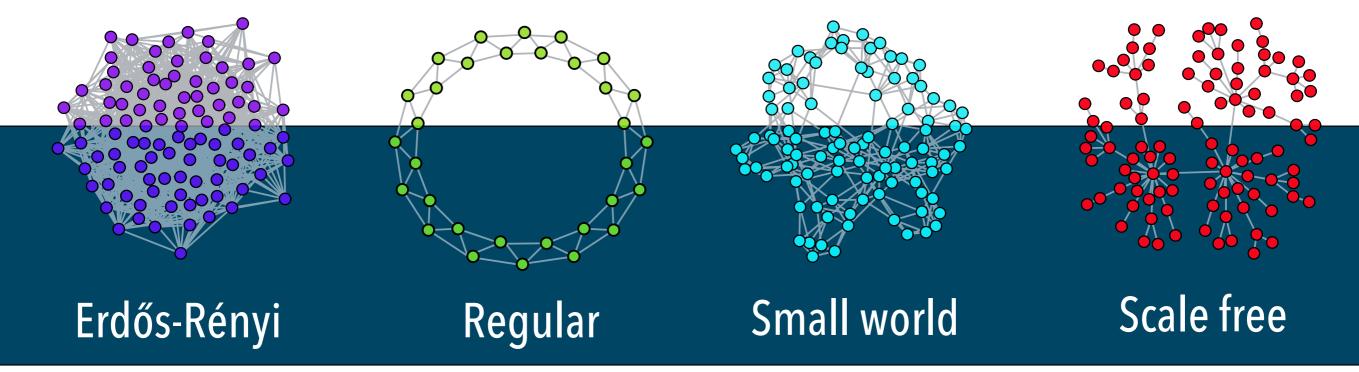


Social Network Analysis (SNA): The study of social structures through the use of networks and graph theory. It is focused on the importance of relationships among interacting units/ actors (people, groups, organizations), and provides a powerful analysis of the patterns of human interactions.

Social Networks: Nodes of individuals, organizations, groups which are connected by different kinds of interdependencies. Social Networks can be studied by means of **Graph Theory.**

Graph: A mathematical abstraction consisting of a set of *N* nodes or vertices, connected by a set of *E* edges or links.

Network analysis gives a solid and detailed vision of different aspects of social phenomena.



Different techniques, tools and methods for **collecting data**, **visualization** and **statistical analysis are used in the field**.

APPLICATIONS: Disease spreading (how does an infectious disease spread in a population?), interorganizational collaboration, decision-making processes...

WHAT WE HAVE...

Large Datasets (Increasing amount of data) Big Data era - New data sources: GIS data, sensor data, economic data, etc... New Technologies and Computing Devices (Tablets, Smartphones,..) Social Media Platforms (Facebook, Twitter, MySpace...) Advances in High Performance Computing (HPC)

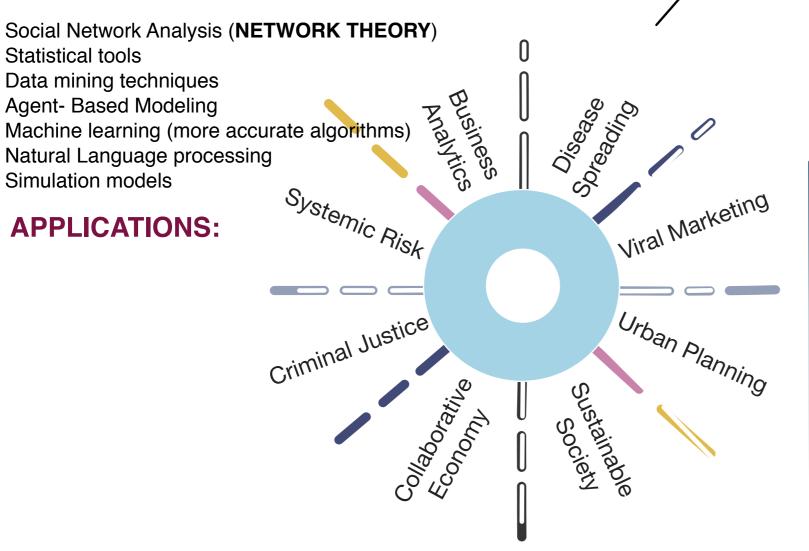
APPROACH:

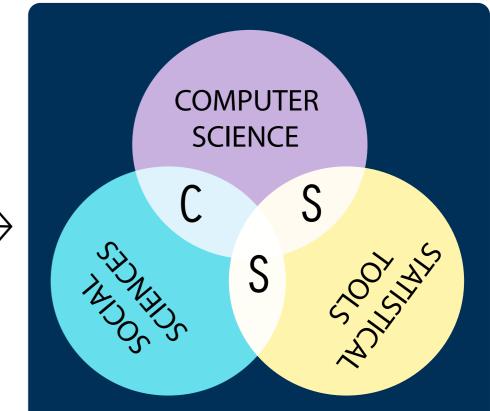
Integrative/ Multidisciplinary Approach (COMPLEXITY SCIENCE)

SOCIAL SCIENCES:

The study of social phenomena by means of Computing and Statistical Data Processing

THEORIES, METHODS, COMPUTATIONAL TOOLS:





COMPUTATIONAL SOCIAL SCIENCES

PRESSING CHALLENGES:

The Digital Revolution has generated a significant amount of data, but... Do we know how to make optimal, appropriate and ethical use of that information?

We need to develop new metrics and algorithms and the Mathematics of "Big Data "

Understand how the *Society of the Future* will be and **what we can do to be prepared for the changes** (Security, New Protocols, Governance, Data protection,...)

CHALLENGES:

BIG DATA & SOCIAL SYSTEMS

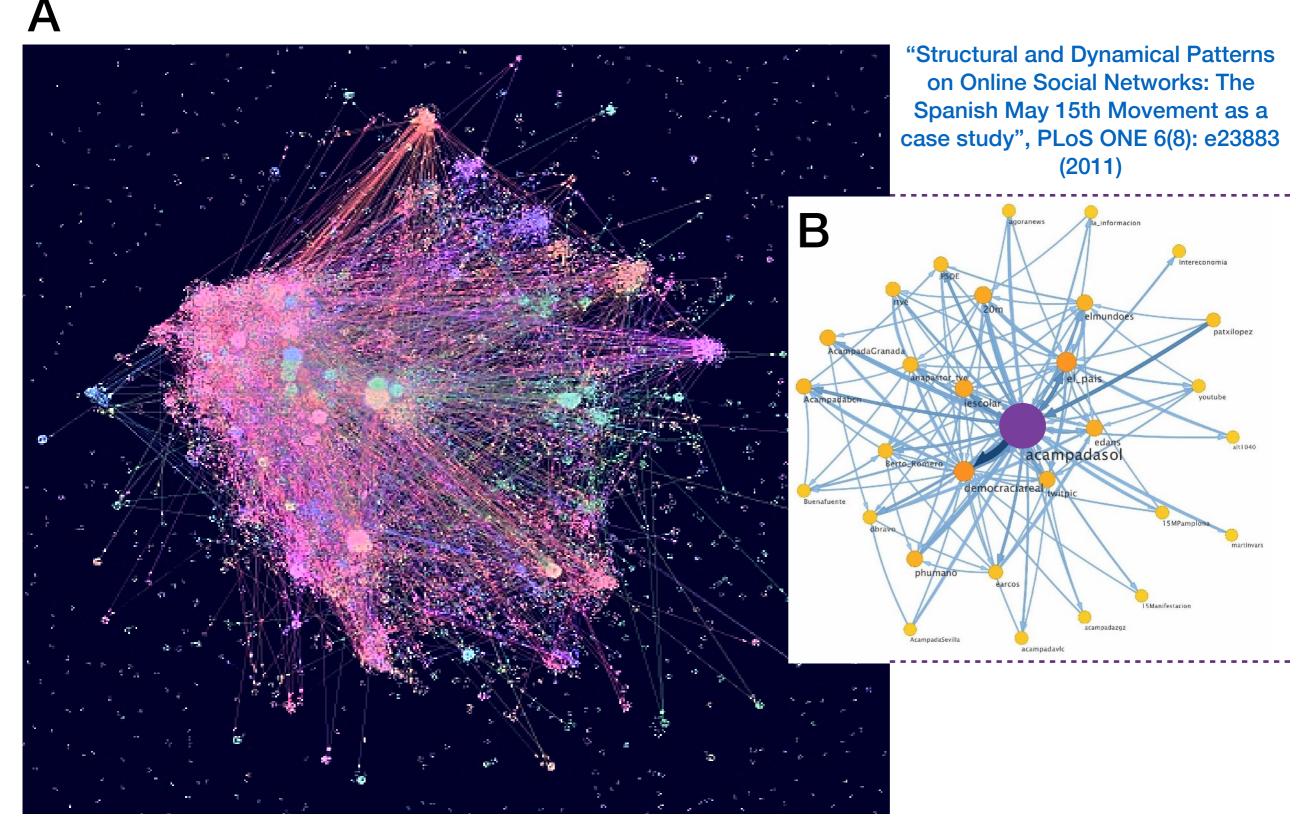
•

Data alone do NOT represent knowledge

- Lay down the foundation for the quantitative modeling of large-scale social phenomena in complex and realistic real-world contexts.
- To face the new challenges to data-driven and data-intensive applications.
- To pave the way for the transition from an analogue to a digital society.
- Patterns of behavior found through extensive data mining will feed improved versions of current models which can then be used to implement policies aimed at improving the citizens quality of life.
- Identify the best course of action to transfer the acquired knowledge from basic sciences to the application level.
- \cdot To develop new algorithms and standards for data processing.
- To be able to anticipate the consequences of new regulations, actions or systems failures.
- Guarantee the quality and consistency of data.
- How to store, protect and catalog the data?
- To work on ethical and legal regulations.

ONLINE Social Systems

STRUCTURE & DYNAMICS OF INFORMATION IN ONLINE SOCIAL NETWORKS

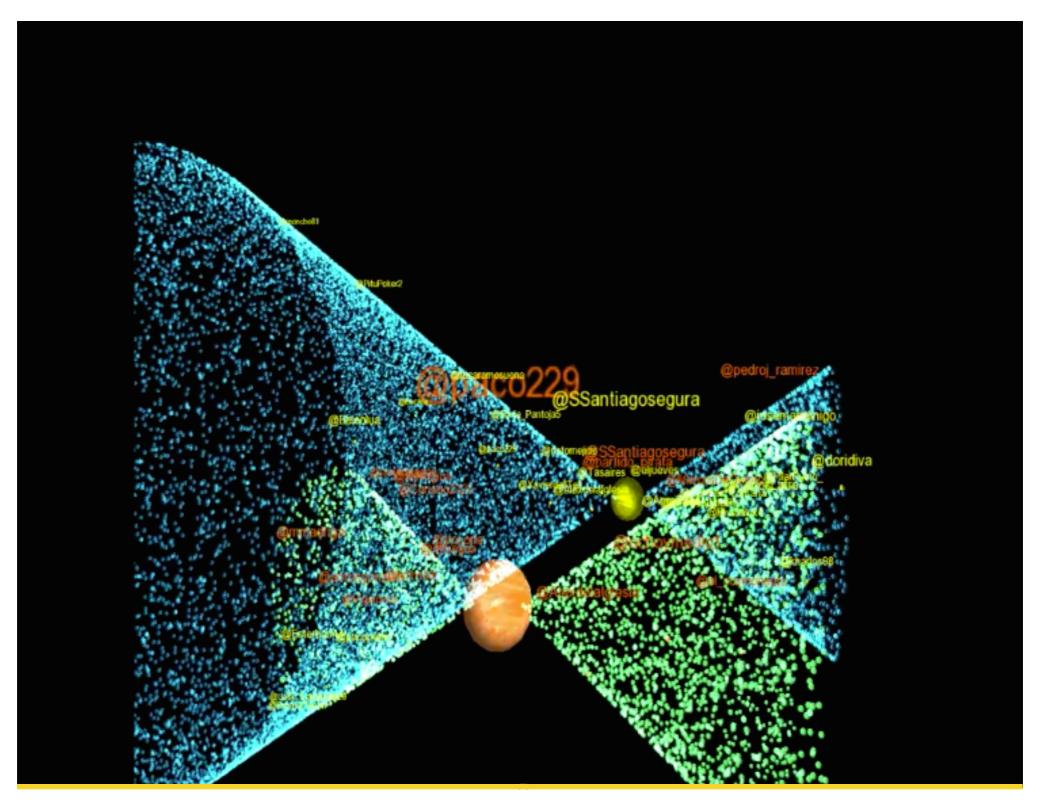


Indignados Movement, Spain.

A: Interaction Network between Twitter users/B: Most representative communities

In the online world, the "average individual approach" does not work in general.

Few individuals receive a lot of messages while others are mainly senders.



Video: BIFI Institute ©

This has allowed:



- Blurring the geographical boundaries.
- Quick access to information.

- Strengthen the relationship between individuals (online world).

- To group together many social agents around a common issue/goal.

- Deeply modify the dynamics of Information.
- Making our World more global.

ACCESS TO FREE LARGE DATA SETS storing, processing data & also make sense of all the information available

> **RENOVATED** INTEREST IN THE **STUDY** OF LARGE SOCIAL SYSTEMS

STRUCTURE

- Formation and evolution of interaction networks.
- Topological properties of individuals.

-Emergent properties of the system.

- DYNAMICS
- Propagation of information. Adoption of certain behaviors.

MODELING CONTAGION IN ONLINE SOCIAL SYSTEMS

An idea, behavior, product is adopted and transmitted in a population by individuals known as **Potencial Adopters**. They are previously exposed to the behavior of other individuals in their contact networks. **It is a local process**.

DIFFERENT CONTAGION PATHS (Epidemic & Rumor Models, Threshold Models...)

Epidemic- and Rumor- like Dynamics MODELS

The decisions to adopt are taken independently with probability *p* for each successive contact.

Contagion dynamics: At each time step, infected individuals propagate the contagion to susceptible neighbours with probability λ . Infected individuals can recover at a rate μ (SIR models), or they can revert to the susceptible state with probability μ (SIS models).

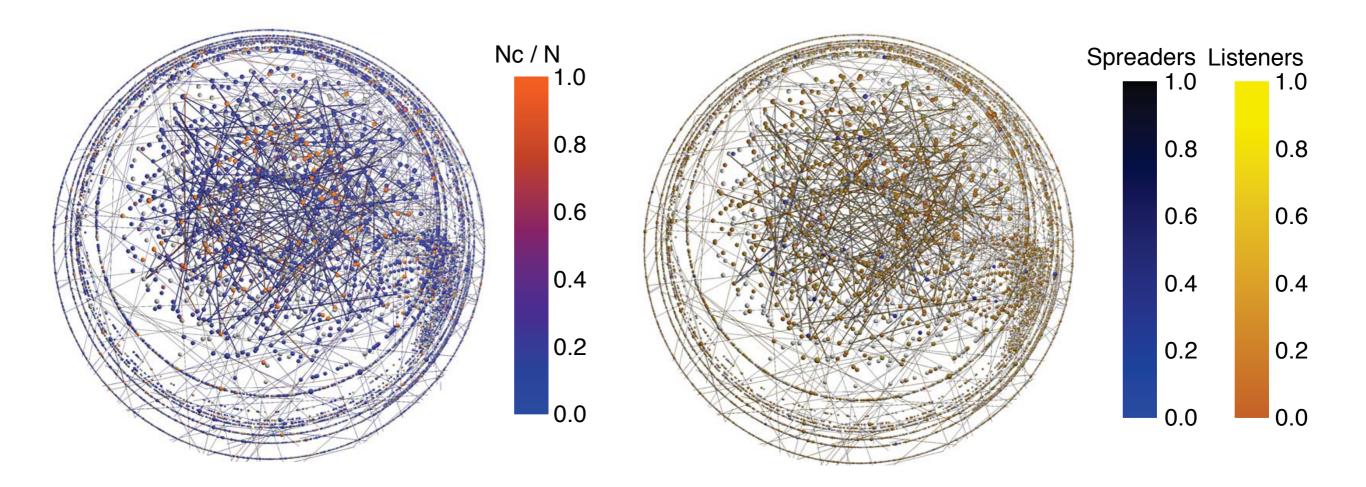
THRESHOLD MODELS

In the Threshold models the decision of adopting certain behavior/idea depends on a critical proportion of contacts that have already adopted such behavior, so that a particular individual/agent will only adopt it, if **his number of active contacts is higher than a certain threshold**.



snet

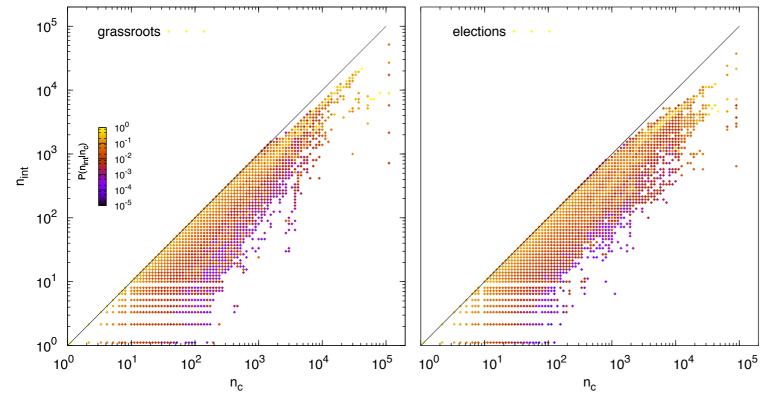
"The Dynamics of Protest Recruitment through an Online Network", Scientific Reports 1, 197 (2011). S. González-Bailon, J. Borge-Holthoefer, A. Rivero, and Y. Moreno



We study recruitment patterns in Twitter network and find evidence of Social Influence and Complex Contagion.

We also identify the network position of early participants (i.e. the leaders of the recruitment process) and of the users who acted as seeds of message cascades (i.e. the spreaders of information).

"The Role of Hidden Influentials in the Diffusion of Online Information Cascades", EPJ Data Science 2:6 (2013). R. A Baños, J. Borge-Holthoefer and Y. Moreno



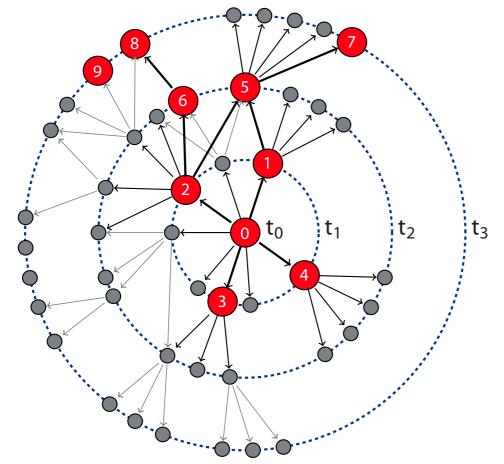
Inter- and intra-modular cascading events for the topics under consideration (left: 'grassroots'; right: 'elections'). Binned representation of the relationship between the number of nodes in a cascade that unfolds in the same community of the initial seed and the size of the cascade itself. Proportions have been normalized column-wise, i.e. by the total number of cascades with the same size.

Time-constrained information cascade

- Time-constrained cascades: Nodes are disposed in concentric circles indicating the time when they received a specific tweet.

- Links between them represent the follower/friend relationship: an arrow from *i* to *j* indicates that *j* follows *i*, as any tweet posted by *i* is automatically received by *j*.

- **Red nodes** are those who posted a new message at the corresponding time, whereas **gray nodes** only listened to their friends.



HIDDEN INFLUENTIALS

The success of an activity cascade **might greatly depend on intermediate spreaders characteristics, and not only on the properties of the seed nodes**. That being so, a large seed k_{out} (*i.e.* its follower set) may be a sufficient but not a necessary condition for the generation of large-scale cascades.

- Here, we have to consider the important role of a new class of actors: The Hidden Influentials:

Hidden influentials, *i.e.*, relatively smaller (in terms of connectivity and centrality) nodes which, on the aggregate, can make chain reactions turn into global cascades.

These nodes **do not occupy key topological positions** that would *a priori* identify them as influential, and yet they play a major role promoting system-wide events. Therefore, getting these nodes involved **has a multiplicative impact on the size of the cascades.**

- Hubs often act as cascade firewalls rather than spawners

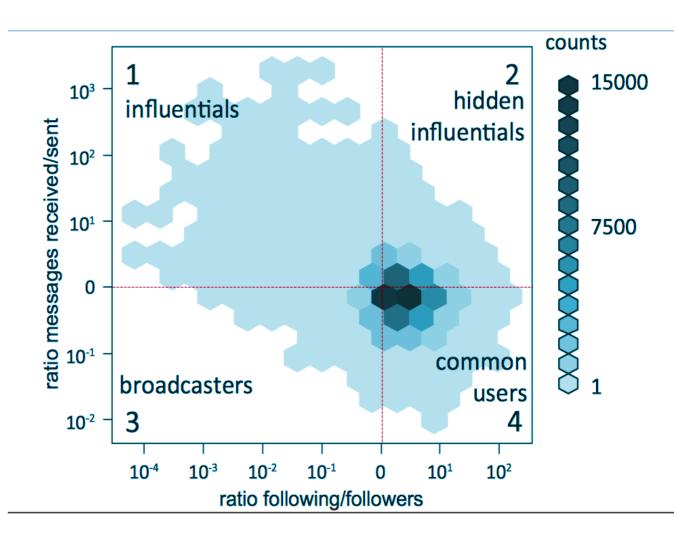
For a cascade to be successful in terms of the number of users involved in it, key nodes should be engaged. These nodes are not the hubs, which more than often behave as firewalls, but belong to a middle class that either has a high multiplicative capacity or bridges the modules that make up the system.

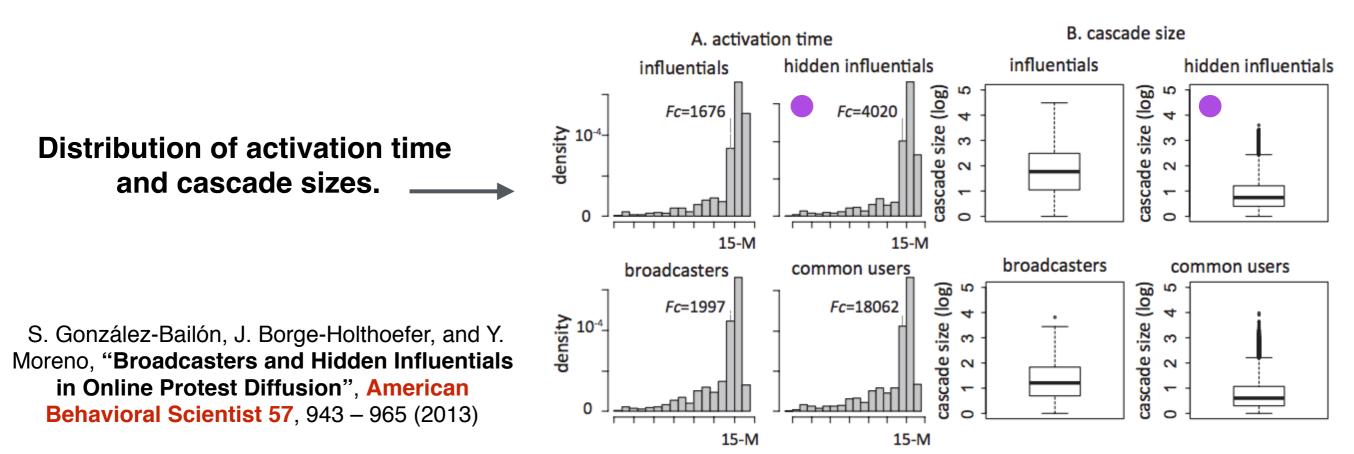
Distribution of users according to Network Position and Message Activity

The vertical axis tracks the number of protest messages that users received by the number of protest messages they sent; the most visible users (those who were mentioned more often) are above the dashed line.

The horizontal axis tracks the number of other accounts a user is following by the number of followers the user has; the most central and popular users in this baseline network are on the left of the dashed line.

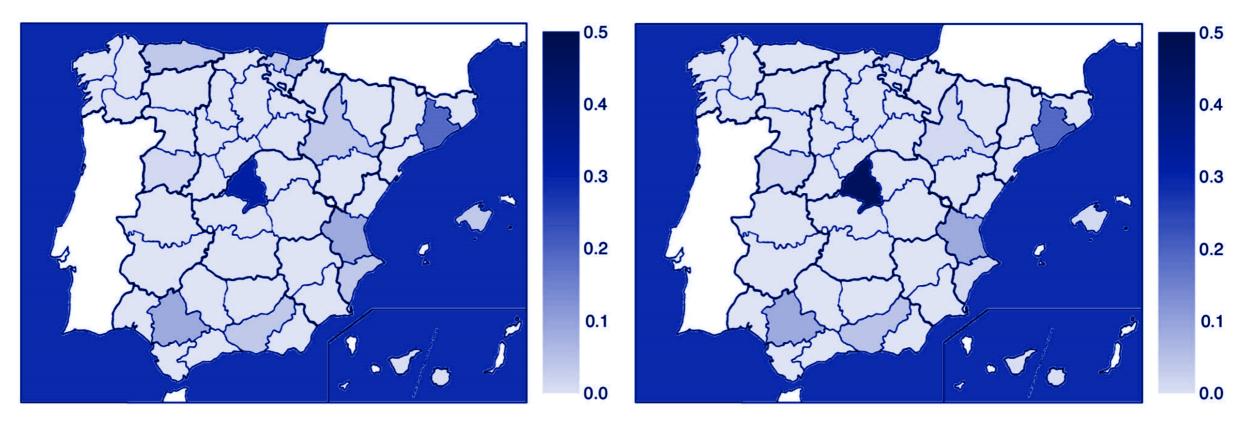
Users who are central in the overall communication network are not necessarily the most visible users in the stream of protest information flow.





Number of Messages Sent (fraction)

Number of Messages Received (fraction)



Geographical distribution of the protests The maps are based on profile location information

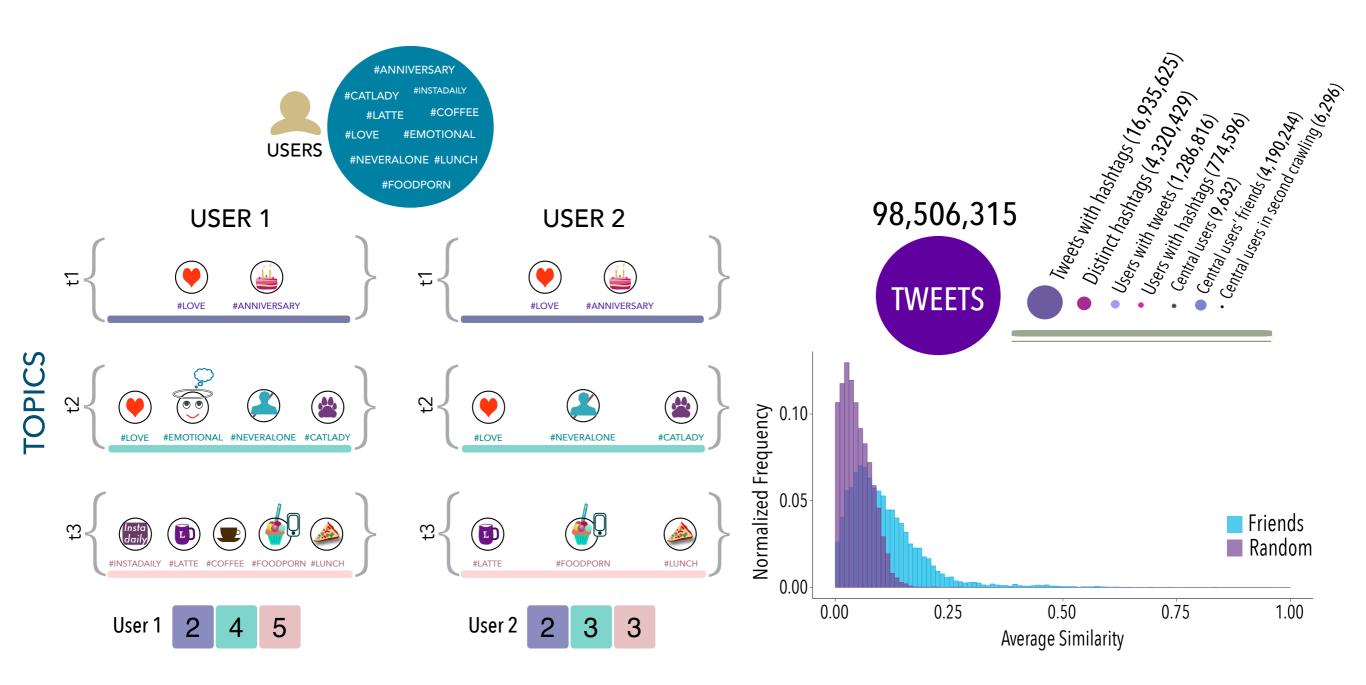
Variable	Following-Follower Network	Mentions Network
N (number of nodes)	87,569	87,569
M (number of arcs)	6,030,459	206,592
<k> (average degree)</k>	69	2.36
$\max(k_{in})$ (maximum indegree)	5,773	29,155
$max(k_{out})$ (maximum outdegree)	31,798	289
C (clustering)	0.22	0.034
I (average path length)	3.24	1.7

Table 1. Network Statistics for the Following-Follower and Mentions Network.

Homophily

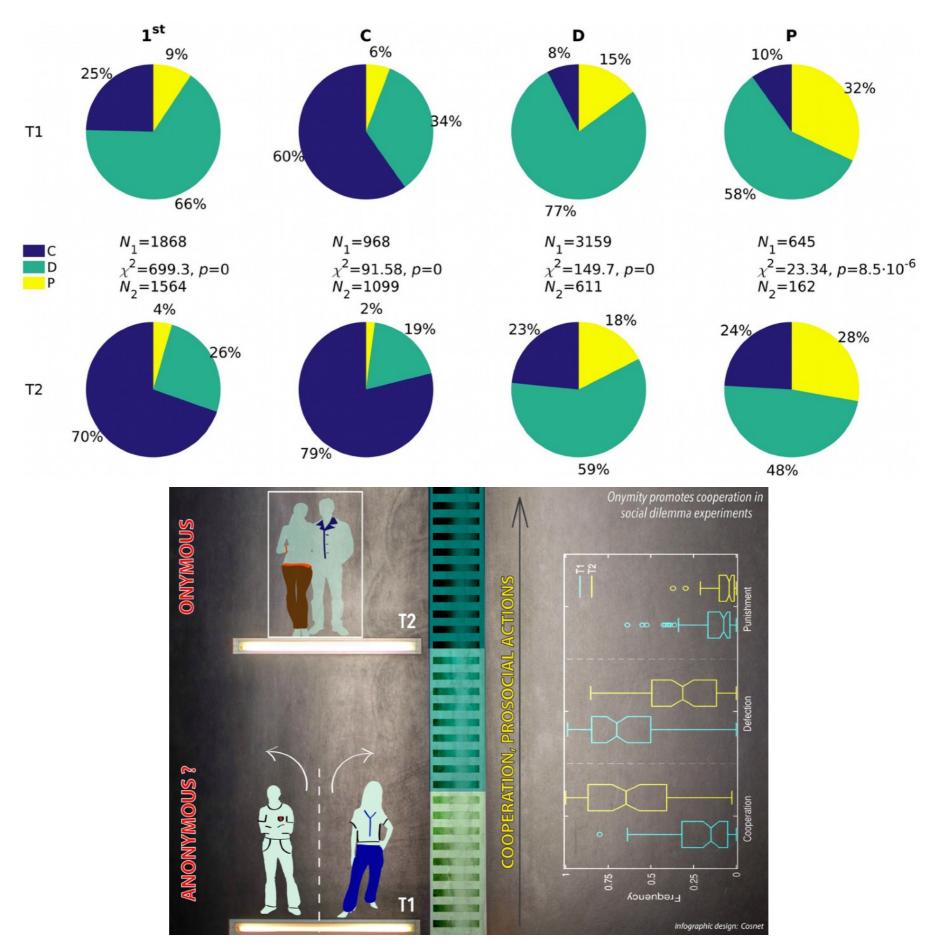
"Topical Homophily in Online Social Systems" Felipe Maciel Cardoso, Sandro Meloni, Andre Santanche & Yamir Moreno

https://arxiv.org/abs/1707.06525

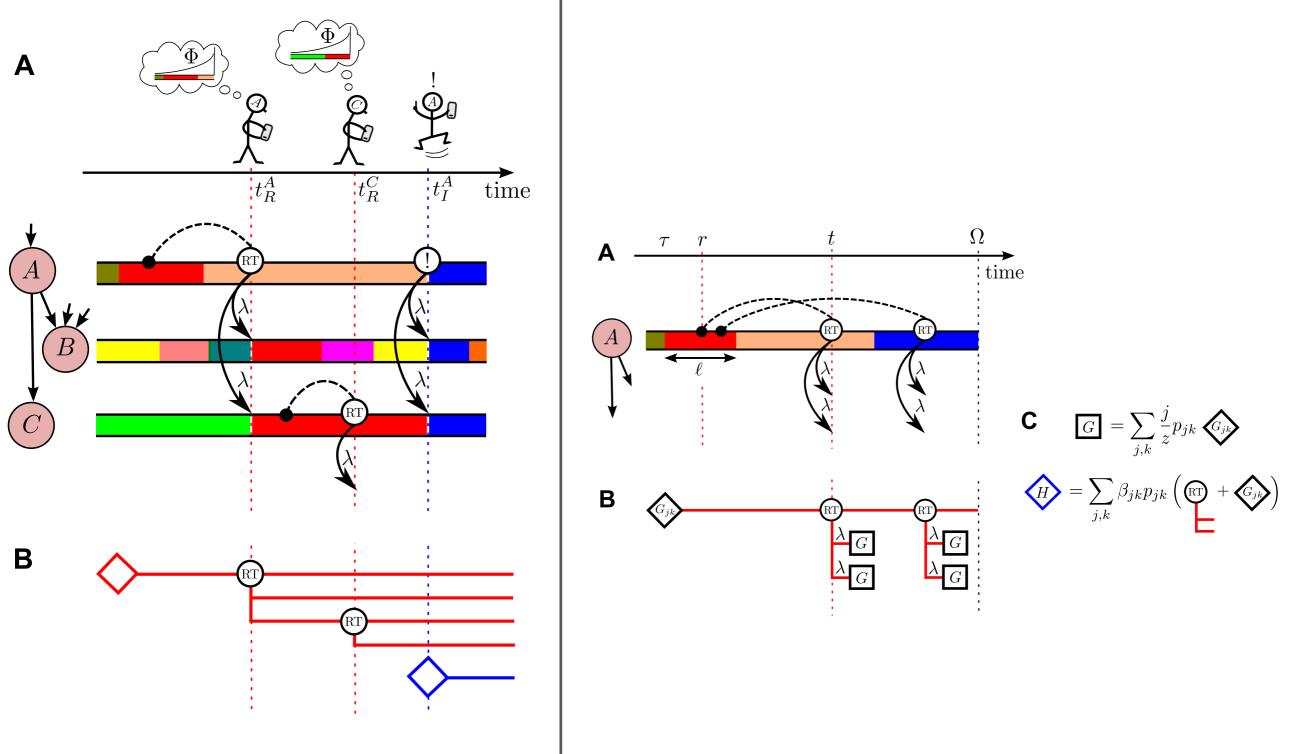


"Onymity promotes cooperation in social dilemma experiments" Jusup, R.-W Kurths. Ē Wang, . כ and Y. Moreno 3:e1601444 (2017), Z. Y. Iwasa, Shi, Advances Wang, Science

Onymity vs **Anonymity**







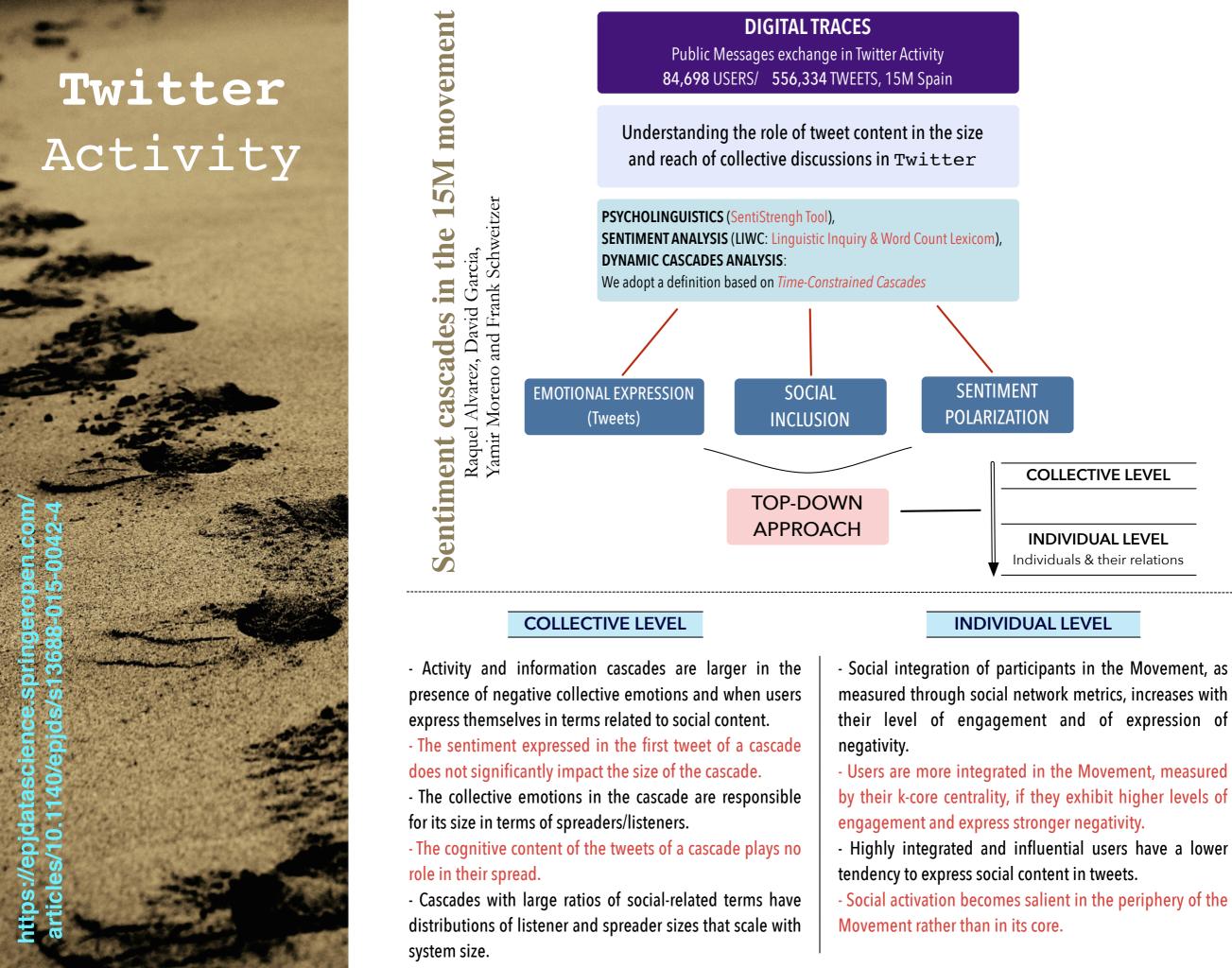
"Effects of Network Structure, Competition and Memory Time on Social Spreading Phenomena", Physical Review X 6, 021019 (2016), J. P. Gleeson, K. P. O'Sullivan, R. A. Baños, Y. Moreno

Sentiment Analysis



Sentiment Analysis is the field of study that analyzes people's opinions, sentiments, attitudes, and emotions towards entities such as products, services, organizations, individuals, etc. It consists of an application of text analytics techniques for the identification of subjective opinions in text data and determines if they are positive, negative or neutral.

One of the most sought-after topics in Computer Science





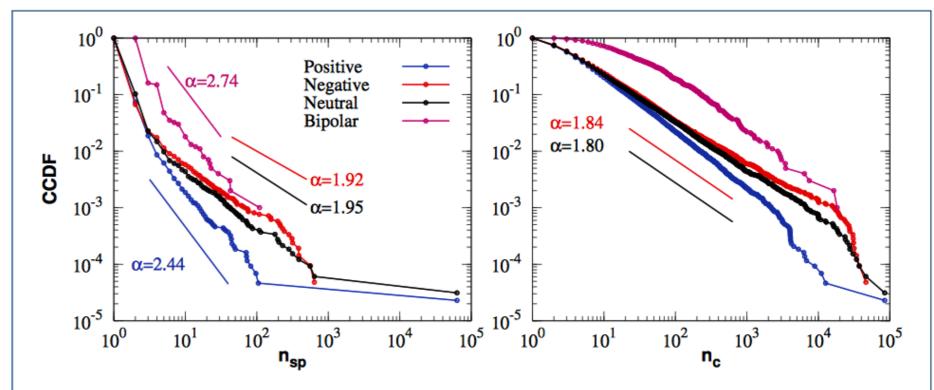
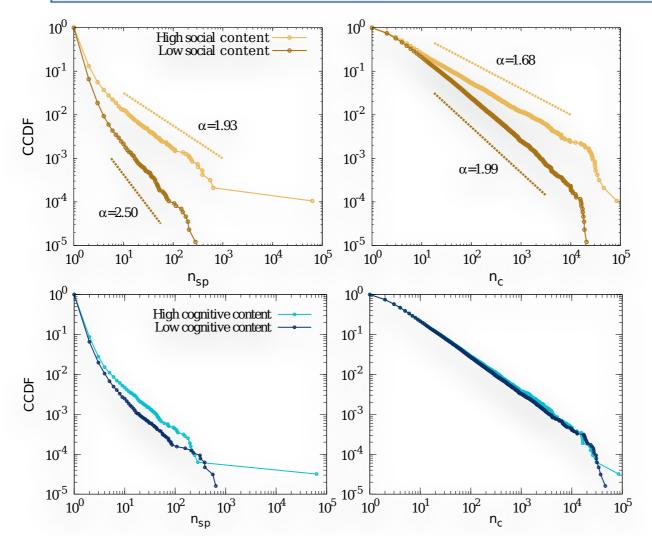


Figure 1 Complementary cumulative density function for activity cascade sizes (left) and information cascade sizes (right). In this case, cascades have been classified according to their aggregate sentiment into positive, negative, neutral and bipolar.



CCDF of activity (left) and information (right) cascade sizes for cascades of high and low social content (top) and high and low cognitive content (bottom). Dashed lines show the result of power-law fits.

Human Behavior

We are complex, heterogeneous, sometimes unpredictable, but with an extraordinary capacity to help strangers in the most unimaginable and unprofitable circumstances. Possibly, that is where our greatness lies.

The are many factors that influence human behavior:



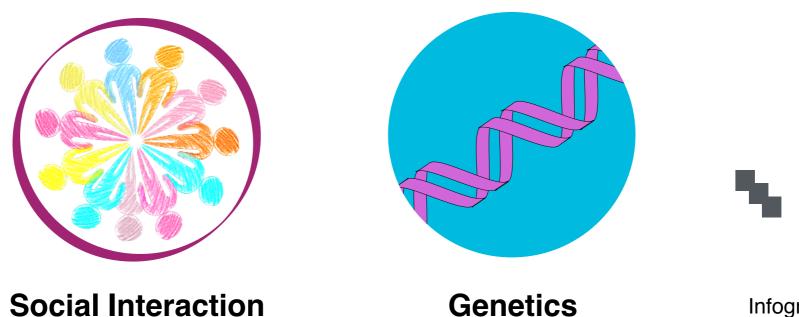




Personal Attitude

Influence of the Environment

Cultural Aspects



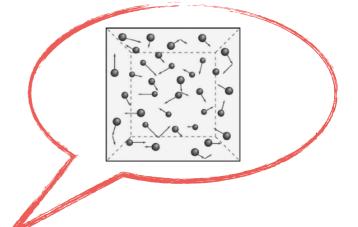


The Physics of Human Behavior: a Conundrum

- We are heterogeneous (still, we are able to reach agreements, consensus)
- We don't know the laws that govern human behavior

•

٠



 Certainly, we do not behave like molecules of an Ideal Gas, and many methods of Physics fail when it comes to analyze collective human behavior.

YAUGU WAA MANINAWIN

- Mean Field Theory———FAILS———The "average individual" does not exist.
- We cooperate more than any other species, but we cooperate a lot with unrelated individulas.
 Coooperative behaviors can not be explained by kin selection. This particular problem related to altruism and cooperation was one of the major difficulties to which Darwin's Evolutionary Theory had to face.
- Game theory (mathematical modeling of strategic interaction among rational and irrational agents) is extremely useful to unravel human behavior and also provides analytical tools for **understanding a wide range of phenomena that occur in real life**. The Prisoner's Dilemma is the best-known and studied model in game theory.

GAME THEORY

In a world in which the interaction networks and relationships between individuals are becoming more and more complex, different hypotheses have emerged to explain the foundations and mechanisms of human cooperation. **Real-world situations can be modeled and analyzed as a game by means of Game Theory**, specially, in fields such as: Biology, Sociology, Economics, etc.

The Prisoner's Dilemma is the best-known and studied model in game theory. It shows how cooperation requires a commitment from both players in order to obtain mutual beneficial outcomes.



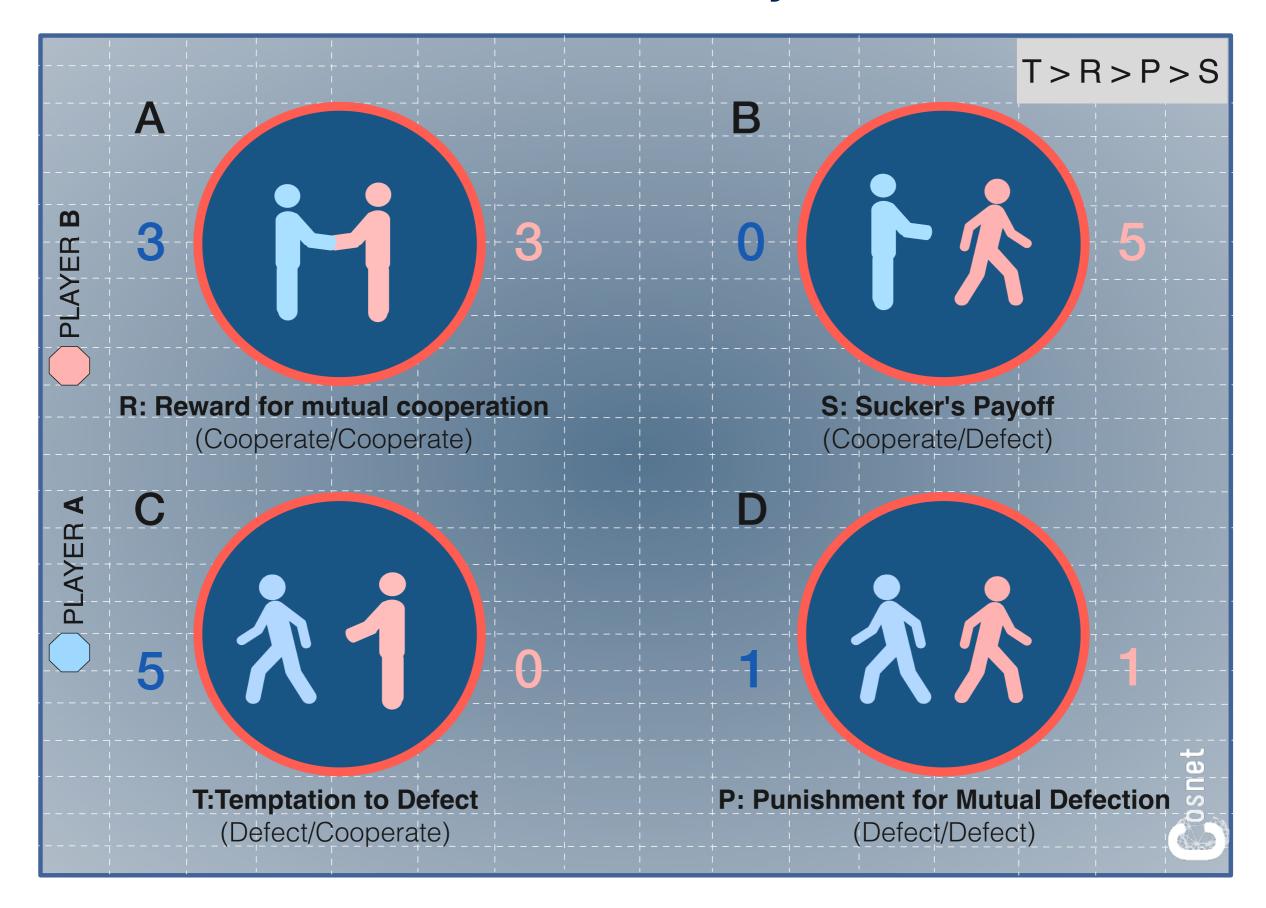
Prisoner Dilemma is a formalized incentives game, which considers two types of players:

- Cooperators: who pay a cost to help other people
- Defectors: who avoid paying this cost while reaping benefits from cooperators

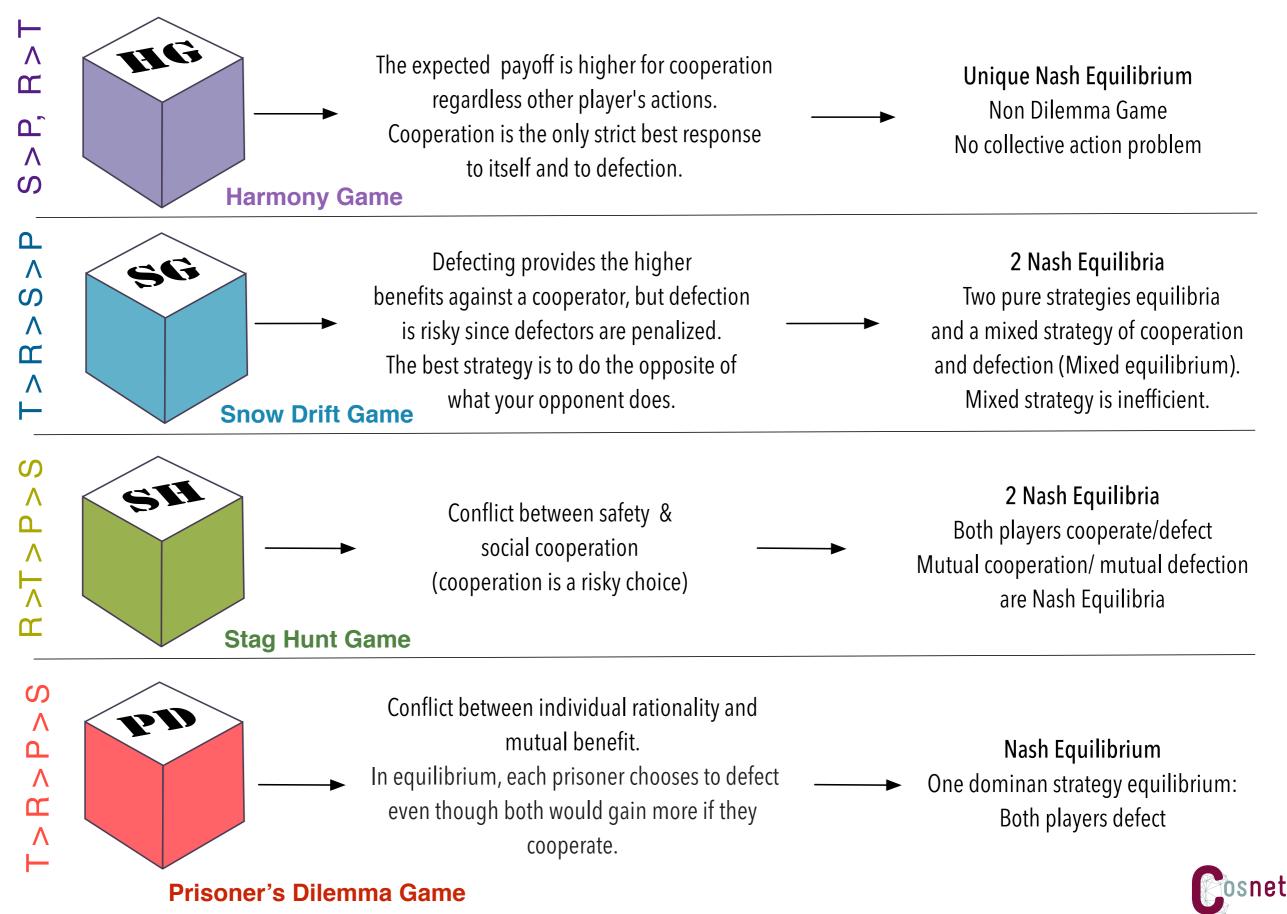
From an individual/unilateral perspective, not cooperating is always the best strategy: If the other player cooperated, I would get more benefits if I would not. If the other player did not cooperate, the best thing for me would also be not to cooperate.

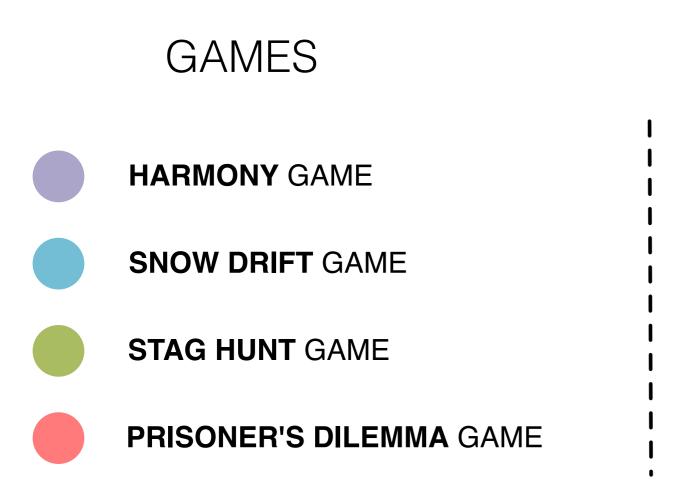
However, mutual cooperation is better than not cooperating unilaterally. This is the dilemma!

Prisoner's Dilemma: Payoff Matrix



GAME THEORY- Different Games





Payoff Matrixes

- **T:** Temptation to defect
- R: Reward for mutual cooperation

P: Punishment for mutual defection

S: Sucker's payoff

Nash equilibrium is a term used to describe a scenario where no player can profitably deviate given the strategies of the other players.

A **pure strategy** is one strategy according to which a certain behavior (or "move") is chosen with certainty in a given context.

"Evolutionary dynamics of group interactions on structured populations: a review"

Journal of the Royal Society Interface 10, 20120997 (2013). M.Perc et al.

Interactions among living organisms, from bacteria colonies to human societies, are inherently more complex than interactions among particles and non-living matter. Group interactions are a particularly important and widespread class, representative of which is the public goods game. In addition, methods of statistical physics have proved valuable for studying pattern formation, equilibrium selection and self-organization in evolutionary games.

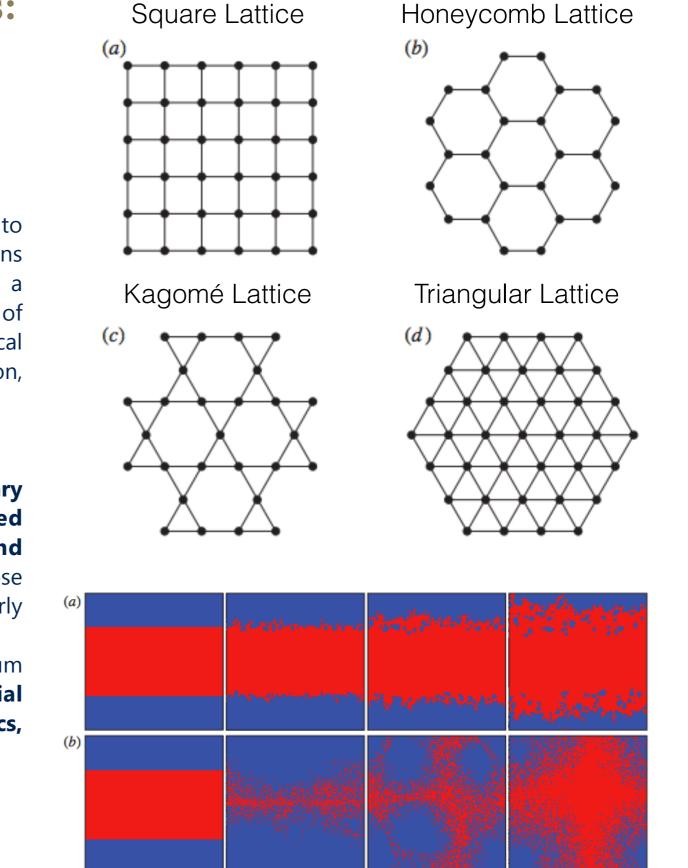
Here, we review recent advances in the study of evolutionary dynamics of group interactions on top of structured populations, including lattices, complex networks and coevolutionary models. We also compare these results with those obtained on well-mixed populations. The review particularly highlights that the study of the dynamics of group interactions, like several other important equilibrium and non-equilibrium dynamical processes in biological, economical and social sciences, benefits from the synergy between Statistical Physics, **Network Science and Evolutionary Game Theory.**

1. Introduction

2. Lattices

4. Coevolutionary rules 5. Outlook

- 3. Complex networks
- 6. Summary



Evolutionary Game Theory

We propose a generalization of the Hawk-Dove Game for an arbitrary number of agents: the **N-person Hawk-Dove Game**

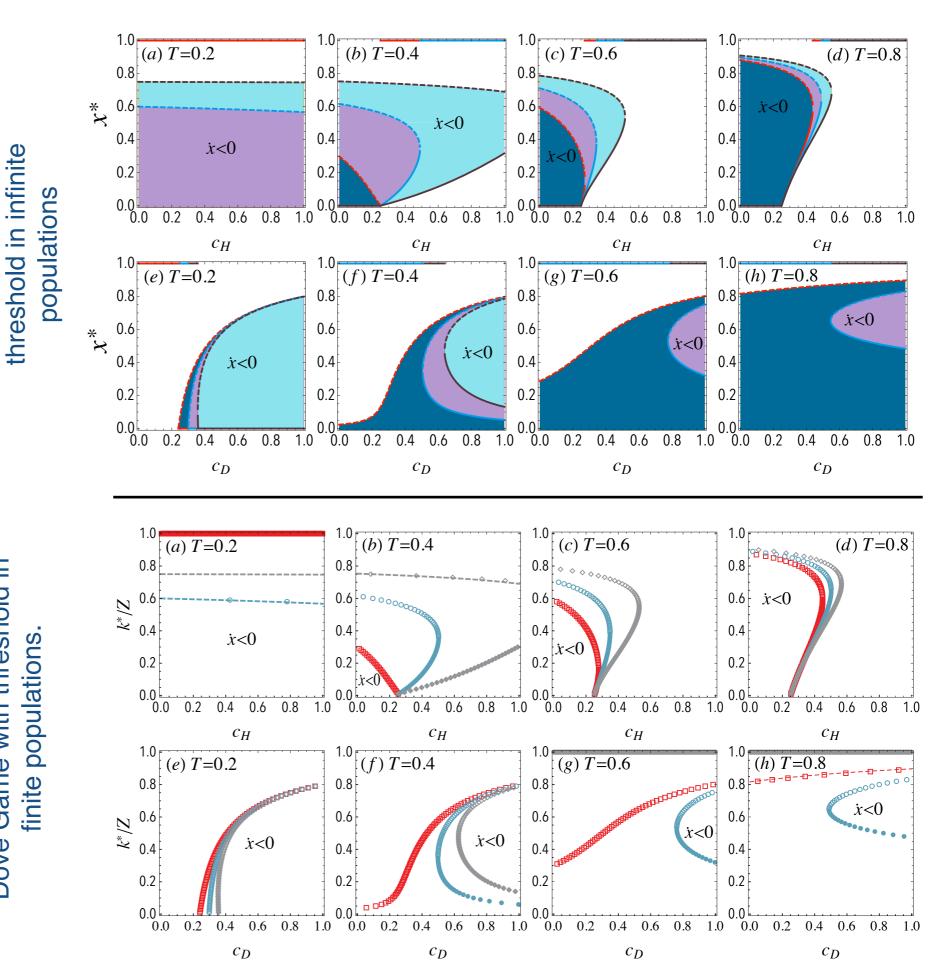
(2017).

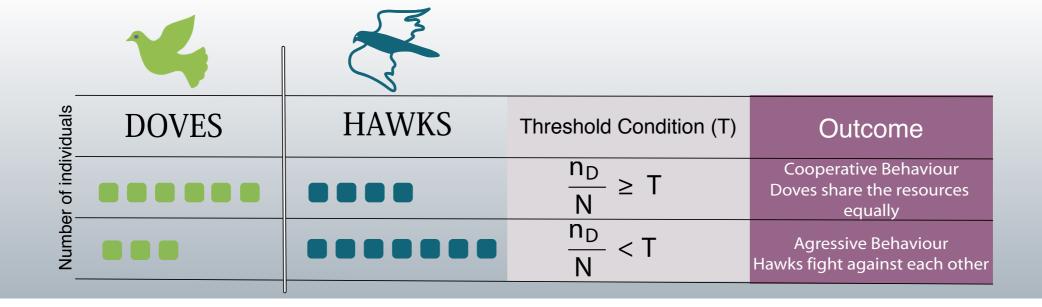


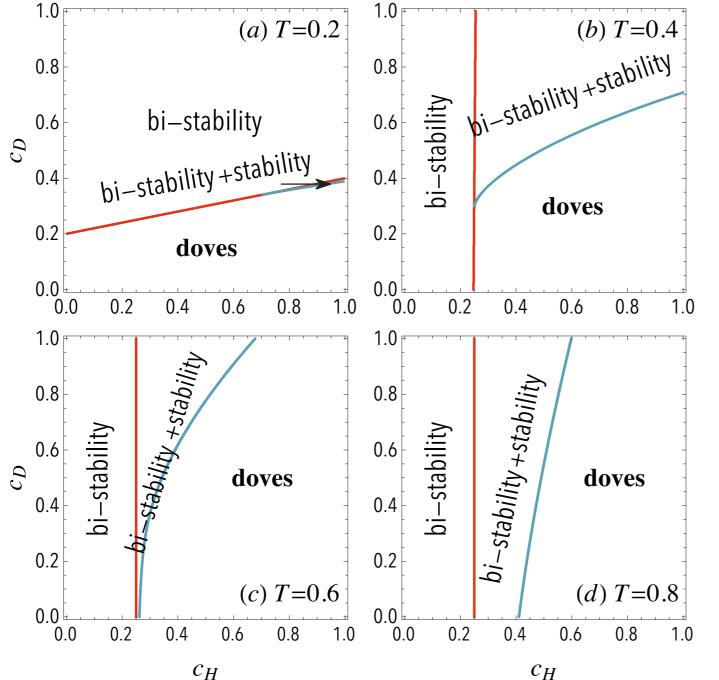
Equilibria of the N-person

Game with

Hawk-Dove







Phase diagrams for the N-person Hawk-Dove Game with threshold. Diagrams show the different regimes in the N-person HDG-T for a sample size of N=5 and an infinite population.

Different panels correspond to different thresholds T=0.2, 0.4, 0.6, 0.8 respectively. Within the bi-stability regime both full-dove state and full-hawk state can be reached from different initial conditions, while the **bi-stability+stability regime** corresponds to coexistence state together with full-dove state. In the *doves* regime, the dynamics always leads to a full-dove state. The resource is taken R=1.

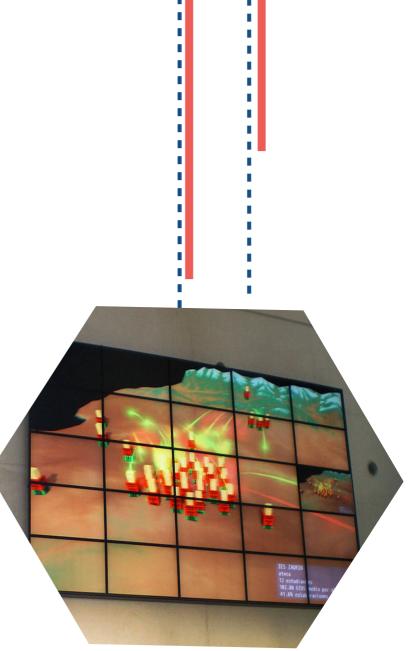
LARGE-SCALE COLLECTIVE EXPERIMENTS

Some experiments...

LARGE-SCALE COLLECTIVE EXPERIMENTS

SO MANY QUESTIONS:

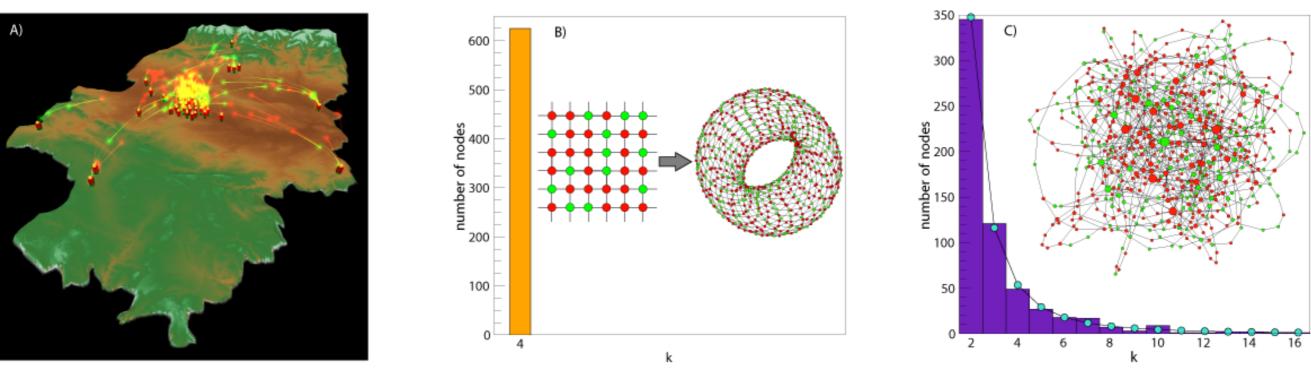
- What are the mechanisms that promote cooperation in humans?
- What is the interplay of social context and cooperative behavior?
- How do we behave in different strategic scenarios?
- Can we build realistic models of how individuals behave and use them to study societal and organizational dynamics?
- Are we able to predict when a collective behavior emerge and unfold?
- Are the laws that govern the Online world the same as those of the Offline world?
- Can we describe with accuracy the different collective phenomena associated to human behavior?



A Large-Scale Experiment in real time with 1229 students, December 2011, Zaragoza, Spain

Heterogeneous Networks

do not promote cooperation among humans playing a PD game



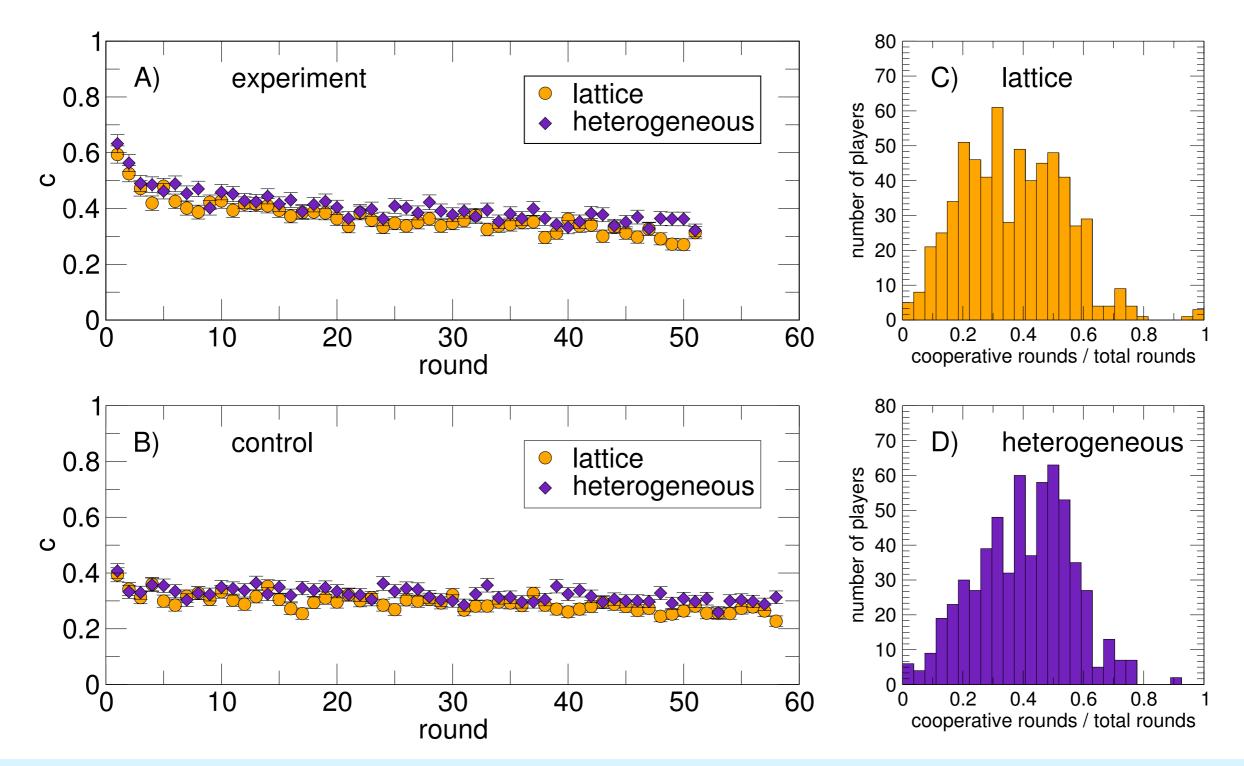
1229 players (coming from 42 schools of Aragon) - 44% male, 56% female

- The level of cooperation reached in both networks is the same, comparable with the level of cooperation of smaller networks or unstructured populations.

- Subjects respond to the cooperation that they observe in a reciprocal manner, being more likely to cooperate if, in the previous round, many of their neighbors and themselves did so.

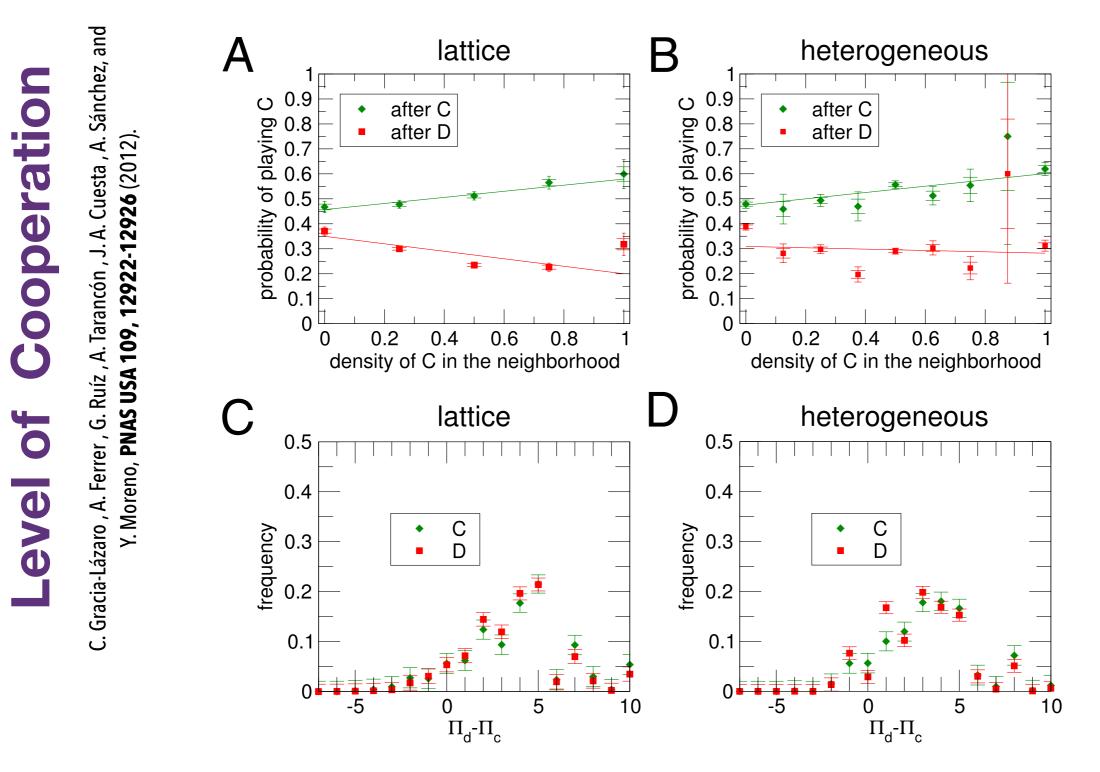
- Our results suggest that population structure has little relevance as a cooperation promoter or inhibitor among humans.

C. Gracia-Lázaro , A. Ferrer , G. Ruíz , A. Tarancón , J. A. Cuesta , A. Sánchez, and Y. Moreno, "Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma", Proceedings of the National Academy of Sciences USA 109, 12922-12926 (2012).



The level of cooperation declines and is independent of the network of contacts. Fraction of cooperative actions (level of cooperation) per round during the experiment (A) and the control (B) for both networks and histograms of cooperative actions in the lattice (C) and the heterogeneous network (D). The histograms (C and D) show the number of subjects ranked according to the fraction of cooperative actions that they perform along the experiment in the two networks.

C. Gracia-Lázaro, A. Ferrer, G. Ruíz, A. Tarancón, J. A. Cuesta, A. Sánchez, and Y. Moreno, "Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma", Proceedings of the National Academy of Sciences USA 109, 12922-12926 (2012).



The structure of the population does not affect the global level of cooperation.
Why?: because players' behavior does not depend on the payoff differences.

C. Gracia-Lázaro , A. Ferrer , G. Ruíz , A. Tarancón , J. A. Cuesta , A. Sánchez, and Y. Moreno, "Heterogeneous networks do not promote cooperation when humans play a Prisoner's Dilemma", Proceedings of the National Academy of Sciences USA 109, 12922-12926 (2012).

What is Reputation?

Objectives:

- Is reputation the driving mechanism observed behind the levels of cooperation when individuals play a PD' game?

and

Ferrer, Y. Moreno,

Gracia-Lazaro, A.

Cuesta, C.

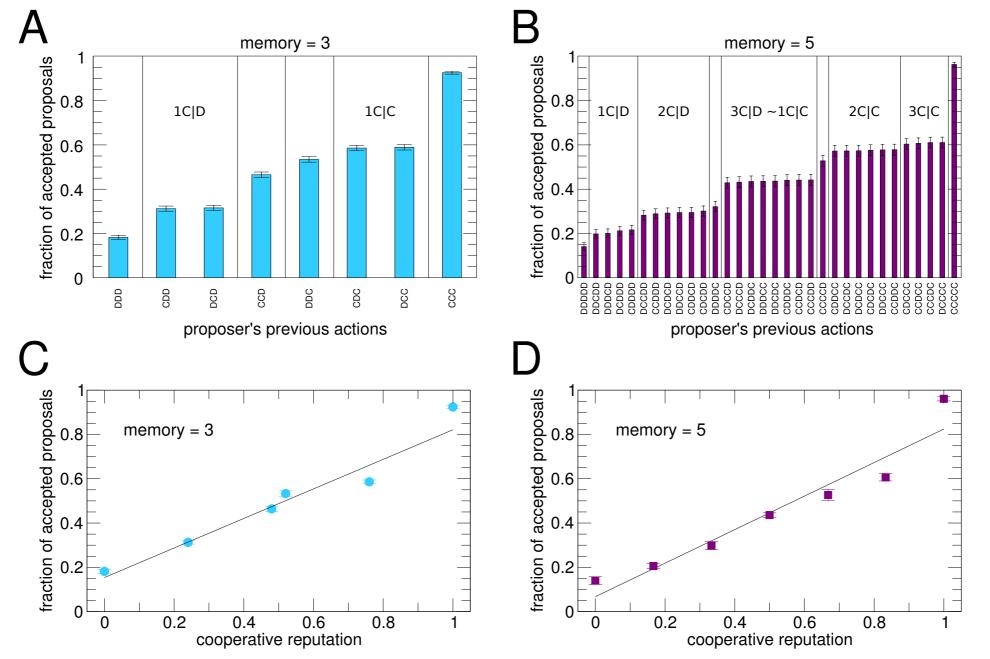
J. A.

Scientific Reports 5:7843 (2015)

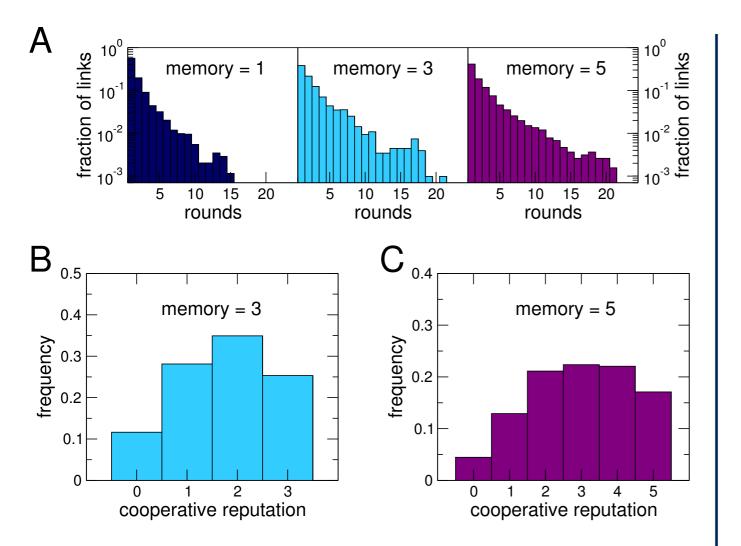
Sanchez,

Ā

- How can we quantify reputation?



Reputation is a weighted combination of average cooperation and last action, and it strongly conditions linking.

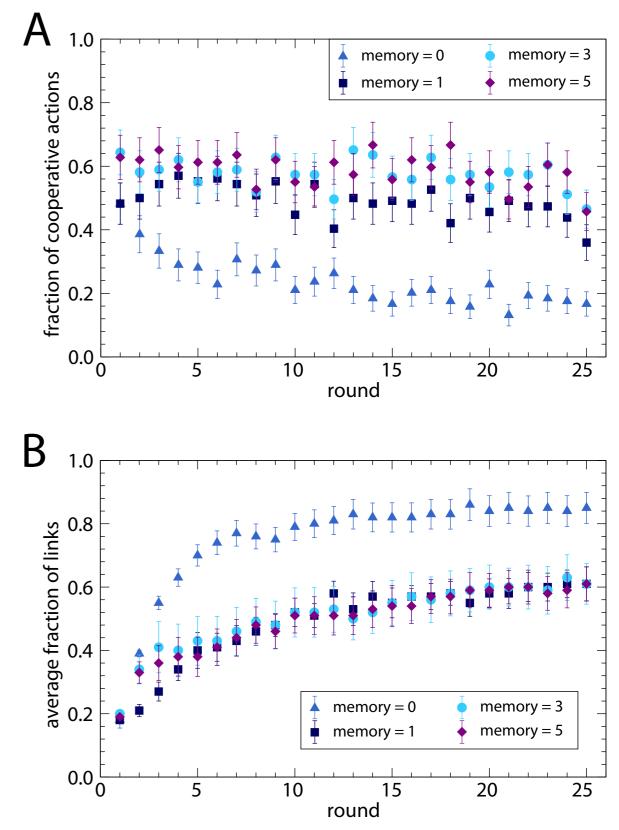


Subjects try to hold a high reputation, but not the highest. The histogram of link lifetimes shows a fast exponential decay.

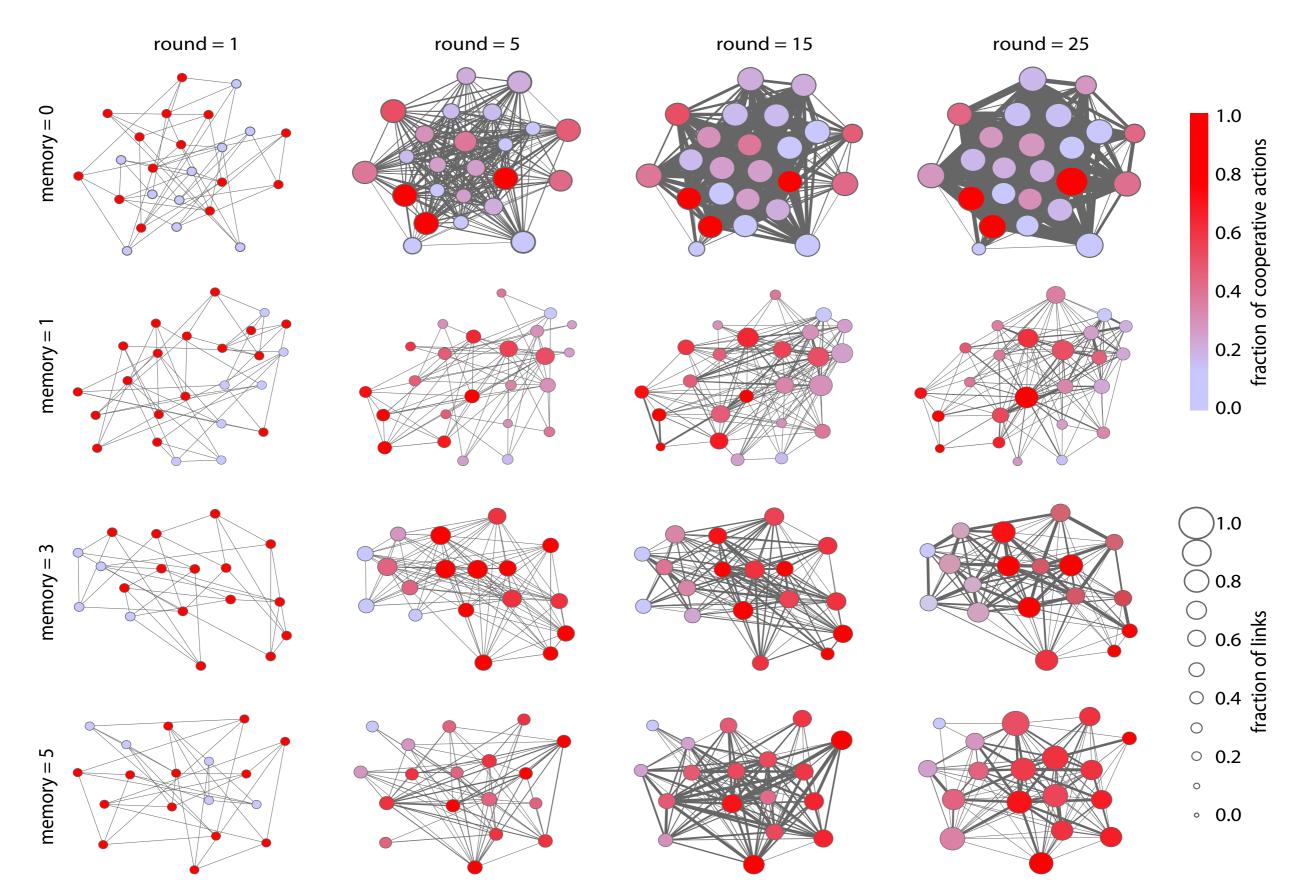
(A; average lifetimes are 2.75 rounds for m= 1, 3.21 for m= 3, and 3.23 for m= 5). This is a consequence of the fact that most individuals keep a record of 2 cooperative actions out of 3.

(**B**) or 2–4 out of 5.

 (\mathbf{C}) ; in other words, subjects often defect but not too much as that would ruin their reputation. This sporadic defection has a drastic effect on the linking dynamics because reputation is very much influenced by subjects' last action.



The level of cooperation is significantly higher when the past actions record of players to whom to connect is available.



"What it lies in our power to do, It lies in our power not to do"

Aristotle



Social Dilemmas & Human Behavior

Nectunt Lab (Experimental Economics Lab, BIFI Institute -University of Zaragoza)



http://cosnet.bifi.es @cosnet_bifi

Complex Systems & Networks Lab (COSNET)

Main Collaborators:

- Carlos Gracia-Lázaro (Complex Systems and Networks Lab -COSNET-, Nectunt Lab ZGZ, Institute for Biocomputation & Physics of Complex Systems, BIFI, Zaragoza)
- Raquel Alvarez Baños (Scientific Software Developer at STFC, Oxford, United Kingdom)
- Alfredo Ferrer Marco (Complex Systems and Networks Lab -COSNET-, Institute for Biocomputation & Physics of Complex Systems, BIFI, Zaragoza and Kampal, Software developer)
- Sandra González Bailón (Annenberg School for Communication & Warren Center for Network and Data Sciences)
- Javier Borge-Holthoefer (Interdisciplinary Internet Institute (IN3) Universitat Oberta de Catalunya- UOC)
- Sandro Meloni (Complex Systems and Networks Lab -COSNET-, Institute for Biocomputation & Physics of Complex Systems, BIFI, Zaragoza)
- Angel Sánchez (Interdisciplinary Complex Systems Group -GISC- & BIFI Institute)
- José Angel Cuesta (Interdisciplinary Complex Systems Group -GISC- & BIFI Institute)
- James Gleeson (MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland)
- Matjaz Perc (Director of the Complex Systems Center Maribor, University of Maribor, Slovenia)
- COSNET Lab PhD: Alberto Aleta y Felipe M. Cardoso