# **Geography and community in complex networks**

Vincent Blondel LIDS, MIT On leave from UCLouvain (Belgium)





Sontag Fest May 2011







VB, Sontag, Vidyasagar, Willems, 1999

### Panel on Theoretical Challenges Moderator: M.Vidyasagar Panel: Peter Caines, Jan van Schuppen

**3. System approximation in the 2-norm** A.C. Antoulas

**16. A stabilization problem** Roger Brockett

**25. Input-output gains of switched linear systems** J. P. Hespanha, A. S. Morse

**40. Control-Lyapunov functions** Eduardo D. Sontag

**46. Shift policies in QR-like algorithms and feedback control of self-simil**<sup>*a*</sup> Paul Van Dooren, Rodolphe Sepulchre

**50. Lyapunov theory for high order differential systems** Jan C. Willems

**51. Performance lower bound for a sampled-data signal reconstruction** Yutaka Yamamoto, Shinji Hara

**77. Important computational complexity problems in systems and control** *Peter Caines* 



VB, Sontag, Vidyasagar, Willems, 1999

Panel on Theoretical Challenges Moderator: M.Vidyasagar Panel: Peter Caines, Jan van Schuppen

«Clearly, the most important and pressing problems in system and control are problems related to computational complexity. These are problems that have had an enormous impact in control and will continue to do so for the next decades.»

Peter Caines at SontagFest, May 2011

**77. Important computational complexity problems in systems and control** *Peter Caines* 

Technology Review, in English | en Experier | ast Deutach, in Italiano | 41% | in India



NetMob	Workshop on the Analysis of Mobile Phone Networks A satellite workshop to NetSci 2010 Tuesday, May 11, 2010 MIT, Cambridge, MA
NetMob2011	Given the success of NetMob2010, we consider the possibility of organizing a NetMob2011. If you wish to be included on the NetMob mailing list, please send an email to sympa2@listes.uclouvain.be with "subscribe netmob yourname" in the subject line (where "yourname" is your first and last name). You can also subscribe/unsubscribe by going to https://listes- 2.sipr.ucl.ac.be/sympa/info/netmob
Introduction	Mobile phone datasets have become widely available in recent years and have opened the possibility to improve our understanding of large-scale social networks by investigating how people exchange information, build trust, create markets and develop social interactions. Mobile phone data is also helping us understand complex processes such as the spread of information and viruses or transportation and the use of urban infrastructures.
	This workshop will consist of a number of contributed talks on the analysis of mobile phone networks. The workshop format is flexible: no registration fees, a simplified submission procedure, and the possibility to present recent results or results submitted elsewhere.
Practical	Date: Tuesday May 11, 2010 (this is the day prior to the conference NetSci).
information	Location: On the sixth floor of the newly built Media Lab (building E14 on MIT campus, map available here).
	Registration: Attendance is free of charge but, due to limited seating, registration is compulsory. If you wish to register please send an email to notmob@uclouvain.bo. Registration will be processed on a first-come first-serve basis. Although there is no registration fee for the workshop, participants are of course encouraged to also participate (and register) in the NetSci conference.
	We have have received an unexectedly large number of registrations to the workshop. The workshop has been moved to a larger space (the multi-media hall of the Media Lab). All those who have registered by sending an email or through the NetSci website are welcome to attend.
Submissions	All contributions that deal with the analysis of mobile phone datasets are welcome.
	Authors are invited to submit an abstract (one to three pages) by the deadline of March 5, 2010. Submissions should include the title, author(s), affiliation(s) and e-mail address(es) on the first page. There will be no published proceedings; the material submitted to the workshop may also be submitted elsewhere.
	Electronic submission of manuscripts in PDF format is required. Please send your manuscript directly to netmob@uclouvain.be by March 5, 2010.
	The evaluation of submitted abstracts will be organized by the scientific committee and decisions will be made by March 26, 2010. Once an abstract has been accepted for presentation, at least one author is required to attend the workshop and present the paper. In case too many abstracts are selected, some of these may be moved to a special session taking place the next day at the NetSci 2010 conference.
Program	The program is available here (PDF format).
Book of abstracts	The book of abstracts is available here (5:5 MB, PDF format).
Scientific committee	Chair: <u>Vincent Blondel</u> , UCLouvain (Belgium) Laszlo Barabasi, Northeastern University Rob Clusters, Retirch Telecom (UK)



Recent technological and mathematical developments have opened the possibility to considerably i our understanding of how information flows and decisions are made in large social networks. workshop, we bring together researchers from different communities working on information prop and decision making in social networks to investigate both rigorous models that highlight capabili limitations of such networks as well as empirical and simulations studies of how people ex information, influence each other, make decisions and develop social interactions.

This workshop is being organized by the Laboratory for Information and Decision Systems.

#### Organizers

Vincent Blondel, UCLouvain (Belgium) and LIDS, MIT <u>Munther Dahleh</u>, LIDS, MIT <u>Asu Ozdaglar</u>, LIDS, MIT John Tsitsiklis, LIDS, MIT

#### Scientific committee

Chair: <u>Vincent Blondel</u>, UCLouvain (Belgium) and LIDS, MIT <u>Daron Acemoglu</u>, Economics, MIT <u>Sinan Aral</u>, New York University <u>Albert-László Barabási</u>, Northeastern University



# Large networks

- Web graph
- Internet graph
- Email exchange networks
- Blogs networks
- Citation networks
- Social networking sites (facebook, linkedIn, etc)
- Instant messages
- Mobile phone networks
- Twitter
- Collaboration graphs
- •

# **Communities in networks**



# Why look for communities?

- Visualisation
- Structural organisation of the network
- Analysis (information propagation, robustness, cohesive,...)
- Time-evolution



Physics Reports 486 (2010) 75-174



#### Community detection in graphs

#### Santo Fortunato\*

Complex Networks and Systems Lagrange Laboratory, ISI Foundation, Viale S. Severo 65, 10133, Torino, I, Italy

#### ARTICLE INFO

#### ABSTRACT

Article history: Accepted 5 November 2009 Available online 4 December 2009 editor: I. Procaccia

Keywords: Graphs Clusters Statistical physics The modern science of networks has brought significant advances to our understanding of complex systems. One of the most relevant features of graphs representing real systems is community structure, or clustering, i.e. the organization of vertices in clusters, with many edges joining vertices of the same cluster and comparatively few edges joining vertices of different clusters. Such clusters, or communities, can be considered as fairly independent compartments of a graph, playing a similar role like, e.g., the tissues or the organs in the human body. Detecting communities is of great importance in sociology, biology and computer science, disciplines where systems are often represented as graphs. This problem is very hard and not yet satisfactorily solved, despite the huge effort of a large interdisciplinary community of scientists working on it over the past few years. We will attempt a thorough exposition of the topic, from the definition of the main elements of the problem, to the presentation of most methods developed, with a special focus on techniques designed by statistical physicists, from the discussion of crucial issues like the significance of clustering and how methods should be tested and compared against each other, to the description of applications to real networks.

© 2009 Elsevier B.V. All rights reserved.

#### Contents

1.	Introduction			
2.	Communities in real-world networks			
3.	Elements of community detection			82
	3.1. Computational complexity			83
	3.2.	Communities		
		3.2.1.	Basics	83
		3.2.2.	Local definitions	84
		3.2.3.	Global definitions	85
		3.2.4.	Definitions based on vertex similarity	
	3.3. Partitions			
		3.3.1.	Basics	87
		3.3.2.	Quality functions: Modularity	88
4.	Traditional methods			90
	4.1. Graph partitioning			90
	4.2. Hierarchical clustering			93
	4.3.	Partitio	nal clustering	93
	4.4	Spectra	clustering	94
5. Divisive algorithms				96
4.4. Spectral clustering.     5. Divisive algorithms     5.1. The algorithm of Girvan and Newman			prithm of Girvan and Newman	97
			nethods	

\* Tel.: +39 011 6603090; fax: +39 011 6600049. E-mail address: fortunato@isi.it.

E-mail address: fortu

0370-1573/\$ - see front matter © 2009 Elsevier B.V. All rights reserved. doi:10.1016/j.physrep.2009.11.002

#### Communities in Networks

Mason A. Porter, Jukka-Pekka Onnela, and Peter J. Mucha

Introduction: Networks and Communities "But although, as a matter of history, statistical mechanics owes its origin to investigations in thermodynamics, it seems eminently worthy of an independent development, both on account of the elegance and simplicity of its principles, and beccuse it yields new results and places old truths in a new light in departments quite outside of thermodynamics."

 Josiah Willard Gibbs, Elementary Principles in Statistical Mechanics, 1902 [47]

ROM AN ABSTRACT PERSPECTIVE, the term network is used as a synonym for a mathematical graph. However, to scientists across avariety of fields, this label means so much more [13, 20,4-8, 33, 88, 120, 124, in sociology, each *nade* (or vertex) of a network represents an *agent*, and a pair of nodes can be connected by a link (or degle) that signifies some social interaction or tie between them (see Figure 1 Mason A. Porter, Oxford Centre for Industrial and Applied Mathematics, Mathematical Institute, University of Oxford, and CABDyN Complexity Centre, University of Oxford, and CABDyN Complexity Centre, University of Oxford. His email address is porterminatins. ox. a.e. uk.

Jukka-Pekka Onnela, Harvard Kennedy School, Harvard University; Department of Physics, University of Oxford; CABDyN Complexity: Centre, University of Oxford; and Department of Biomedical Engineering and Computational Science, Helsinki University of Technology. His email address is jp\_onne1a@ksg\_harvard.edu.

Peter J. Mucha, Carolina Center for Interdisciplinary Applied Mathematics, Department of Mathematics, University of North Carolina, and Institute for Advanced Materials, Nanoscience and Technology, University of North Carolina. His email address is muchalumc, edu.

the number of edges connected to it and a strength given by the total weight of those edges. Graphs can represent either man-made or natural constructs, such as the World Wide Web on neuronal synaptic networks in the brain. Agents in such networked systems are like particles in traditional statistical mechanics that we all know and (presumably) lowe, and the structure of interactions between agents reflects the microscopic rules that govern their behavior. The simplest types of links are binary pairwise connections, in which one only cares about the presence or absence of a tic. However, in many situations, links can also be assigned a direction and a (positive or negative) weight to designate different statistical physics is concerned with Traditional statistical physics is concerned with

for an example). Each node has a *dearee* given by

Traditional statistical physics is concerned with the dynamics of ensembles of interacting and noninteracting particles. Rather than tracking the motion of all of the particles simultaneously, which is an impossible task due to their tremendous number, one averages (in some appropriate manner) the microscopic rules that govern the dynamics of individual particles to make precise statements of macroscopic particles to make precise statements and density [112]. It is also sometimes possible to make comments about intermediate (*massocopi cost statutes*, which the between the microscopic and macroscopic worlds; they are large enough that it is reasonable to discuss their collective properties but small enough so that those properties are obtained through averaging over smaller numbers of constituent items. One can similarly the nodes of a network, with some set of microscopic interaction rules and attempt to derive the southmer socole and attempt to derive the

1082

VOLUME 56, NUMBER 9

What the papers say

The art of community detection

NOTICES OF THE AMS

Natali Gulbahce<sup>1,2</sup>\* and Sune Lehmann<sup>1,2</sup>

#### Summary

Networks in nature possess a remarkable amount of structure. Via a series of data-driven discoveries, the cutting edge of network science has recently progressed from positing that the random graphs of mathematical graph theory might accurately describe real networks to the current viewpoint that networks in nature are highly complex and structured entities. The identification of high order structures in networks unveils insights into their functional organization, Recently, Clauset, Moore, and Newman.<sup>(1)</sup> introduced a new algorithm that identifies such heterogeneities in complex networks by utilizing the hierarchy that necessarily organizes the many levels of structure. Here, we anchor their algorithm in a general community detection framework and discuss the future of community detection. BioEssays 30:934-938, 2008. © 2008 Wiley Periodicals, Inc.

#### Structure everywhere

The view that natworks are accentially random was challenged

interact with more than 200 other proteins, contrary to proteins that interact with only a few other proteins.

Various local to global measures have been introdu unveil the organizational principles of complex networf Maslov and Snepen<sup>(10)</sup> discovered that who links to can depend on node degree; in many biological network degree nodes systematically link to nodes of low degree disassortativity decreases the likelihood of cross talk be functional modules inside the cell and increases i robustness. Other networks, for example social networ are highly assortative—in these networks nodes with : degree tend to link to each other.

Going beyond the properties of single nodes and p nodes, the natural next step is to consider structure include several nodes. Interestingly, a few select mc three to four nodes are ubiquitous in real networks<sup>(12)</sup> most others occur only as often as they would at ra or are actively suppressed. Other measures of local str

# Geography and community in complex networks





### Louvain method for community detection (and modularity)

[VB, Guillaume, Lambiotte, Lefèvre, 2008]

### **Communities in a mobile phone network**

[Lambiotte, VB et al., 2009]

### **Geography in community detection**

[Krings, Calabrese, Ratti, VB, 2009] [VB, Krings, Thomas, 2010] [Expert, Evans, VB, Lambiotte, PNAS, 2011]

### **Eigenvectors and communities**

[Cucuringu, Vandooren, VB, 2011]



Vincent D. Blondel, Jean-Loup Guillaume, Renaud Lambiotte, Etienne Lefebvre, Fast unfolding of communites in large networks, Journal of Statistical Mechanics: Theory and Experiment, 1742-5468, P10008, 2008. Web: 118M/1G

Divide and Conquer: Partitioning Online Social Networks Josep M. Pujol, Vijay Erramilli, Pablo Rodriguez arXiv, 2010 **Twitter: 2.4M/38M** 

Mapping search relevance to social networks Jonathan Haynes, Igor Perisic Proceedings of the 3rd Workshop on Social Network Mining and Analysis, 2010 Linkedin, 21M

Tracking the Evolution of Communities in Dynamic Social Networks Greene, D.; Doyle, D.; Cunningham, P.; International Conference on Advances in Social Networks Analysis and Mining (ASONAM), 2010 **Mobile phone, 4M/100M** 

Real World Routing Using Virtual World Information Pan Hui, Sastry N. International Conference on Computational Science and Engineering, 2009 Flickr 1.8M/22M, LiveJournal 5.3M/77M, YouTube 1.1M/4.5M

Community structure in audio clip sharing Gerard Roma, Perfecto Herrera International Conference on Intelligent Networking and Collaborative Systems, INCoS 2010 **Freesound** 

Subject clustering analysis based on ISI category classification Lin Zhang, Xinhai Liu, Frizo Janssens, Liming Liang and Wolfgang Glänzel Journal of Informetrics, Volume 4, Issue 2, April 2010 ISI 6M papers





LIVEJOURNAL



twitter

flickr

You Tube



NetworkX Home | Download | Developer Zone | Documentation | Blog »

#### High productivity software for complex networks

NetworkX is a Python package for the creation, manipulation, and study of the structure, dynamics, and functions of complex networks.

#### **Quick Example**





#### Download

Current version: 1.3

Get NetworkX from the Python Package Index, or install it with:

modules | index

easy\_install networkx

#### Questions? Suggestions?

Join the Google group:

your@email Subscribe

You can also open a ticket at the NetworkX Developer Zone.











Quality of a partition of a network in communities



### Quality of a partition of a network in communities



# edges in communities - expected # edges in communities

total number of edges

### Quality of a partition of a network in communities



# edges in communities - expected # edges in communities

total number of edges

### Quality of a partition of a network in communities



### Quality of a partition of a network in communities



m edges,  $d_i$  degree of node i

expected # edges between node *i* and node  $j = (d_i d_j)/(2m)$ 

### Quality of a partition of a network in communities



# edges in communities - expected # edges in communities

total number of edges

*m* = 25, expect **0.3** edges

m edges,  $d_i$  degree of node i

expected # edges between node *i* and node  $j = (d_i d_j)/(2m)$ 

### Quality of a partition of a network in communities



*m* nodes,  $d_i$  degree of node *i* 

expected # edges between node *i* and node  $j = (d_i d_j)/(2m)$ 

expected # edges in community  $Sum_{i, j \text{ in } C} (d_i d_j)/(2m)$ 

### Quality of a partition of a network in communities



# edges in communities - expected # edges in communities

total number of edges

Expected **11.56 2.56** Observed **15** 6

m nodes,  $d_i$  degree of node i

expected # edges between node *i* and node  $j = (d_i d_j)/(2m)$ 

expected # edges in community Sum<sub>i, j in C</sub> (d<sub>i</sub> d<sub>j</sub>)/(2m)

### Quality of a partition of a network in communities



 Expected
 11.56
 2.56

 Observed
 15
 6

*m* nodes,  $d_i$  degree of node *i* expected # edges between node *i* and node  $j = (d_i d_j)/(2m)$ expected # edges in community  $Sum_{i, j \text{ in } C} (d_i d_j)/(2m)$ 



0.275





0.32





# The (greedy) Louvain method

Initially every node forms a community.

For every node *i*, insert node *i* in a neighboring community that maximizes the resulting modularity gain.

Repeat until a local maximum is attained.

Construct the resulting network of communities and repeat the construction on the network of communities.

This way we construct a hierarchy of communities.



# The (greedy) Louvain method



118M nodes/1B links in 152mn

- Belgian phone call network
  - 6 months of communications
  - One Belgian main operator
- Network :
  - 2.6 M customers
  - 800M voice/text messages
  - language information (Dutch, English, French, German)
  - location information (ZIP)

### Mobile phone network



### **Distribution of calls received**



### **Duration with distance**


# **Connection with distance**



[Onnela, Arbesman, Barabasi, Christakis, 2010]

[Lambiotte, VB et al., 2008]









#### Bye bye Belgium?

With deep divisions between its French and it's a wonder Belgium has stayed united for heading for meltdown. Jon Henley reports



#### No Love Lost: Is Belgium About to Break in Two?

By Leo Cendrowicz / Brussels

#### Jon Henley

#### Tuesday REUTERS

» Print

#### Guardian

Belgium is handsome trains are does the c

This copy is for your personal, non-commercial use only. To order presentation-ready copies for distribution to colleagues, directs or customers, use the Reprints tool at the top of any article or visit; www.reutersreprints.com.

#### Flemish separatists triumph in Belgian election

#### By Philip Blenkinsop

BRUSSELS (Reuters) - A Flemish party that wants to split Belgium triumphed in a parliamentary election on Sunday, a result that could make it hard to form a coalition quickly and deliver austerity to contain a rising national debt.

Belgium can ill afford drawn-out coalition talks because policy paralysis could make the country more vulnerable on financial markets that are closely watching a sovereign debt crisis among the 16 countries that use the euro.

The N-VA (New Flemish Alliance) was the strongest party in the Dutch-speaking Flanders region of northern Belgium. It won more votes there than the French-speaking Socialists (PS) secured in separate voting in the southern, Francophone region of Wallonia, nearly complete results showed.

"The N-VA has won the election today," N-VA leader Bart De Wever, 39, told cheering, flag-waving supporters

# Le Monde.fr

### "Oui, il faut se préparer à la fin de la Belgique"

05.09.10 | 16h29

Près de trois mois après les élections, la Belgique se retrouve de nouveau dans la crise après la démission, entérinée ce week-end, d'Elio di Rupo, chef de file du Parti socialiste francophone qui a renoncé à son tour à tenter de former un gouvernement. Le roi Albert II a accepté samedi soir la démission du socialiste wallon, qui n'a pu combler le fossé entre néerlandophones et franchophones paralysant la vie politique belge depuis plus de trois ans.

Conséquence de ce blocage, le tabou de la scission de la Belgique commence à tomber dans le monde politique francophone, dont plusieurs représentants de premier plan ont ouvertement évoqué cette éventualité, dimanche 5 septembre, en raison des difficultés à s'entendre sur l'avenir du pays avec les Flamands.



t, fourth-round victory over her compatriot country could be to produce two stunningly and even that sports rivalry between the division between the country's Dutch and in's from the French.

ain excellent, of course, as f petrol, which is new.

# Poelvoorde : « Gardons nos barbes jusqu'à ce que la Belgique se relève »

Rédaction en ligne jeudi 13 janvier 2011, 00:09

#### 🖪 📲 😵 🎘 🌀 🔚



Benoît Poelvoorde suggère à la gent masculine de ne plus se raser « jusqu'à ce que la Belgique se relève » et soit dotée d'un gouvernement, alors que le royaume traverse la plus longue crise politique de son histoire.



#### es portfollos

 L'équipe de "Rien à Déclarer" en visite au Soir

#### lire aussi

- Chat : trois jeunes appellent à manifester
- Camping16 rue de la Loi, une manif virtuelle
- Pas de gouvernement ? Appel à manifester le 23 janvier
- La carte blanche : « Si ça continue, je ne vote plus : je tire ! »
- Geluck réagit à la proposition de Poelvoorde
- La Belgique championne du monde ?
- sur le web
- La vidéo sur RTL
- Le site de "Shame"



## Dernie



dép II y a Pier

d'ar d'ar dire



# Geography and community in complex networks





#### Louvain method for community detection (and modularity)

[VB, Guillaume, Lambiotte, Lefèvre, 2008]

#### **Communities in a mobile phone network**

[Lambiotte, VB et al., 2009]

#### **Geography in community detection**

[Krings, Calabrese, Ratti, VB, 2009] [VB, Krings, Thomas, 2010] [Expert, Evans, VB, Lambiotte, PNAS, 2011]









## **Gravitational law of interactions**

[Krings, Calabrese, Ratti, VB, 2008]





#### THE WALL STREET JOURNAL.

WSJ.com

WHAT THEY KNOW APRIL 23, 2011

## The Really Smart Phone

Researchers are harvesting a wealth of intimate detail from our cellphone data, uncovering the hidden patterns of our social lives, travels, risk of disease—even our political views.

By ROBERT LEE HOTZ



Photo-illustration by Adam Magyar

"Phones can know," says an MIT researcher. 'People can get this god's-eye view of human behavior."

Apple and Google may be intensifying p concerns by tracking where and when p use their mobile phones—but the true f of consumer surveillance is taking shap the cellphones at a weather-stained apa complex in Cambridge, Mass.

For almost two years, Alex Pentland at a Massachusetts Institute of Technology l tracked 60 families living in campus qua via sensors and software on their smart —recording their movements, relations moods, health, calling habits and spend this wealth of intimate detail, he is findi patterns of human behavior that could a how millions of people interact at home and play.

Through these and other cellphone rese projects, scientists are able to pinpoint "influencers," the people most likely to others change their minds. The data car