



Medindo Qualidade em Seguro-Saúde Usando Graph Analytics

Claudio Pinhanez - IBM Research Brasil
csantosp@br.ibm.com - @cinhanez

FEBRABAN



Claudio Pinhanez

IBM Research – Brazil
Leader Social Data Analytics Research

Professor, Univ. of São Paulo, 1987-1993

PhD MIT Media Lab 1993-1999

IBM Research, 1999-2008

IBM Research - Brazil, 2008-2016

minerando dados de relações de pessoas no contexto de seguro-saúde

Agenda: Valor



Agenda: Ferramentas

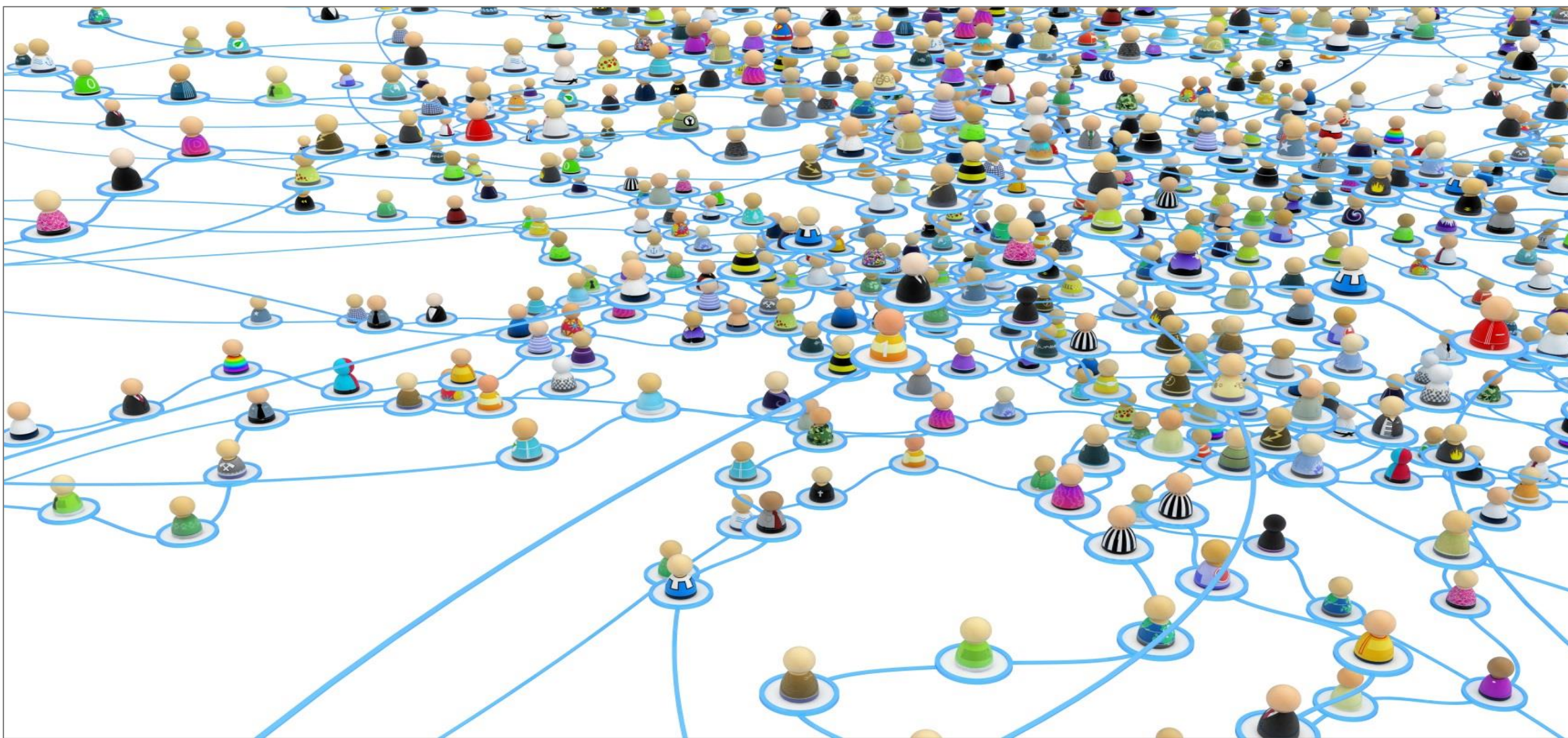


Agenda: Um Caso Real

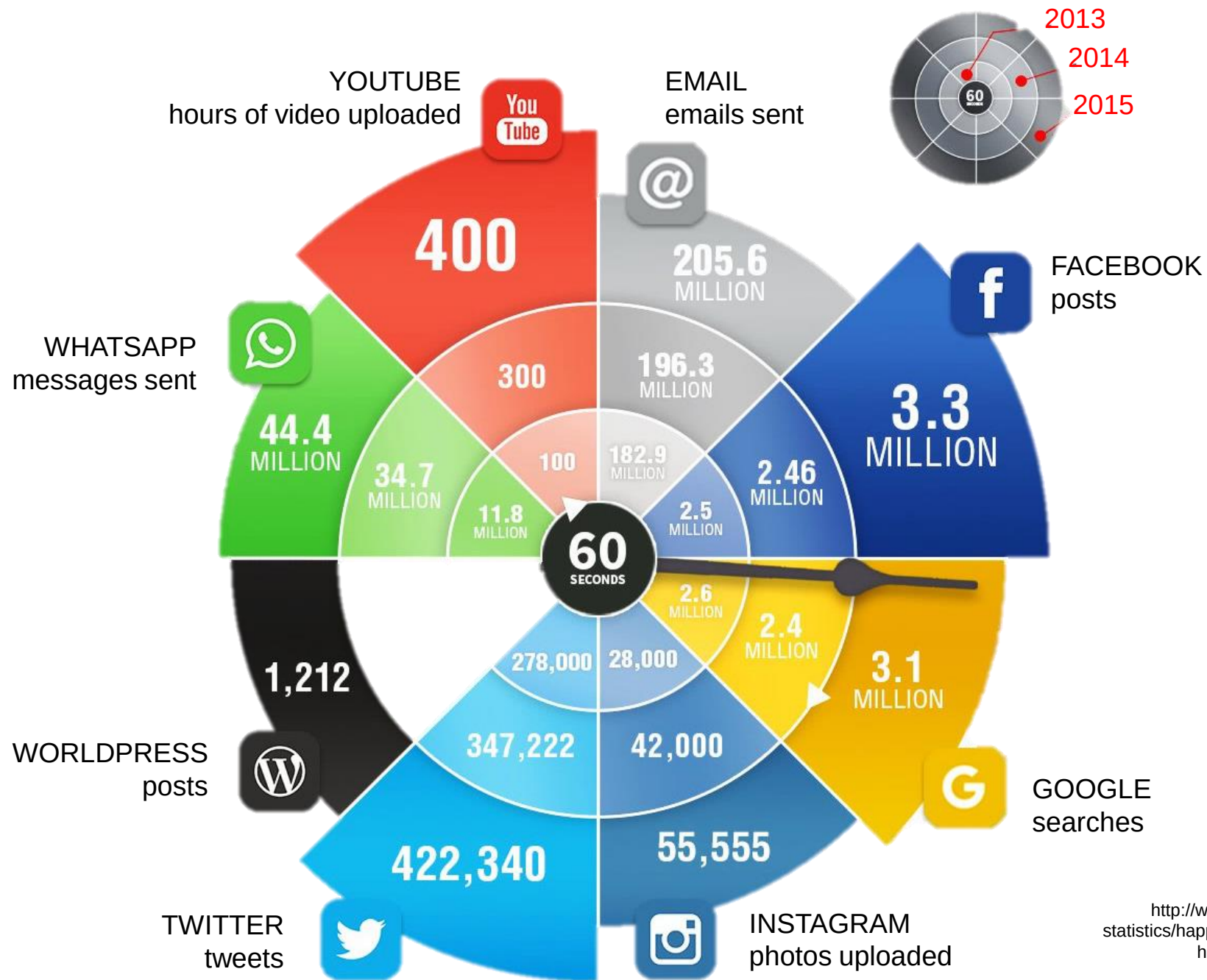




O PODER DAS REDES DE PESSOAS



redes de pessoas transformadas pela digitalização da sociedade

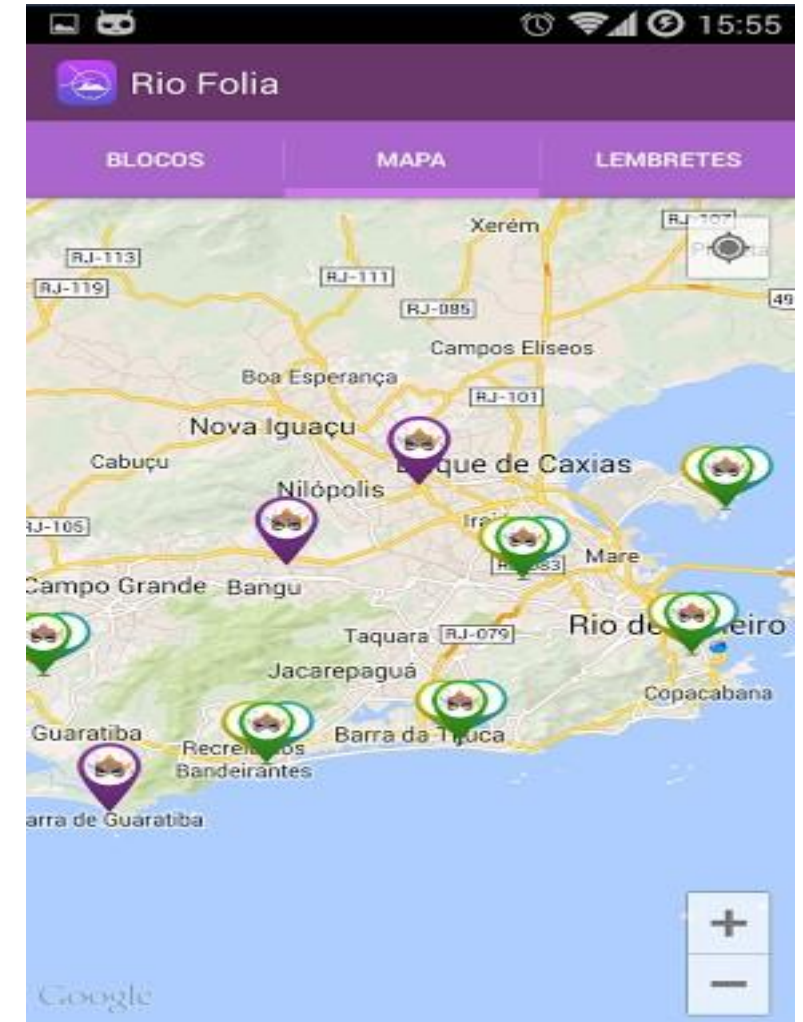


Manifestações de 1 milhão de pessoas sem organização central



Rio, junho 2013

Carnaval de rua em São Paulo: usando mídia social para achar a folia



Trump: maciço uso de “earned media”







Donald J. Trump ✓
@realDonaldTrump
#MakeAmericaGreatAgain #Trump2016
#Instagram:
instagram.com/realdonaldtrum...
#Facebook:
m.facebook.com/DonaldTrump/

TWEETS
32.3K

FOLLOWING
40

FOLLOWERS
9.2M

LIKES
92

Tweets

Tweets & replies

Media



Donald J. Trump @realDonaldTrump · 2h
Crooked Hillary Clinton is totally unfit to be our president-really bad judgement and a temperament, according to new book, which is a mess!

2.7K 7.8K

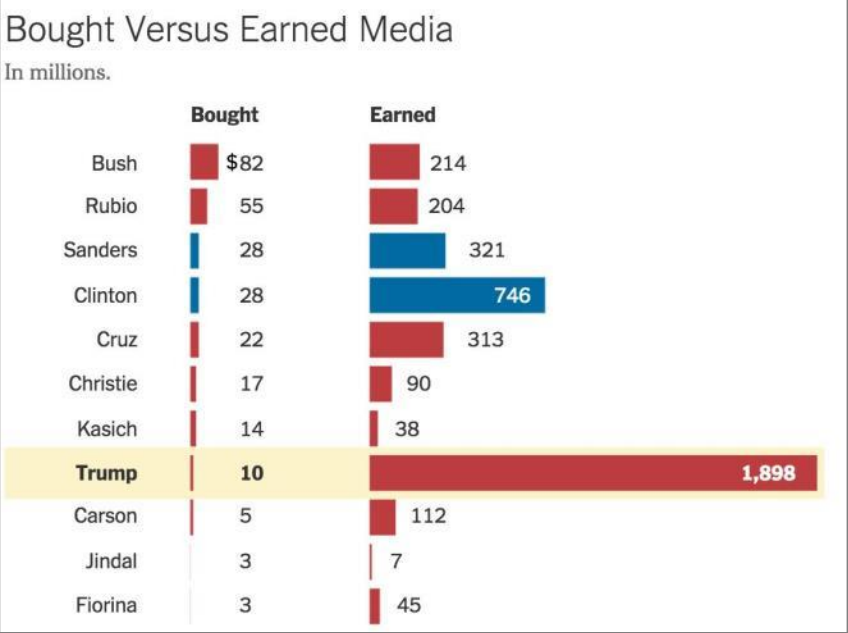
Who to follow · Refr



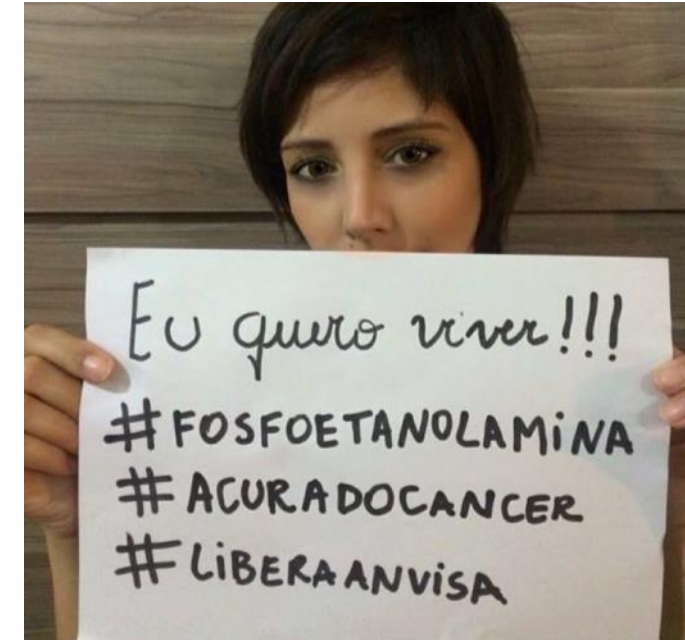
Brian Hazard
Follow



leysia palen
Follow



Pílula do câncer: ciência popular atropela ciência oficial

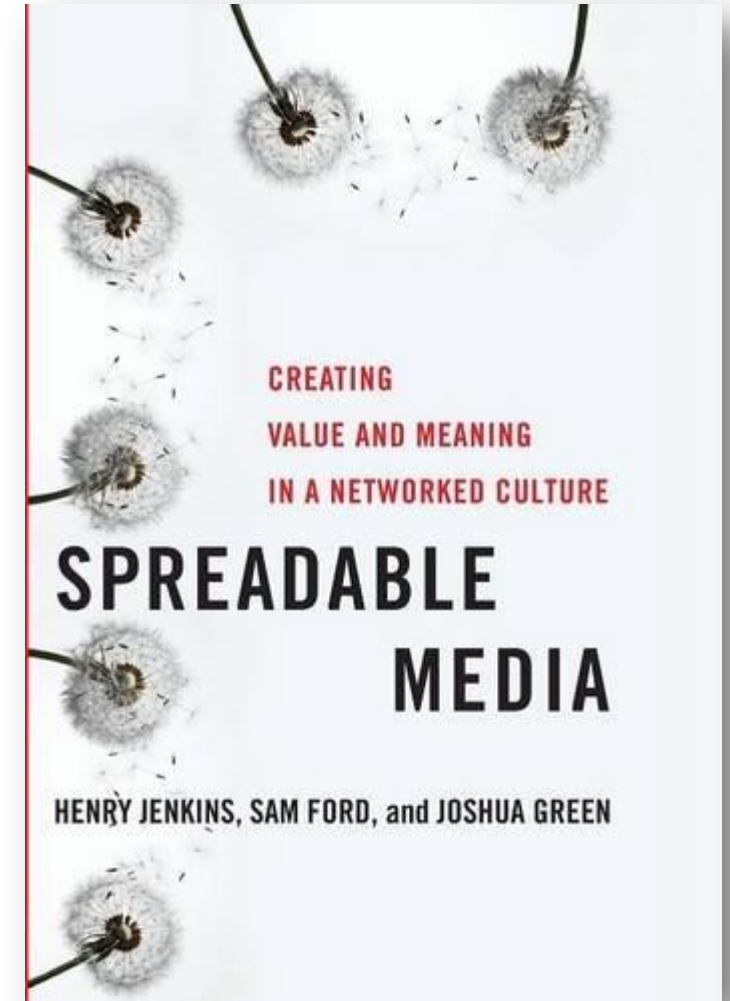
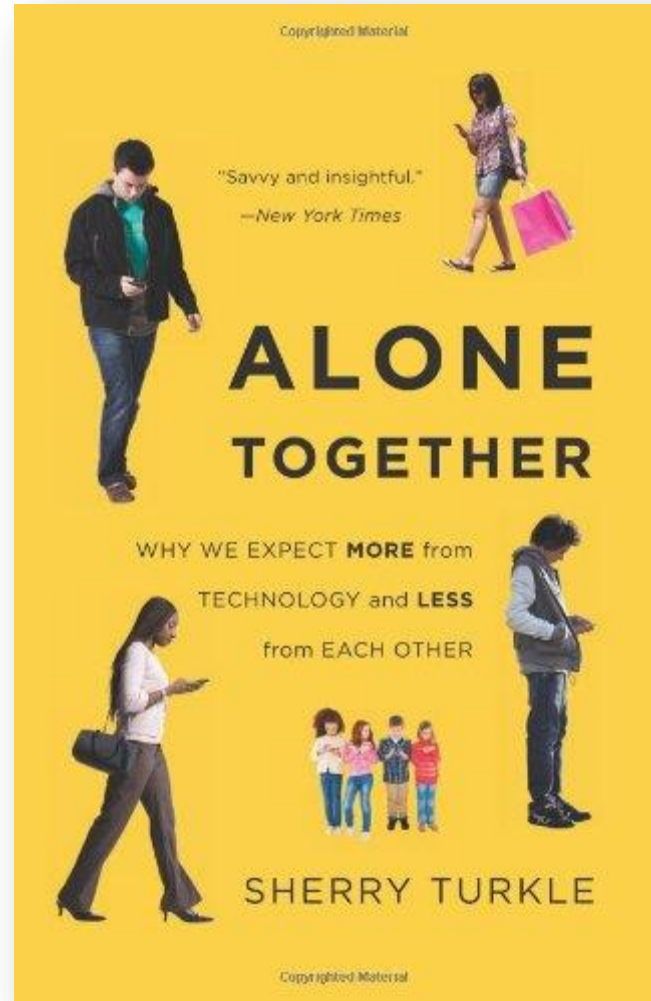
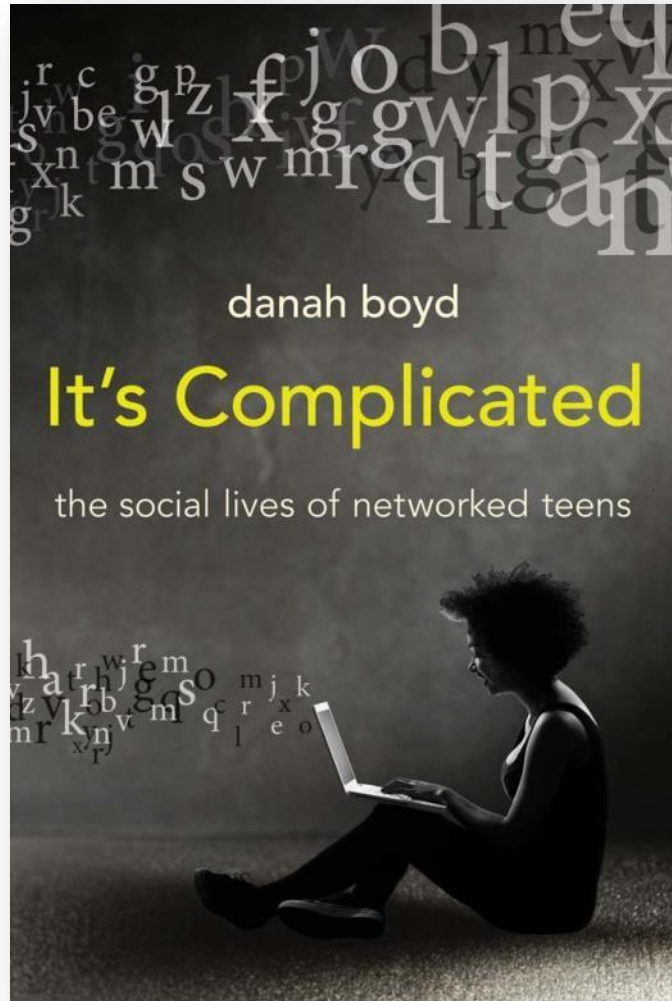


15/03

Matéria aprovada na comissão do Senado

A Comissão de Ciência, Tecnologia, Inovação, Comunicação e Informática do Senado (CCT) aprovou a liberação de produção e uso da Pílula do Câncer. Agora a proposta segue para análise da





Mas seguradoras não são redes sociais



??

=



Redes de pessoas: tipos de informação

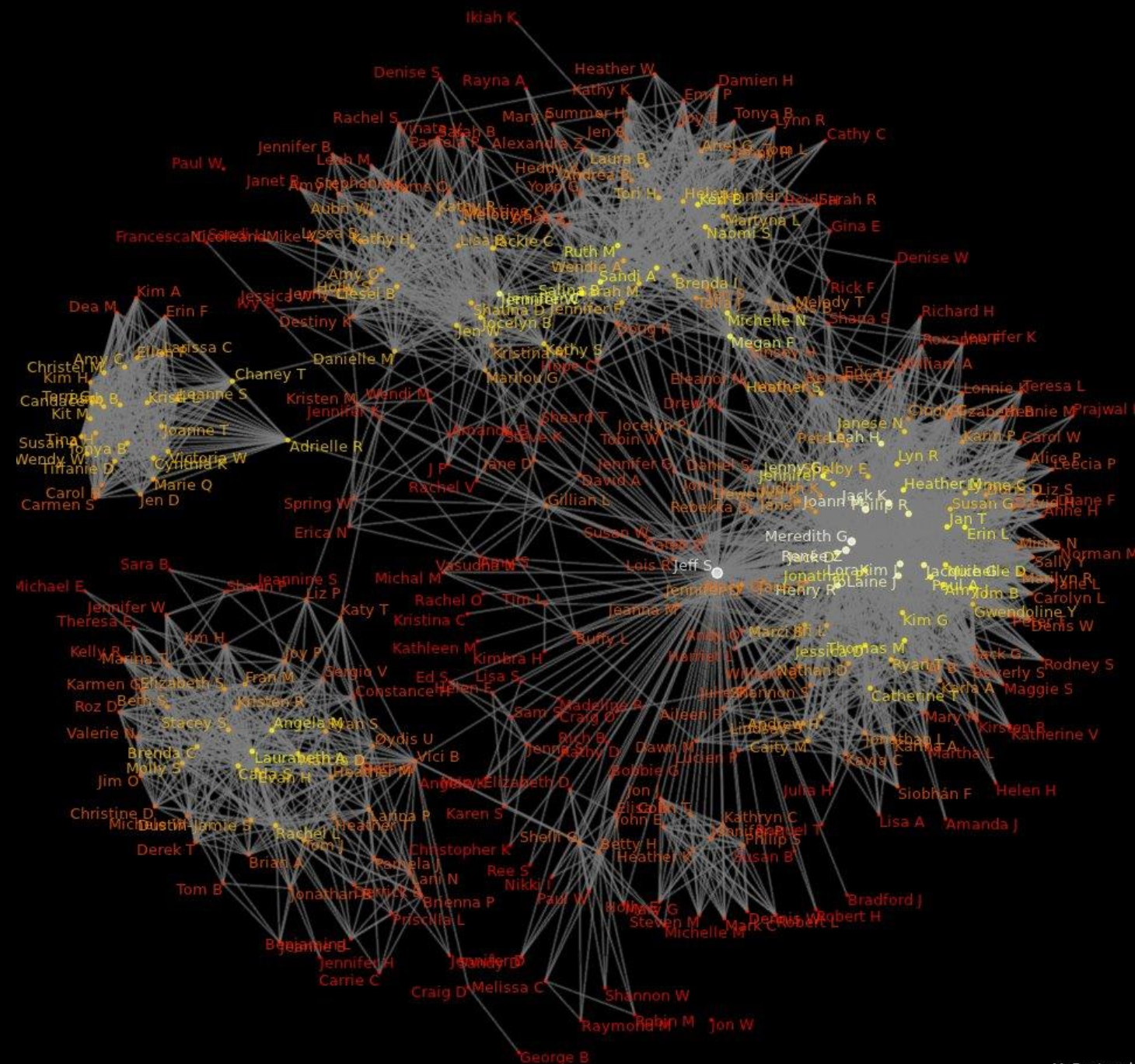
informações individuais

relações entre indivíduos

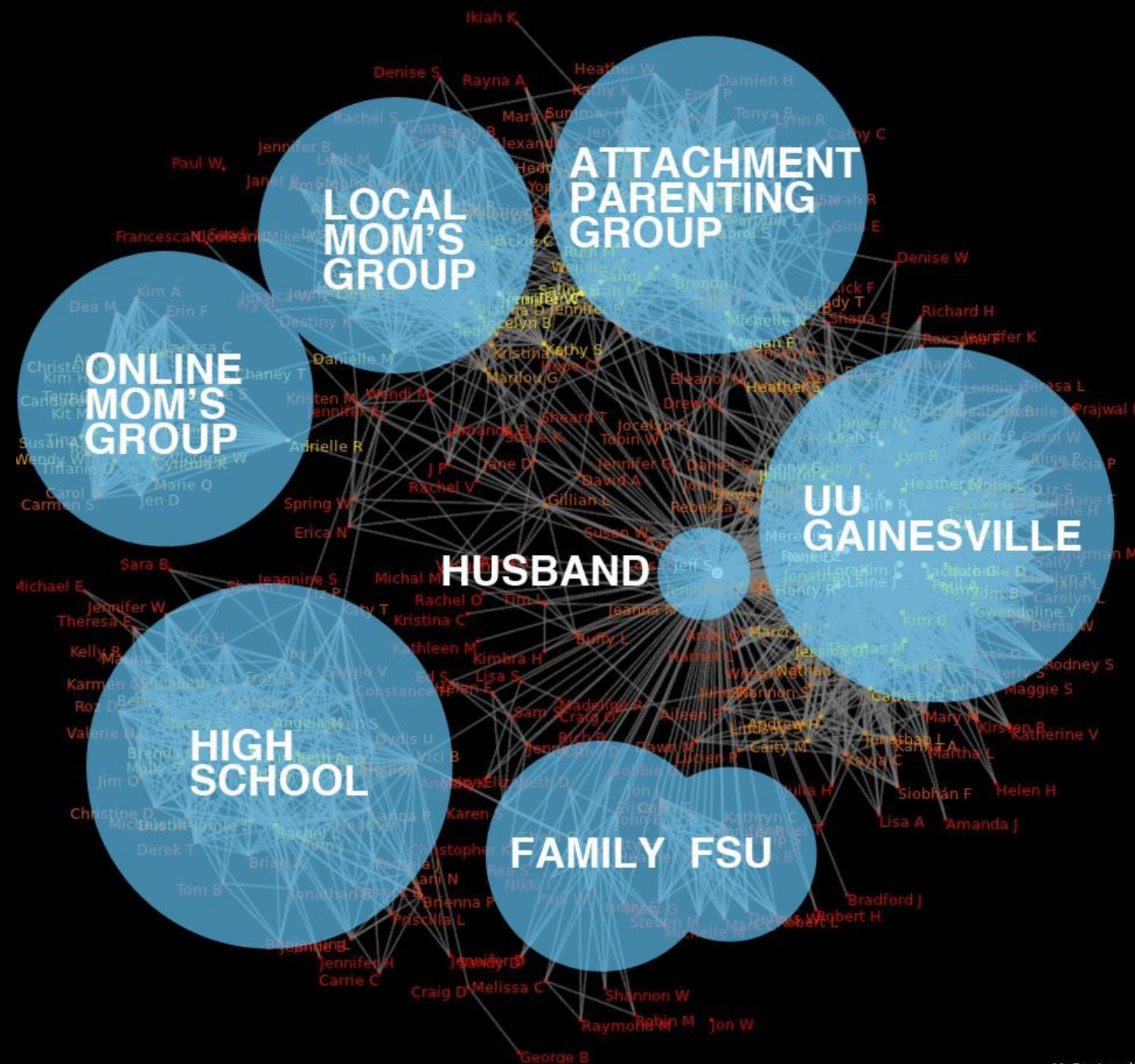
comunicações/ações entre indivíduos



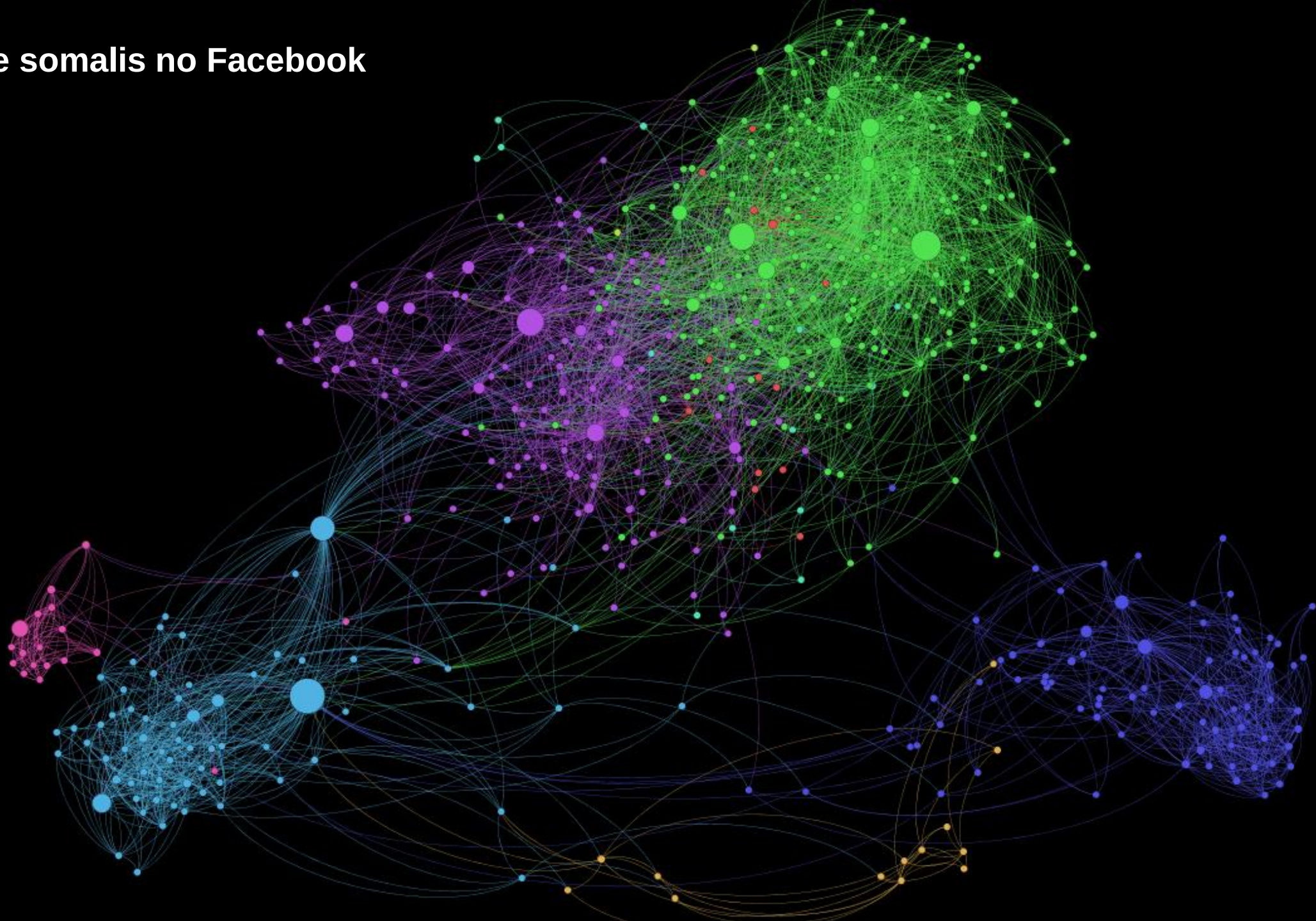
redes de amigos



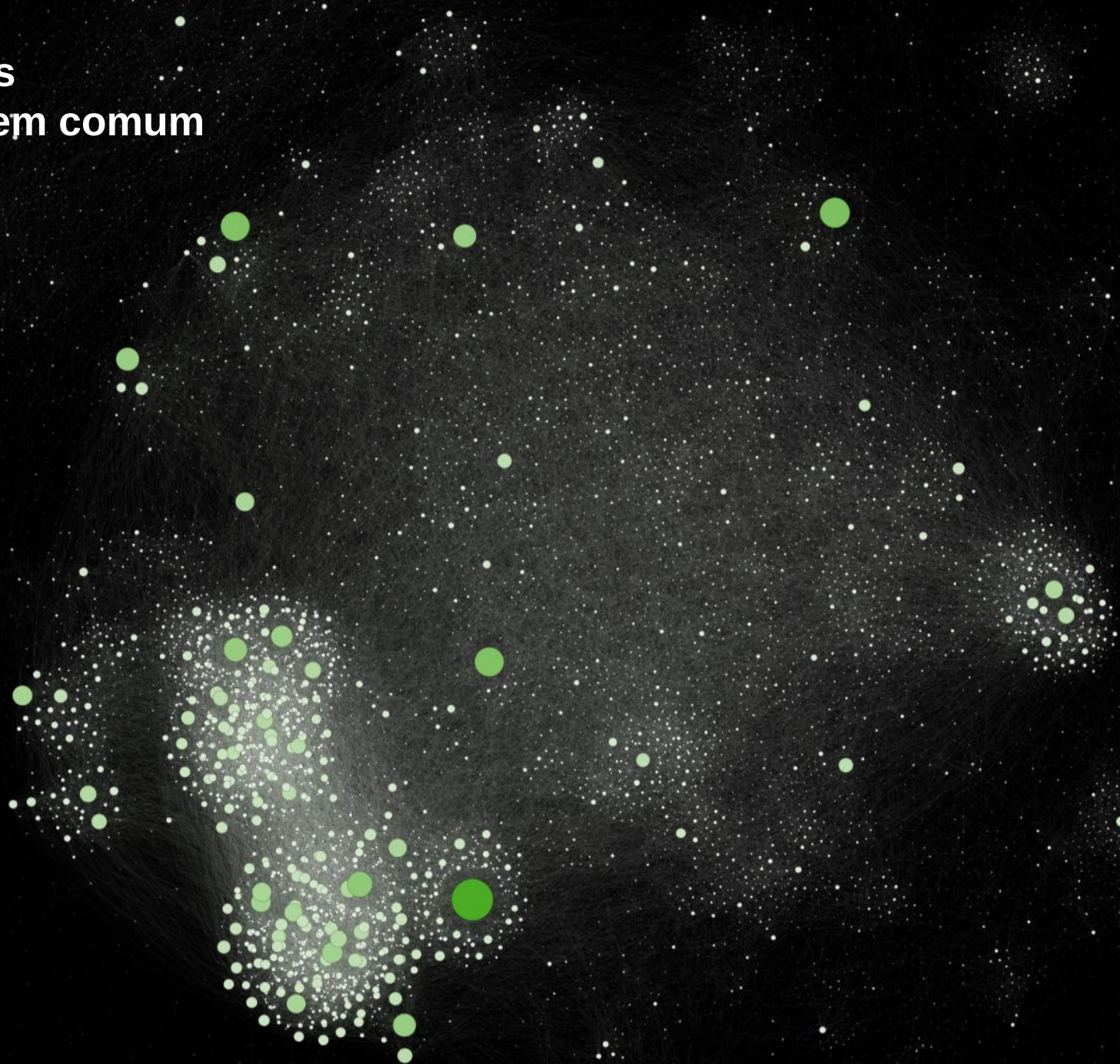
redes de amigos



rede de somalis no Facebook



rede de médicos
com pacientes em comum

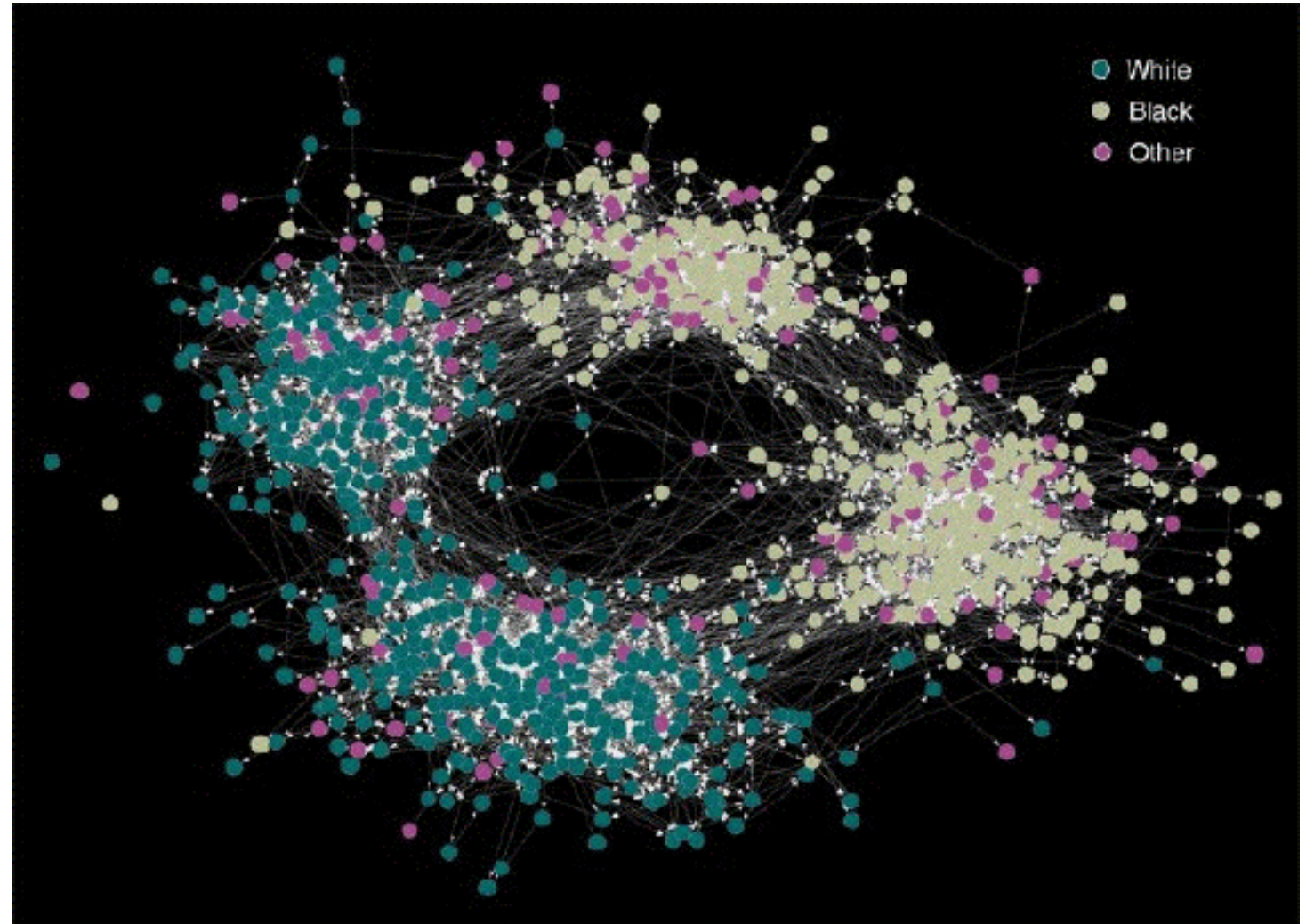




REDES COMPLEXAS

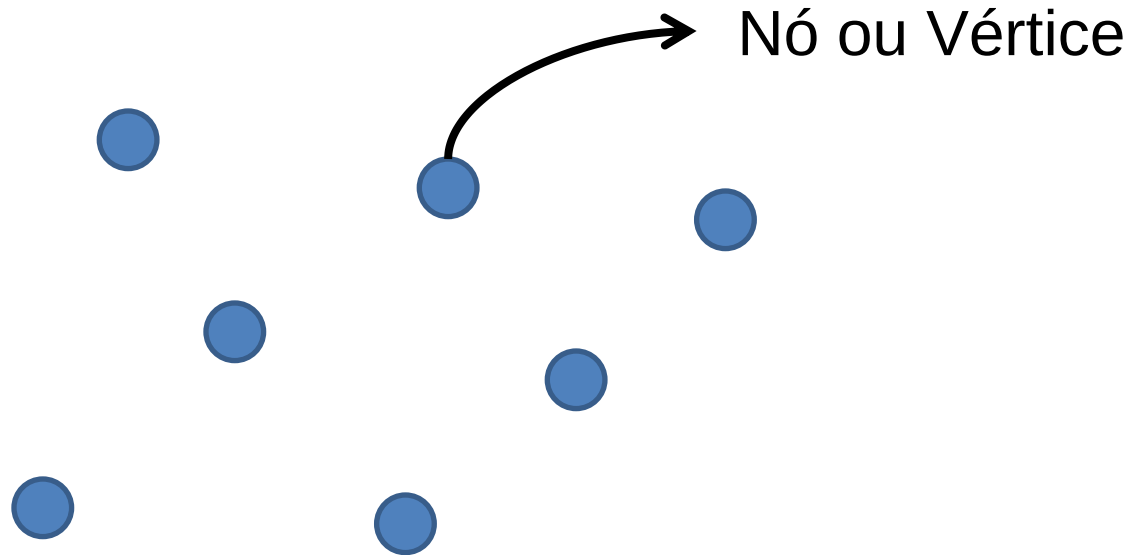
Estrutura de comunidade

- grupos de nós que possuem um alto número de arestas entre si do que com outros grupos de nós
- as pessoas se dividem em grupo naturalmente baseada em interesses, idade, ocupação, ...



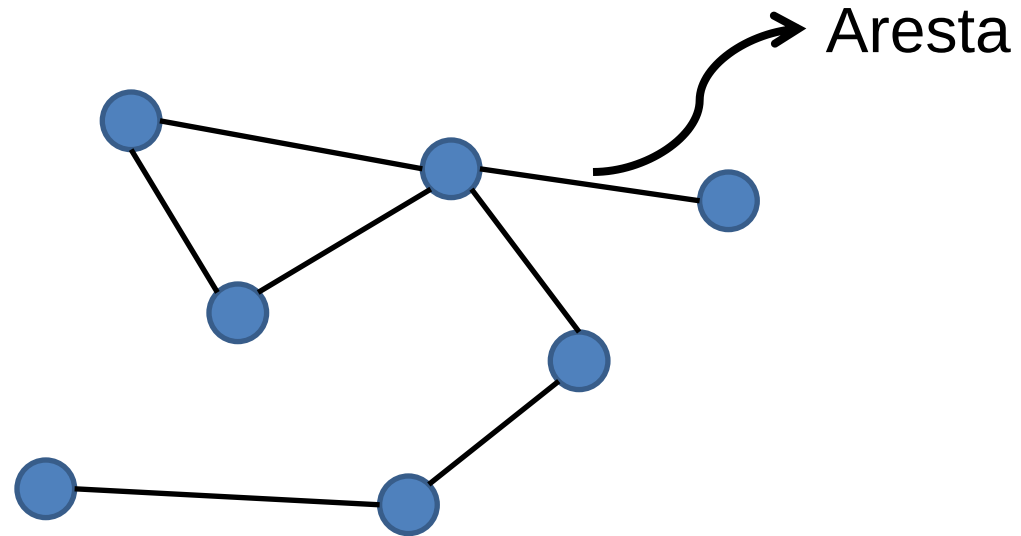
rede de amizade de crianças em uma escola

O que é um grafo?



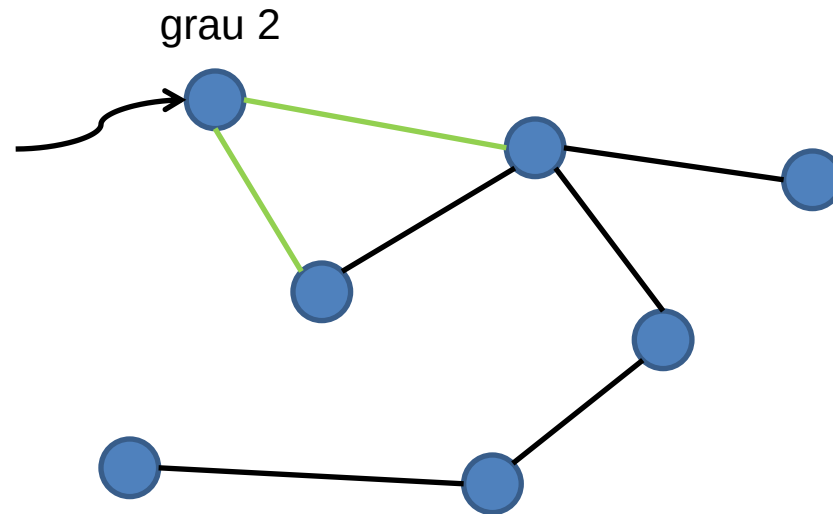
Objeto matemático discreto que contém **nós** e **arestas**.

O que é um grafo?

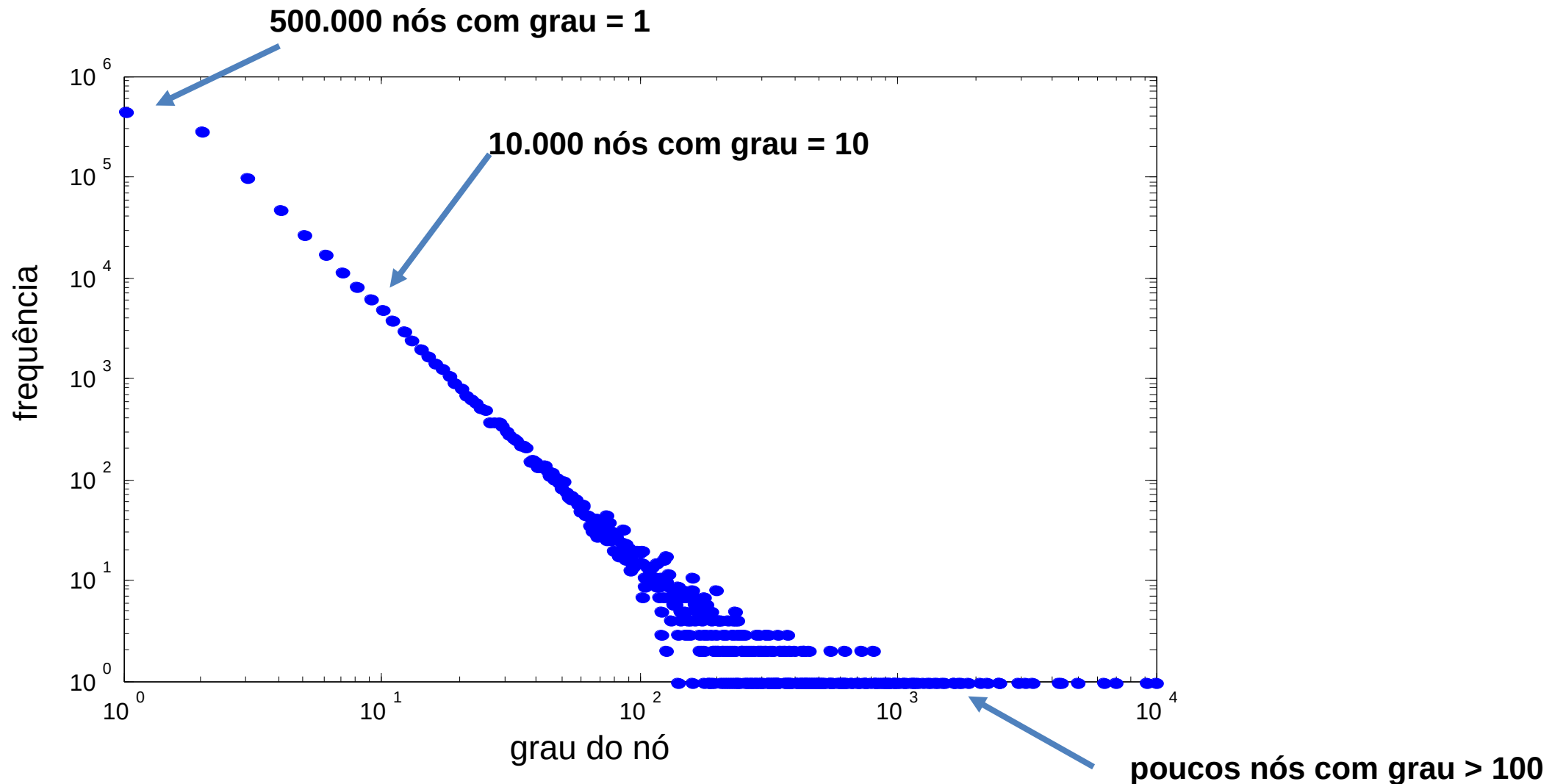


Arestas conectam os nós

Número de arestas incidentes ao vértice é chamado **grau do nó**.



Lei da Potência (Power Law)



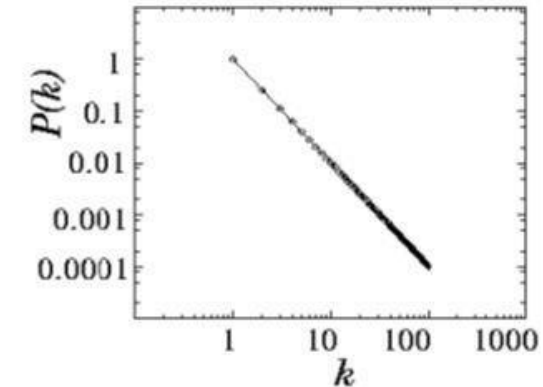
uma lei de potência é uma **reta** em um gráfico na escala **log-log**

Rede complexas são redes (grafos) onde a distribuição dos graus dos nós segue uma lei de potência.

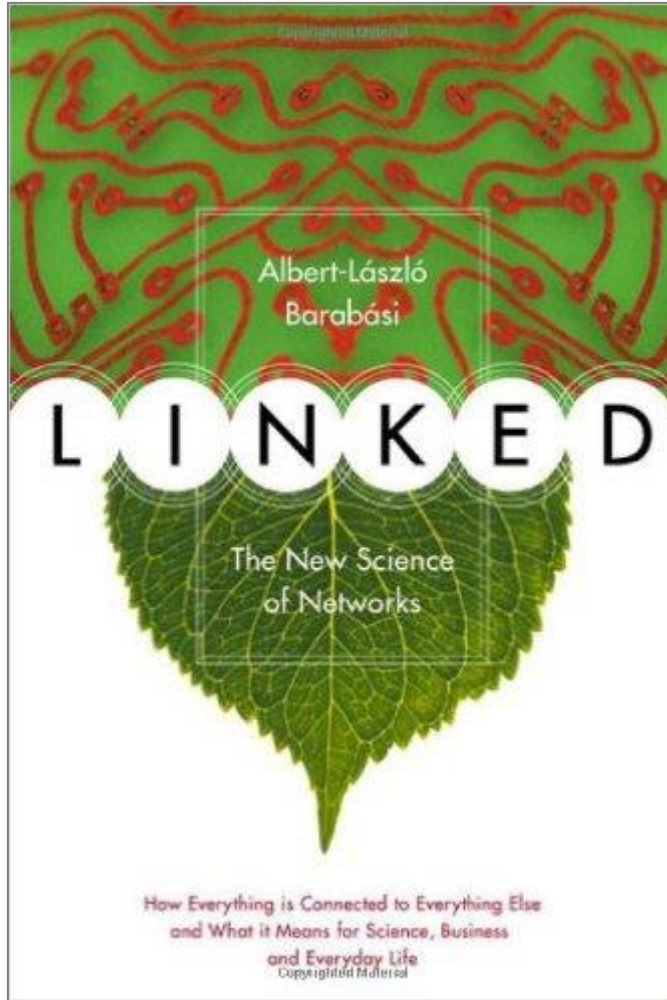
- Possuem estrutura de comunidade.
- Rápida propagação da comunicação.
- Existem membros mais “influentes” (hubs).
- Frágeis em relação aos hubs.

... extremamente estudadas e analisadas.

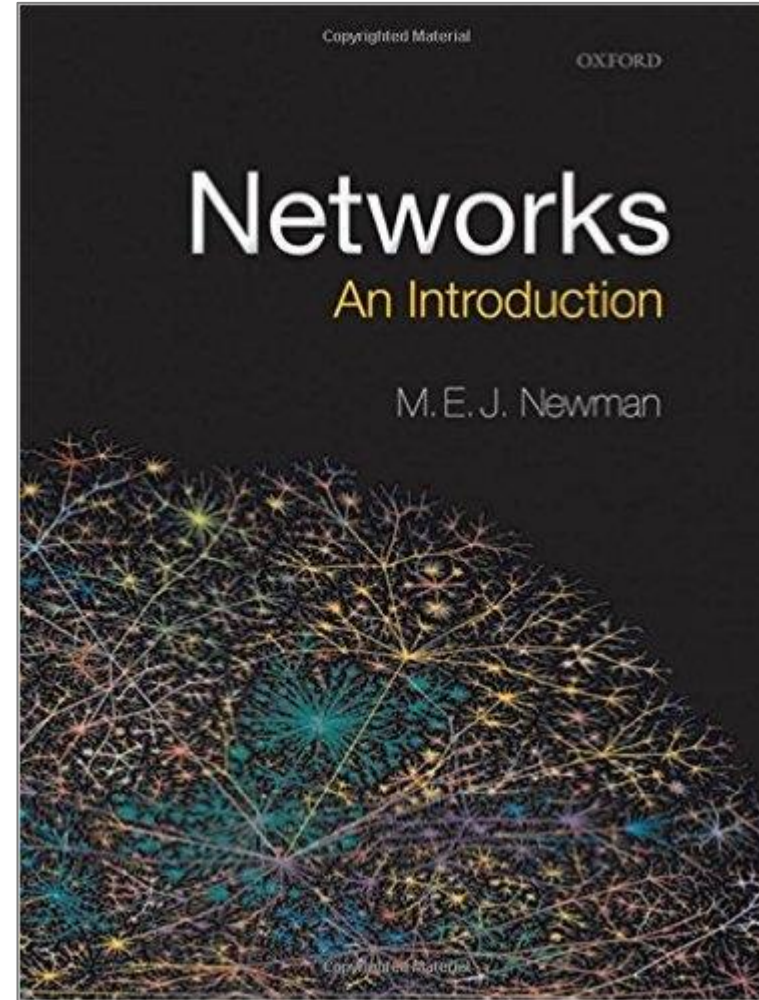
Power law



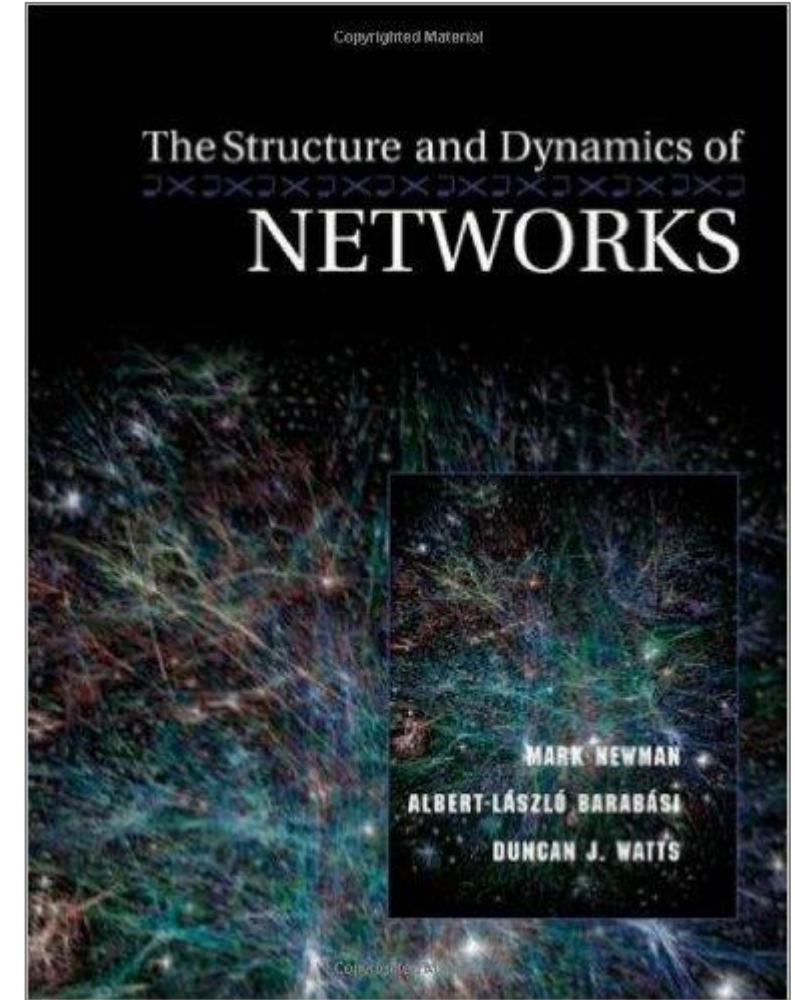
Air traffic network



introdução

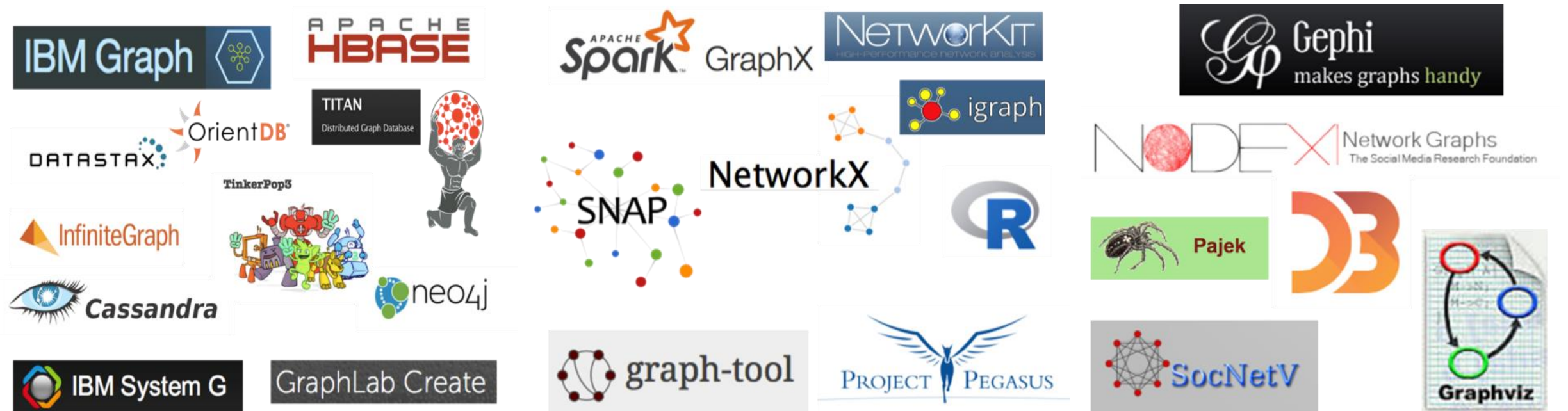


livro-texto



avanzado

GRAPH ANALYTICS



bancos de dados

ferramentas de análise

visualização



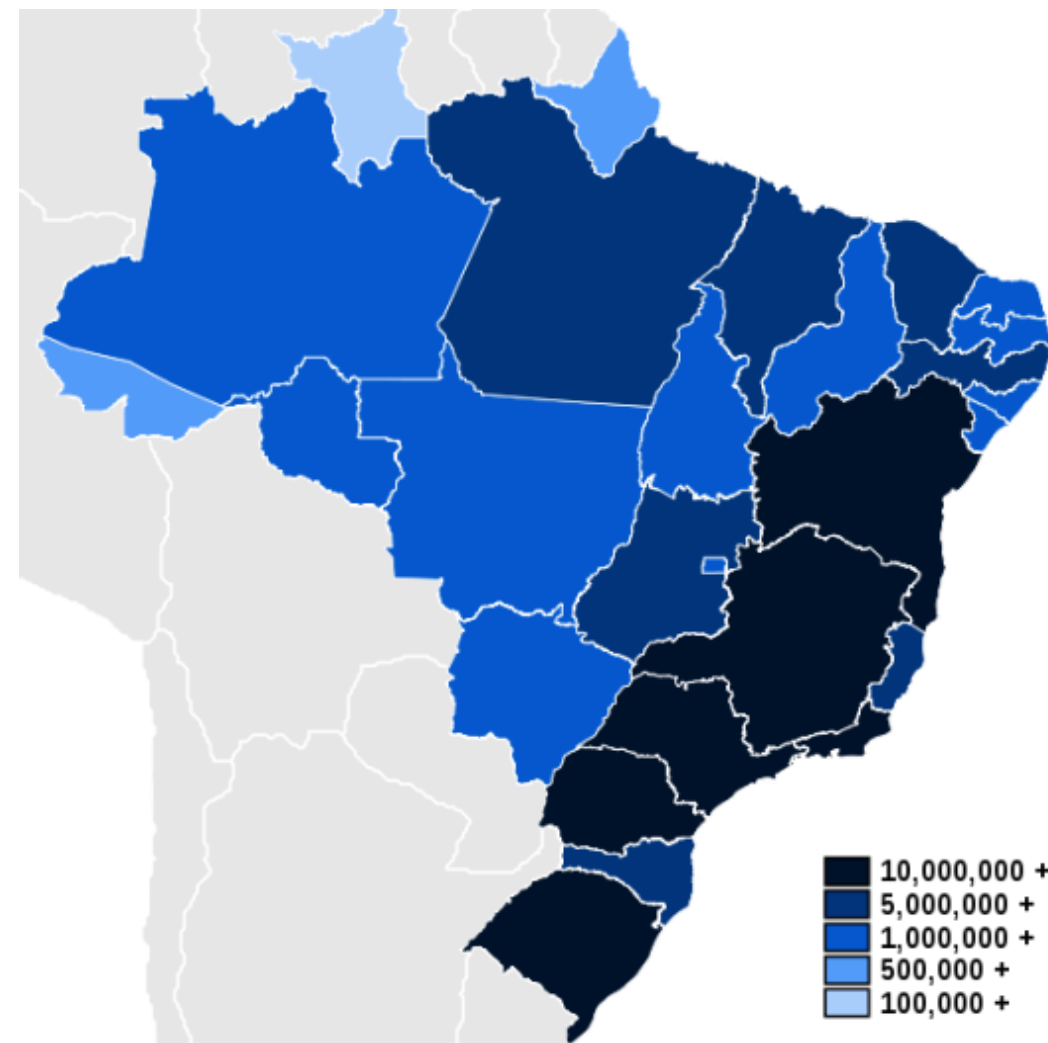
CASE: GRAPH ANALYTICS EM DADOS DE REEMBOLSO DE UMA SEGURADORA NO BRASIL

LUIS MOYANO, ANA APPEL, VAGNER SANTANA, MARCIA ITO

Dados de cobrança/reembolso médico



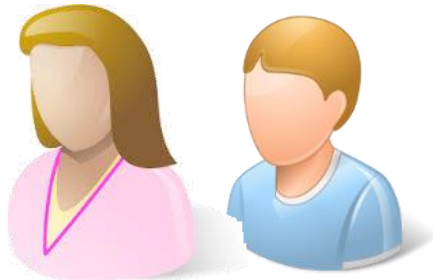
- Dados de transações de seguro saúde
- Uma das maiores seguradoras do Brasil
- Dados envolvendo todo o país
- Intervalo 1 ano e meio (2013-2014)
- 0.6 Tb (compactado)



Stakeholders



médicos



pacientes



seguradora



prestadores



reembolsadores



serviços da área da saúde

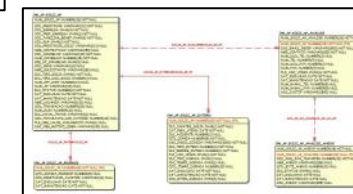
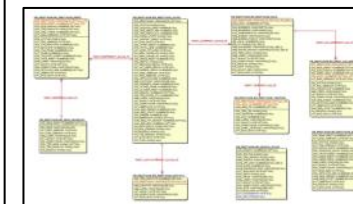
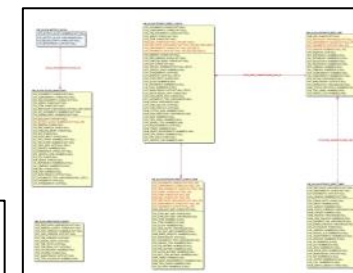
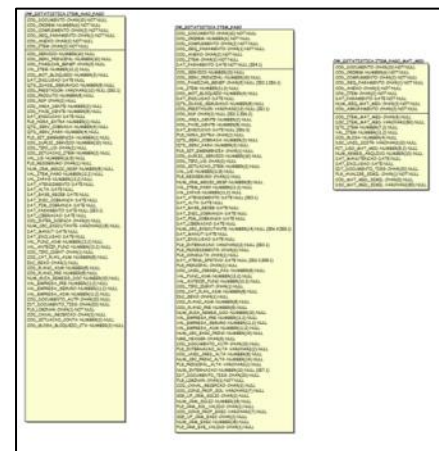
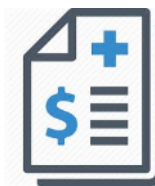
Características dos dados de reembolso



- Cobranças/reembolsos pagos
 - Total: 109M
 - Médicos: 220k (*aprox. metade de todos os médicos do Brasil*)
 - Pacientes: 2.2M
 - Pares Médico-Paciente únicos: 11.6M
- Dados adicionais:
 - Empresas
 - Fornecedores
 - Autorizações ~3M
 - Glosas ~13M
 - Local
 - ...

reembolso

- ID do médico
- ID do paciente
- Timestamp
- Cód. de serviço
- CID
- (80+ linhas)



mais de 40 tabelas,
centenas de campos

- Definir **métricas de qualidade** de médicos em função de suas relações com outros médicos e pacientes.
- Métricas serão combinadas com outras métricas (provenientes de informações individuais e de processos) para classificar, filtrar, convidar e remover médicos do cadastro da seguradora.

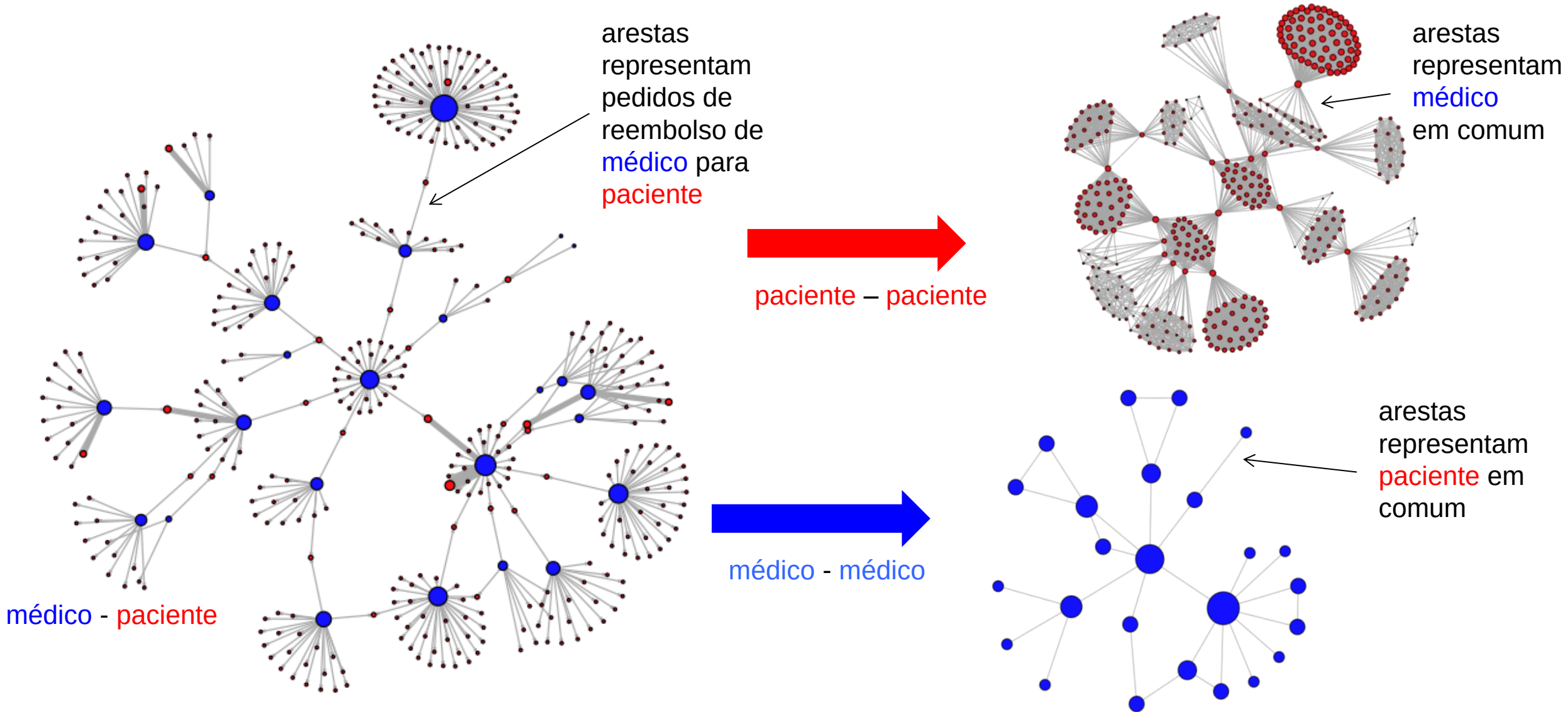


referência mútua	centralidade	fidelização	tempo retorno 2	tempo retorno 7		
35	0.1	3.2	4%	7%	...	...

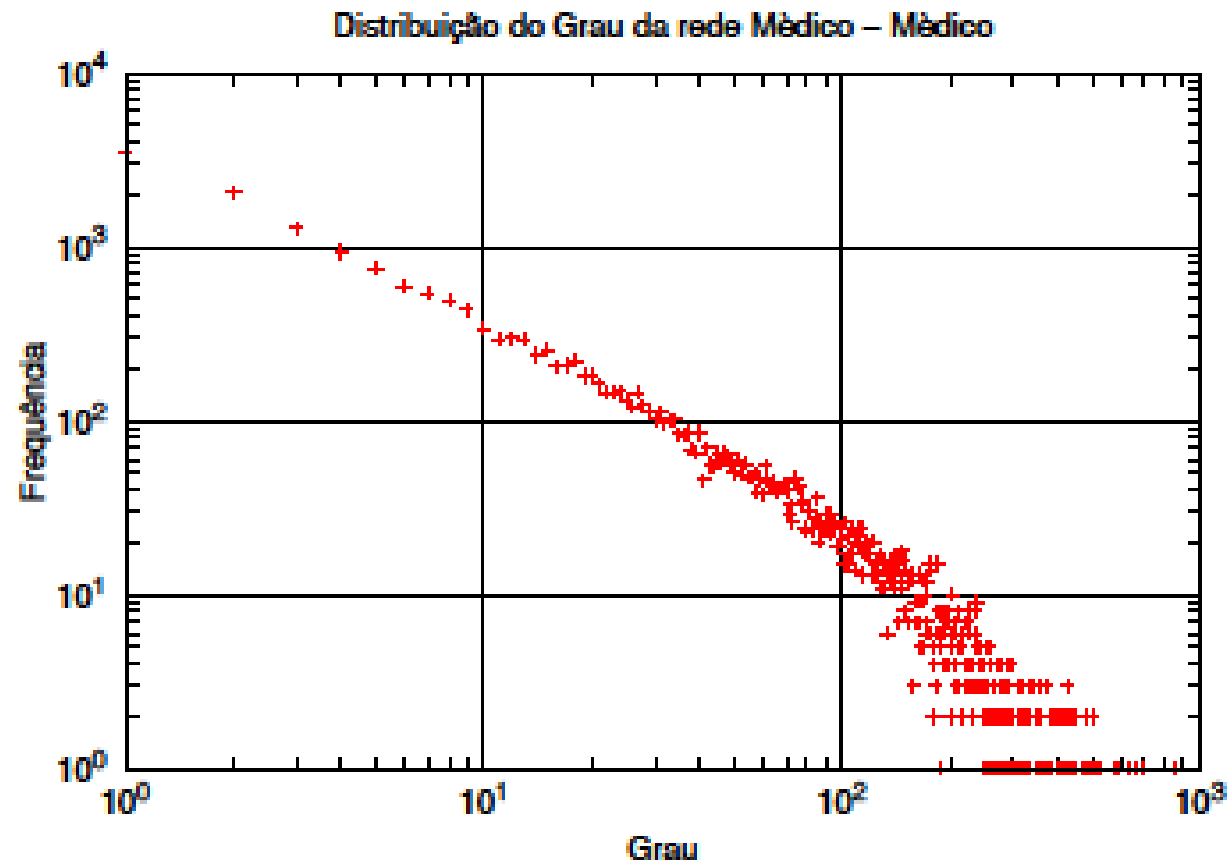


referência mútua	centralidade	fidelização	tempo retorno 2	tempo retorno 7		
17	0.2	5.1	9%	1%	...	...

Exemplos de redes de pacientes/médicos



Distribuição de grau do grafo de médicos - médicos



Métrica 1. Referência Mútua



Objetivo:

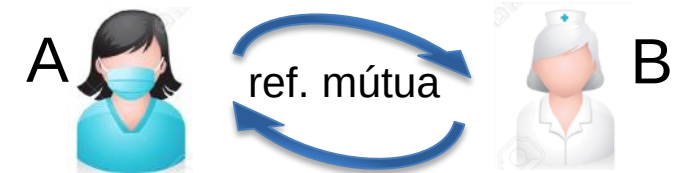
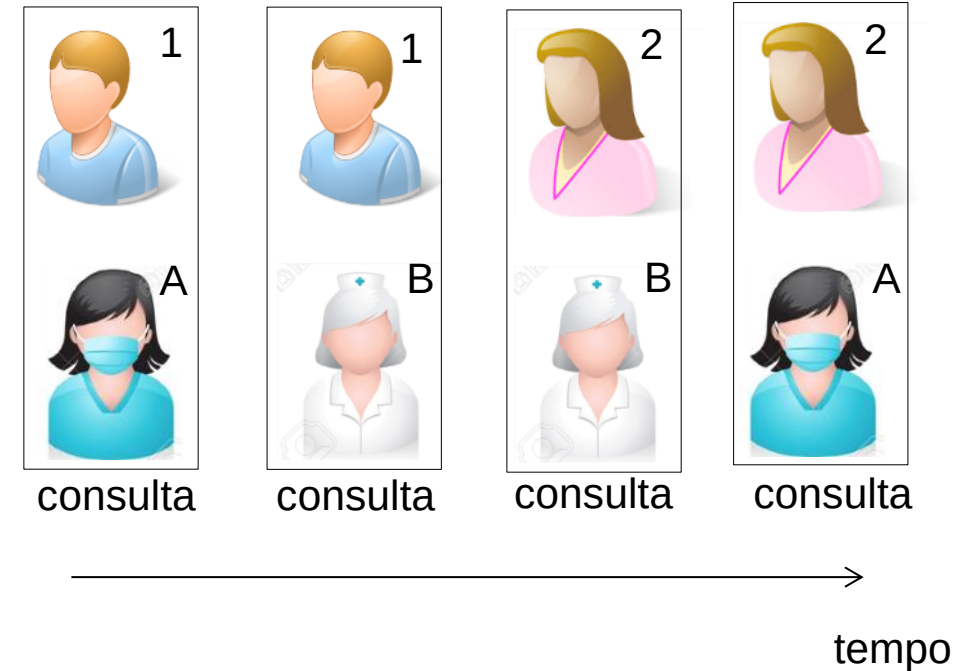
Identificar conexões fortes entre pares de médicos, em particular aquelas além do padrão.

Definição 3.3 (Escore de relação mútua) O escore de relação mútua mede, no intervalo 0 – 1, a relação entre cada par de vértices v_x e v_y em relação à máxima relação mútua identificada em um grafo direcionado G . A máxima relação mútua é representada como:

$$\max(mi(v_x, v_y)) = w(e_{x,y}) + w(e_{y,x}) - |w(e_{x,y}) - w(e_{y,x})|$$

Dessa forma, para cada par de vértices em um grafo direcionado G , o escore de relação mútua é computado da seguinte forma:

$$mis(v_j, v_i) = mis(v_i, v_j) = \frac{mi(v_i, v_j)}{\max(mi(v_x, v_y))}$$

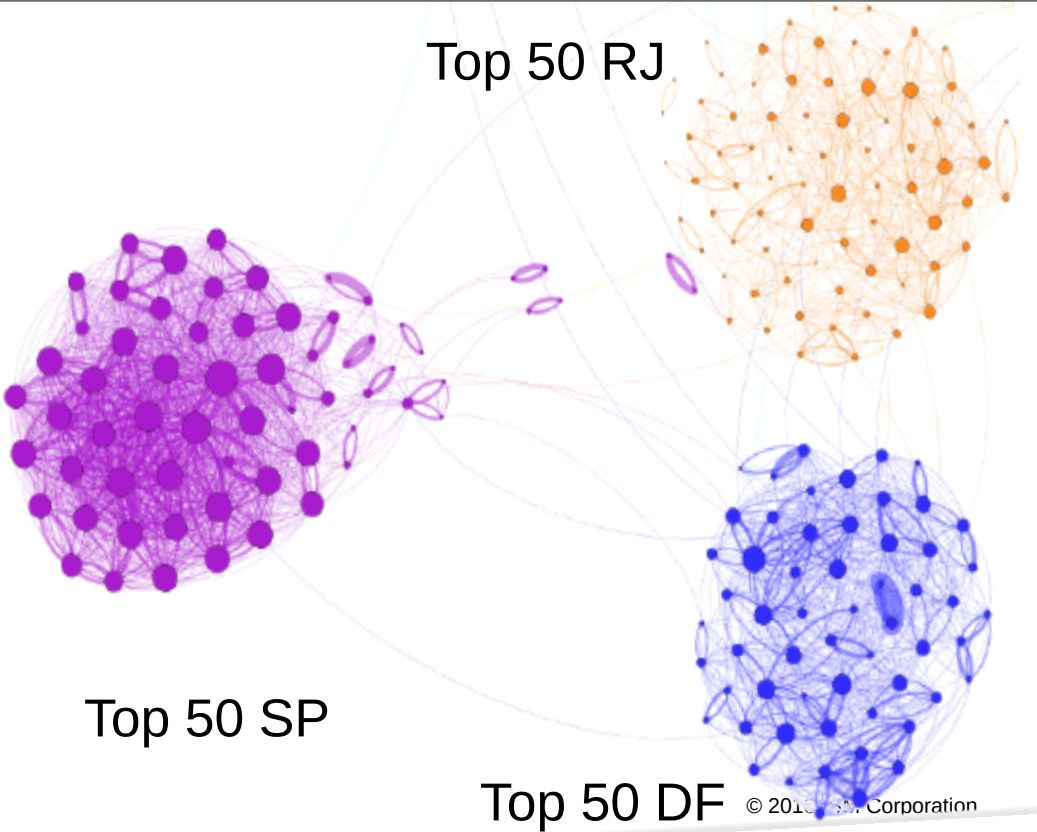


Resultados

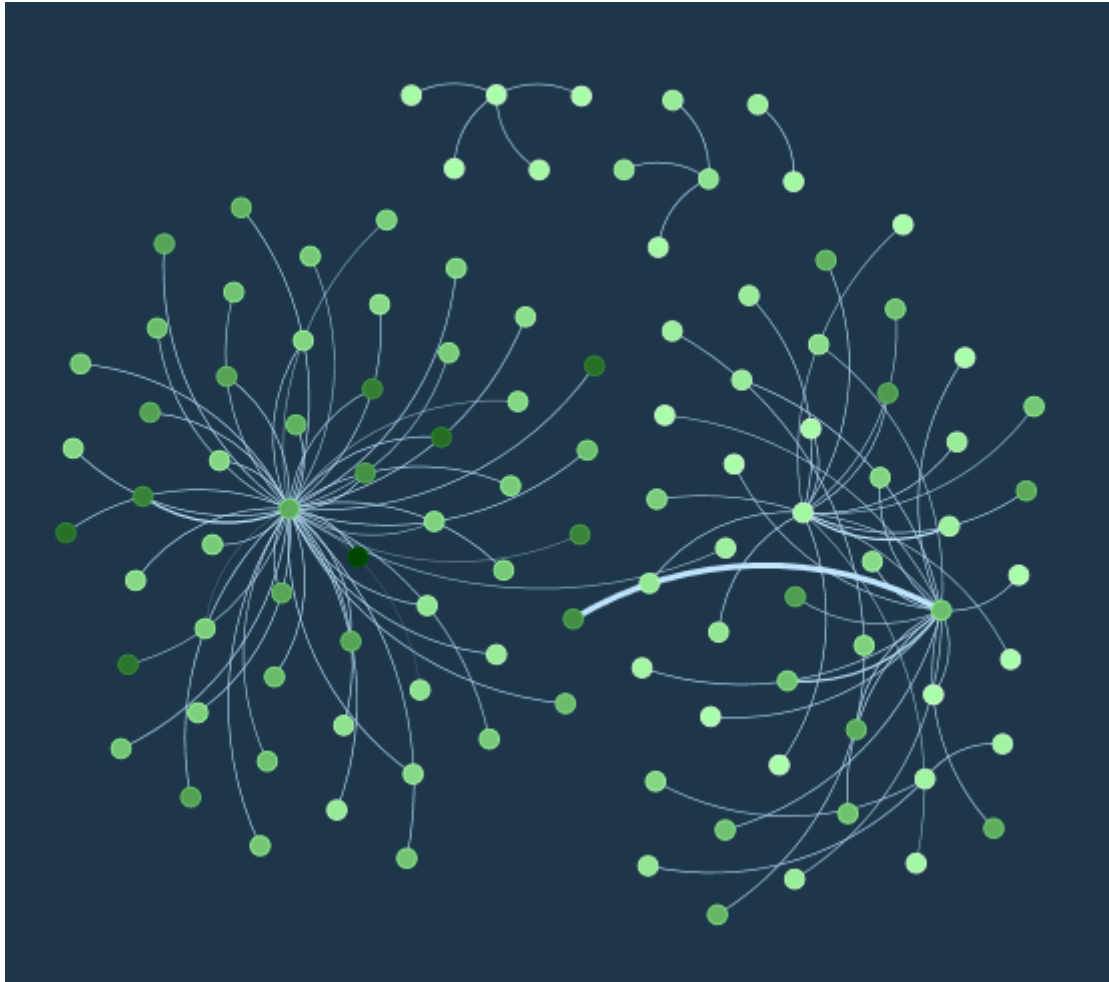
- Especialidades que mais se referenciam
 - Oftalmologia para oftalmologia
 - Ginecologia e obstetrícia
- DF conta consultas em intervalos mais regulares; PE é a que tem menor regularidade
- No relacionamento entre **M_{DF}010** (medicina intensiva) e **M_{DF}009** (especialidade não informada), com 267 consultas, a mediana do número de dias entre as consultas é 0
- O par com maior relacionamento mútuo vem de MS;
 - 205 de **M_{MS}028** para **M_{MS}027**
 - 196 de **M_{MS}027** para **M_{MS}028**

Métrica de referência mútua

	Médico A	Médico B	rm	Ran k
	M_{MS}028	M_{MS}027	1	1
	M_{SP}145	M_{SP}144	0.316	10



Métrica 1. Referência Mútua

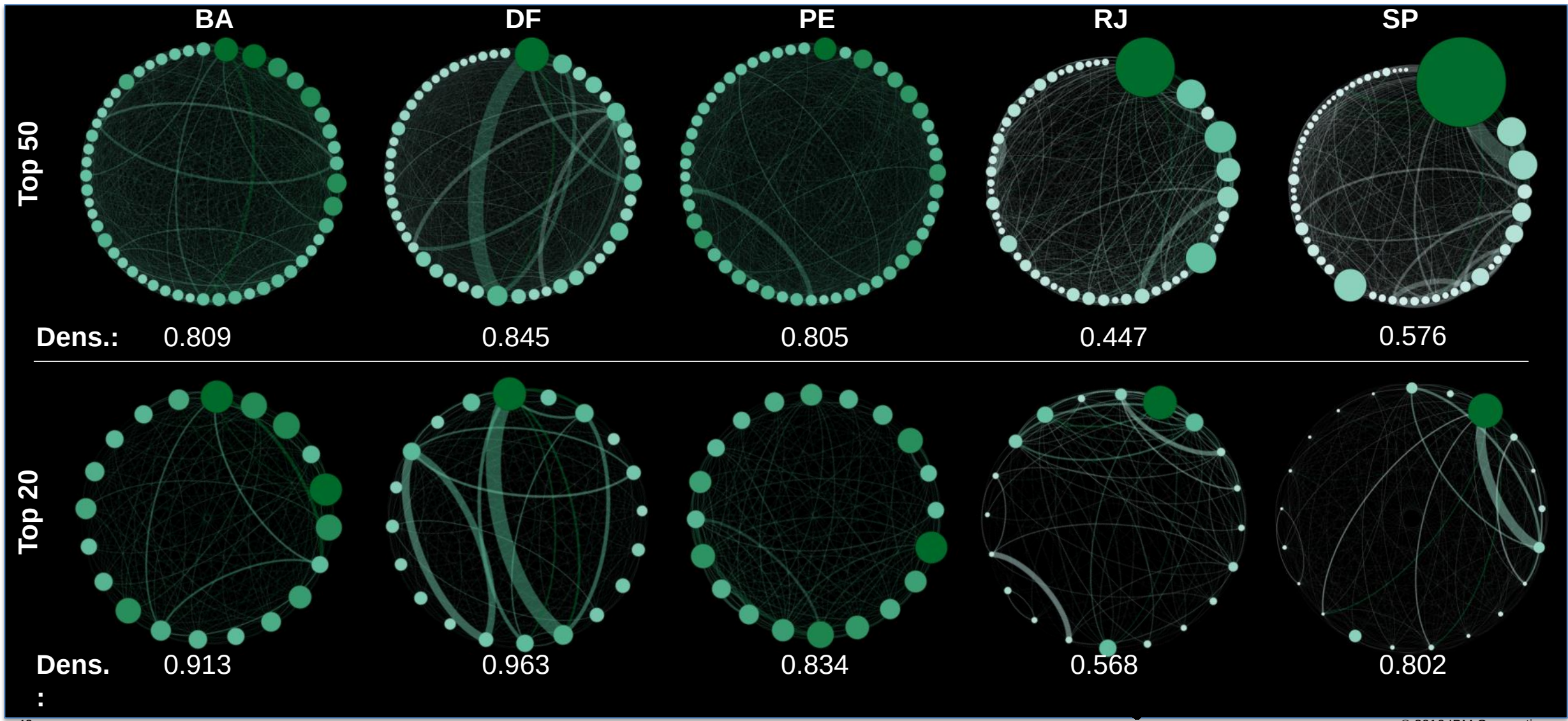


alergia e imunologia



oftalmologia

Métrica 1. Referência Mútua



Possíveis usos da métrica:

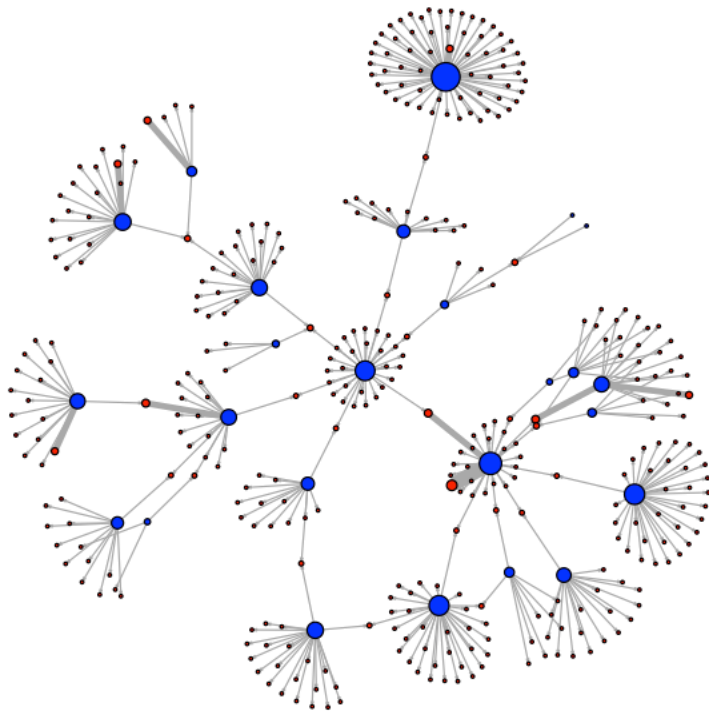
1. **Identificar** de forma recorrente os **pares de médicos** com alta probabilidade de referência mútua para melhor entendimento da sua atividade.
2. Manter e **melhorar relacionamento/experiência** com médicos parceiros que trabalham em conjunto. Fazer campanhas de relacionamento e **propor parcerias** para médicos que trabalham em conjunto.
3. Identificar quando o relacionamento entre médicos ocorre por referência/contra-referência, **investigando os médicos** com valor alto da métrica com maior detalhe .
4. Identificar relações que podem indicar **fraude**.
5. Identificar pares de CRM de estados / cidades diferentes para **entender o fluxo de pacientes**, disponibilidade de serviços, etc.

Métrica 2. Importância dos Médicos na Rede



Objetivo

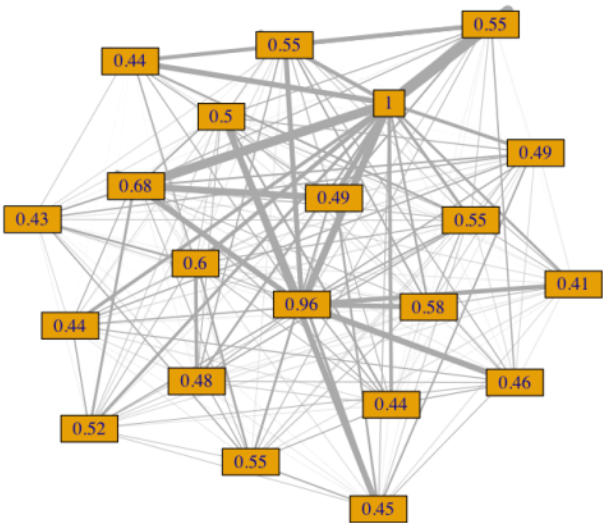
Identificar o papel dos médicos na rede pela sua importância em relação aos outros médicos, especialmente se eles são especialistas procurados.



- **Grau:** Neste caso a importância relativa é proporcional ao número total de segurados compartilhados com outros médicos. Quanto maior o número de pacientes do médico compartilhados com outros médicos, mais central o médico é considerado.
- **Eigenvalue:** Nesta métrica, o médico é considerado como sendo mais destacado quando mais pacientes ele compartilhar com outros médicos também considerados como importantes pela métrica. Assim, um médico que tem um alto número de pacientes em comum com outros médicos com valores baixos da métrica terá também um valor baixo e portanto não será considerado muito influente.
- **Betweenness:** Levando em consideração que médicos podem influenciar médicos próximos, por exemplo, ao indicá-los para seus pacientes, é possível definir uma métrica que capture este efeito de proximidade. Com esta métrica, o médico é considerado como sendo mais importante quanto mais se conecta (potencialmente) com outros médicos.
- **Closeness** Mesmo sem ter conhecimento mútuo, os médicos podem estar relacionados pela quantidade de outros médicos conhecidos entre si. Assim, podemos definir um grau de separação entre os médicos pelo número necessário de outros médicos para conectá-los. Desta maneira, grau 1 de separação corresponde a médicos que têm pacientes em comum. Grau 2 são aqueles médicos que se conectam entre si através de outro médico (amigos de amigos), grau 3 através de 2 médicos e assim sucessivamente. Usando este conceito de grau de separação, entendemos como importante um médico que tem um grau baixo de distância com os outros médicos.

Resultados

- Definimos 4 métricas e calculamos as métricas para a base toda, e para 5 quartis.
- As **métricas são estáveis** entre si. A evolução por quartis do topo dos rankings de médicos (métrica *eigenvalue*) varia pouco demonstrando estabilidade da métrica.
- No **subconjunto central de médicos** com maior valor da métrica, observamos um nível muito forte de compartilhamento de pacientes. O tamanho desse subconjunto varia entre dezenas e centenas de médicos.



Physician	2013			2014		Specialty
	Q2	Q3	Q4	Q1	Q2	
M _{SP} 153	1	1	2	1	1	Cardiologist
M _{SP} 154	2	3	1	2	2	
M _{SP} 164	3	4	7	4	5	
M _{SP} 142	4	5	3	3	3	
M _{SP} 242	5	10	10	14	12	Neurologist
M _{SP} 243	6	8	8	6	8	Trauma and orthopedic
M _{SP} 140	7	2	4	5	4	
M _{SP} 148	8	7	5	13	11	Cardiologist
M _{SP} 149	9	12	14	12	16	Hematology
M _{SP} 244	10	13	12	15	10	
M _{SP} 246	12	6	11	7	15	
M _{SP} 139	13	-	17	18	7	Cardiologist
M _{SP} 247	14	20	19	-	9	
M _{SP} 248	15	9	6	9	18	
M _{SP} 143	19	-	-	10	13	General and vascular surgery
M _{SP} 252	-	14	9	8	6	

Possíveis usos das métricas:

1. Cálculo periódico da métrica, tanto global quanto localmente e por especialidade, para **entender melhor o papel do médico** em relação aos seus pacientes e outros médicos.
2. Uso da métrica em conjunto com outros dados e meta-dados para entender como se formam **blocos de profissionais**, onde seria conveniente o relacionamento coordenado.
3. Uso da métrica para entender o papel dos médicos dentro desses blocos. Identificação de **médicos centrais/periféricos**, médicos que são essenciais para manter o grupo.
4. Uso da métrica **em conjunto com outras informações** de reconhecimento profissional.
5. Estudo mais detalhado de **outras estruturas de médicos** dentro da rede credenciada (subconjuntos com a mesma especialidade, com diferente especialidade, etc.).

Métrica 3. Fidelização Médico-Segurado



Objetivo

Identificar médicos que de maneira sistemática tem pacientes recorrentes no tempo, sugerindo capacidade de fidelização do segurado.

Weight w_{ij} represents number of visits of patient i to dr. j

Strength s_i : sum of weights attached to links belonging to a node (i.e., all visits from i)

$$s_i = \sum_j w_{ij}$$

Relative weight $r_w(ij)$: fraction of weight ij over total

$$r_w(ij) = w_{ij}/s_i = \frac{w_{ij}}{\sum_j w_{ij}}$$



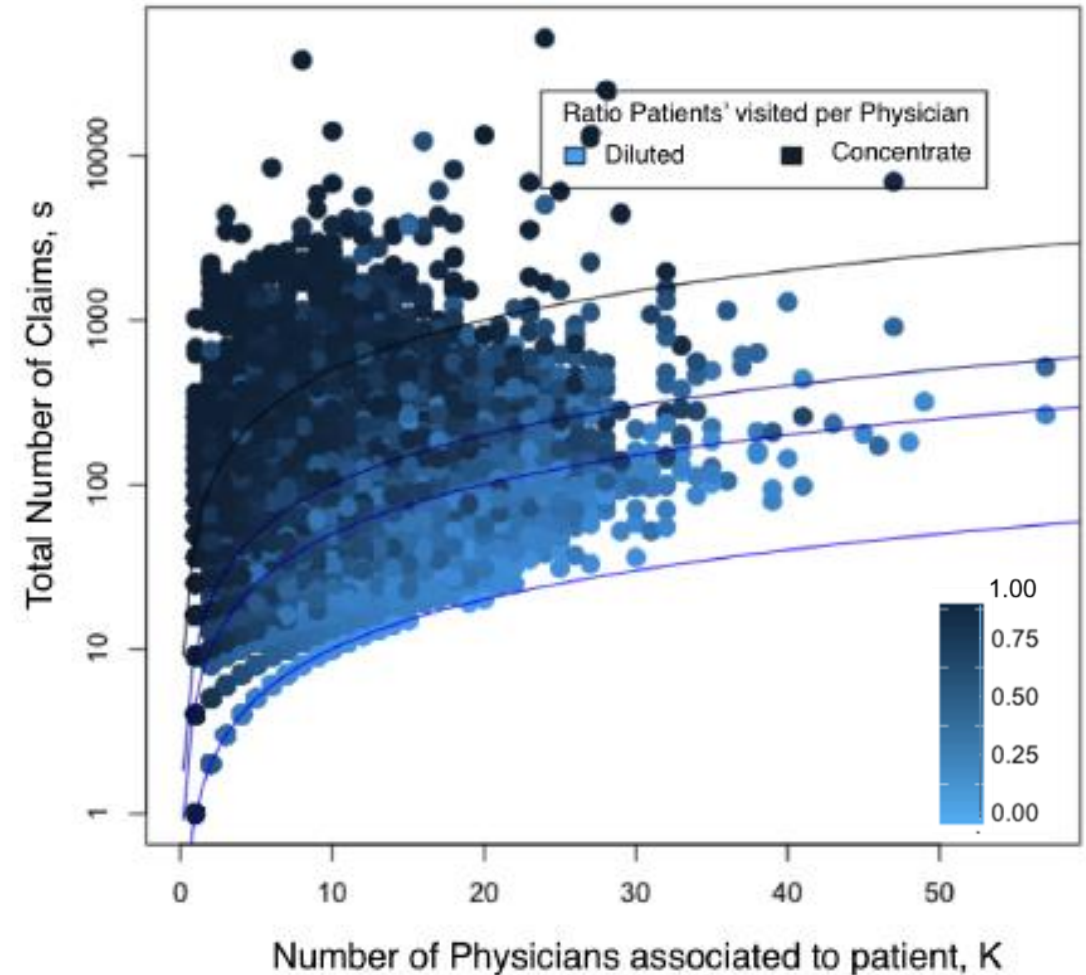
$$s = k$$
$$r_w = 1$$



$$s > k,$$
$$r_w \approx 1$$



$$s > k,$$
$$r_w \approx 1/k$$

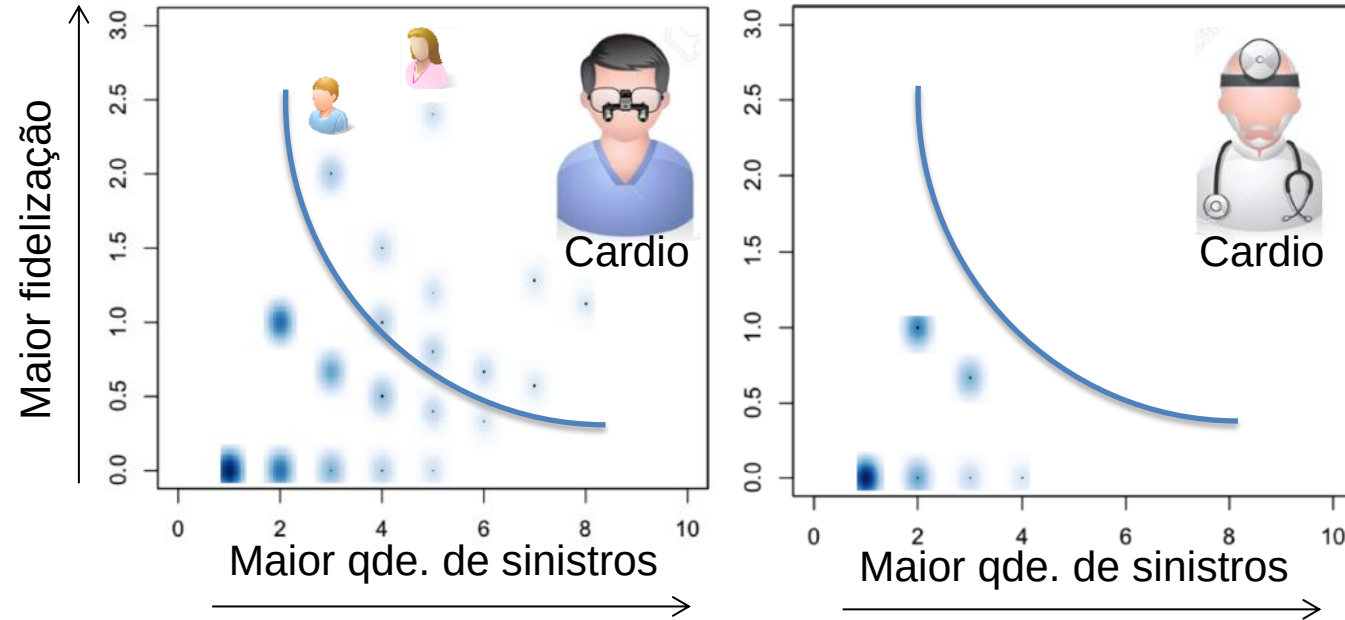


Métrica 3. Fidelização Médico-Segurado



Resultados

- Calculamos **métrica de fidelização** para a base completa
- Muitos médicos mantêm **valores altos sistematicamente**.
 - **M_{BA}261** é número 1 ou 2 durante 5 quartis.
 - Renovação de 4 médicos em média no top 10 em relação ao quartil anterior.
- **Top 5 especialidades** entre médicos de maior fidelização ($mf > 0.5$)
 - Ortopedia e traumatologia (5 no top 10 por freq.)
 - Oftalmologia (3)
 - Ginecologia e obstetrícia(2)
 - Pediatria (1)
 - Otorrinolaringologia (**fora do top 10 por freq.**)



métrica de fidelização



Médico	<i>mf</i>	RANK
M_{SP} 139	1.54	175
M_{SP} 261	1.18	432

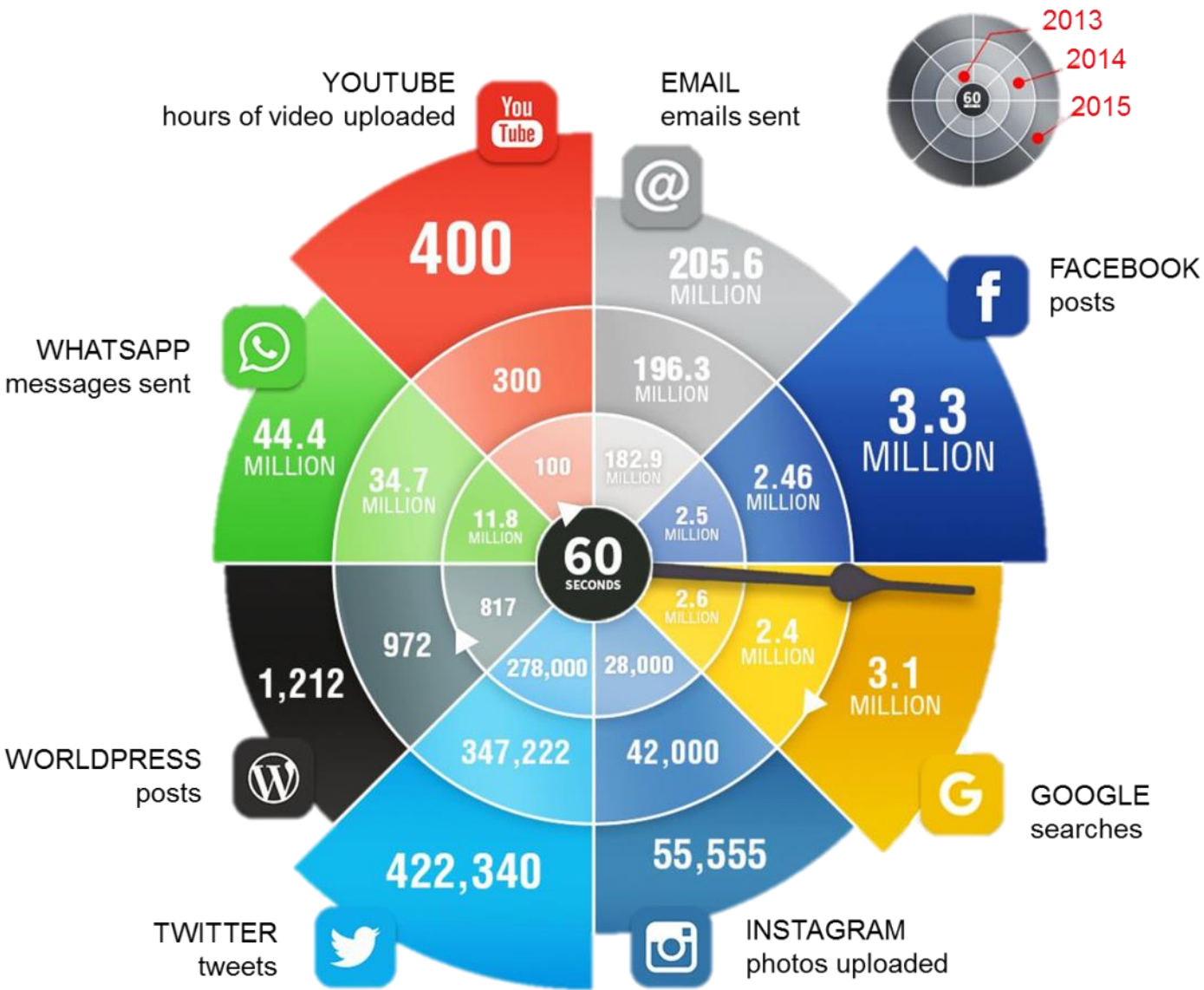
Possíveis usos das métricas:

1. Identificar médicos que tem **maior capacidade de fidelização** para estudo mais detalhado.
2. Desenhar ações de comunicação e **campanhas com os médicos** com maior probabilidade de ter capacidade de fidelização.
3. Identificar **médicos referências** ou que possuem preferências por grupos de pacientes com determinada doença. Esses são médicos que, se associados à seguradora, poderão representar um diferencial no mercado.
4. Colocar médicos com maior valor da métrica de fidelização no **topo da lista de indicação** de médicos do callcenter e do site .

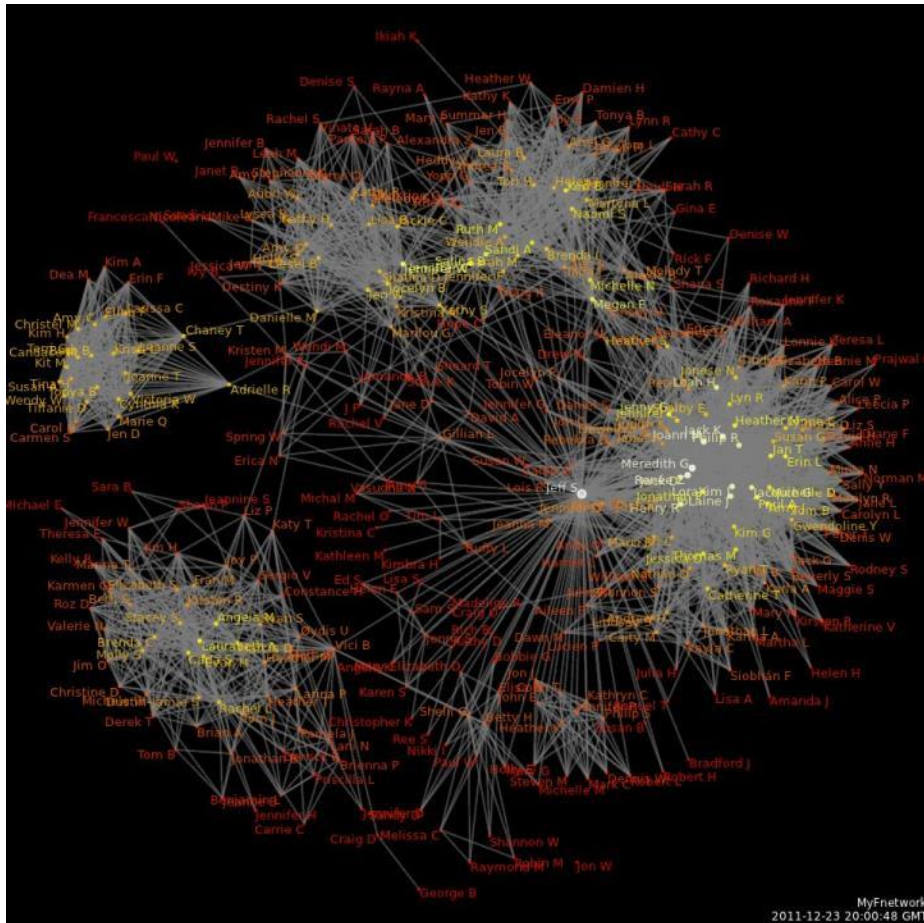


RESUMO

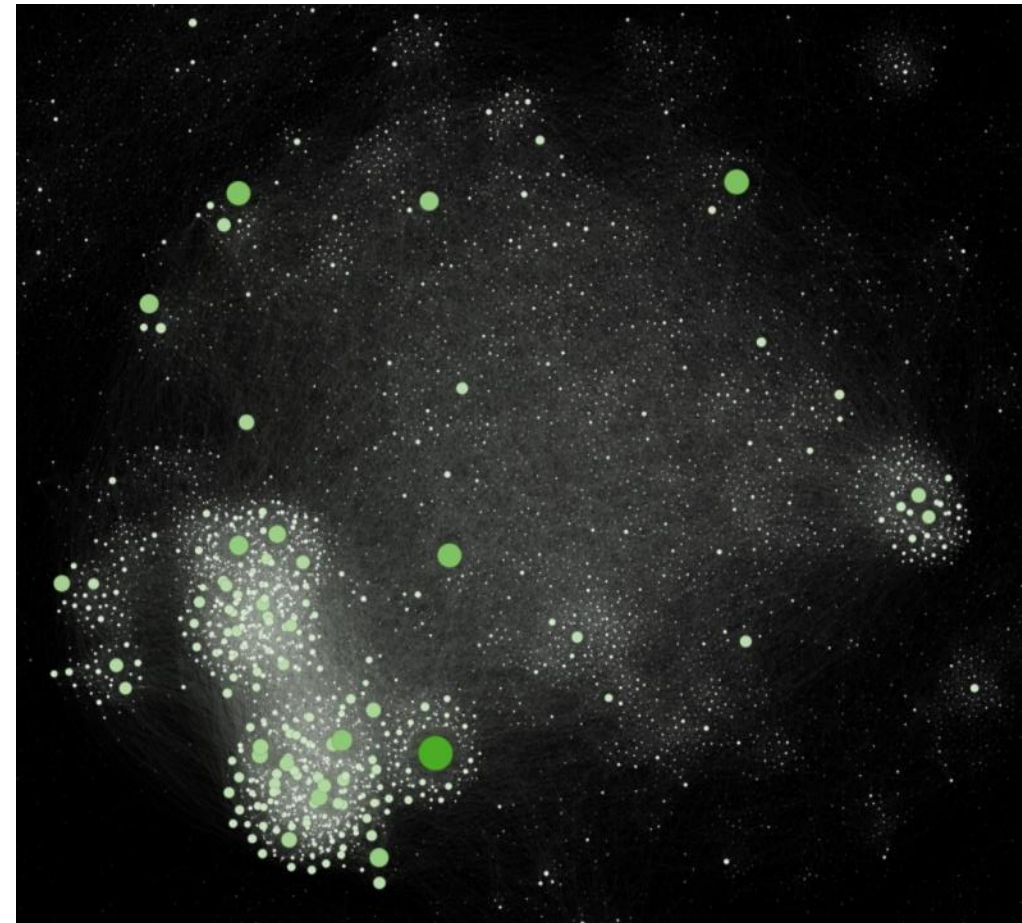
Redes sociais estão transformando as relações entre pessoas



Seguradoras: redes complexas de pessoas semelhantes às redes sociais



redes de amigos

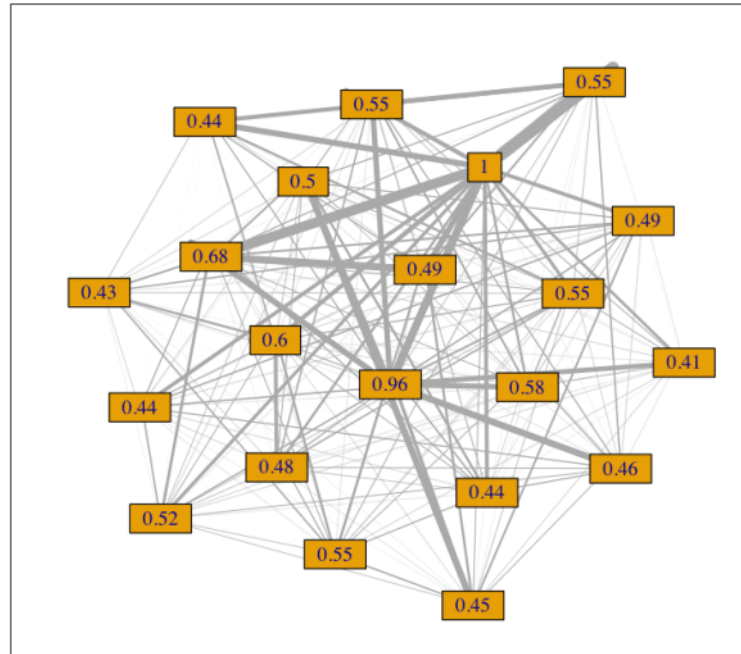


rede de médicos
com pacientes em comum

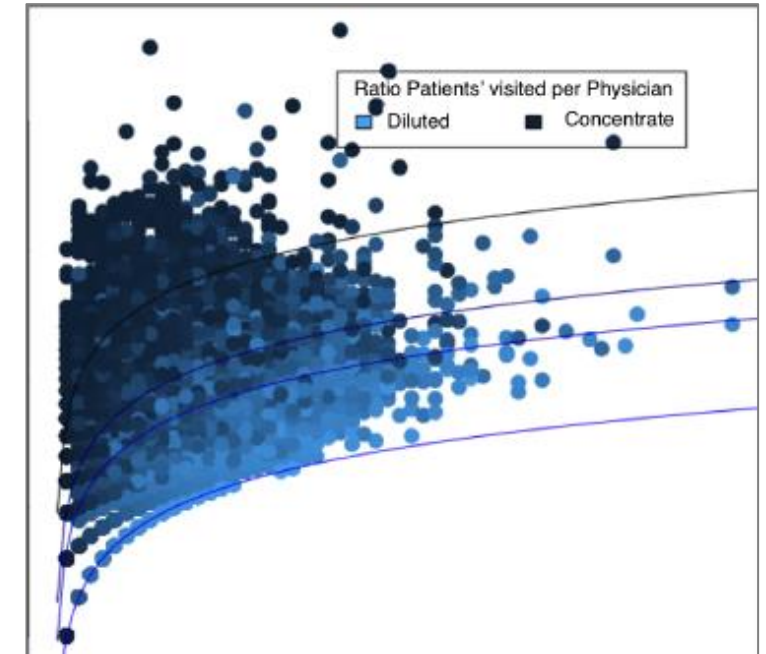
Case: construção de métricas para médicos com redes complexas



referência mútua

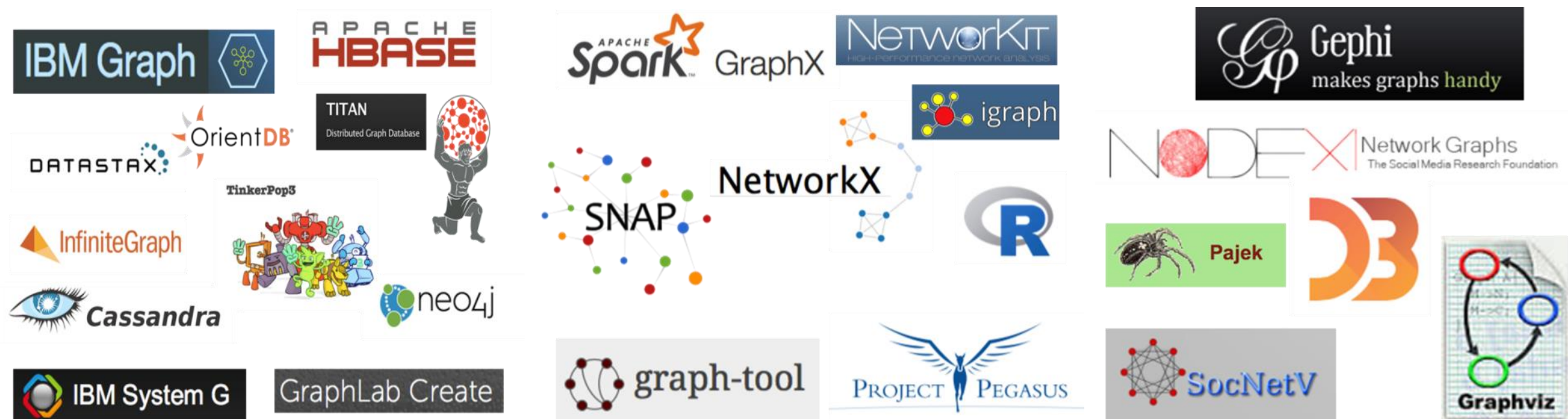


importância na rede



fidelização

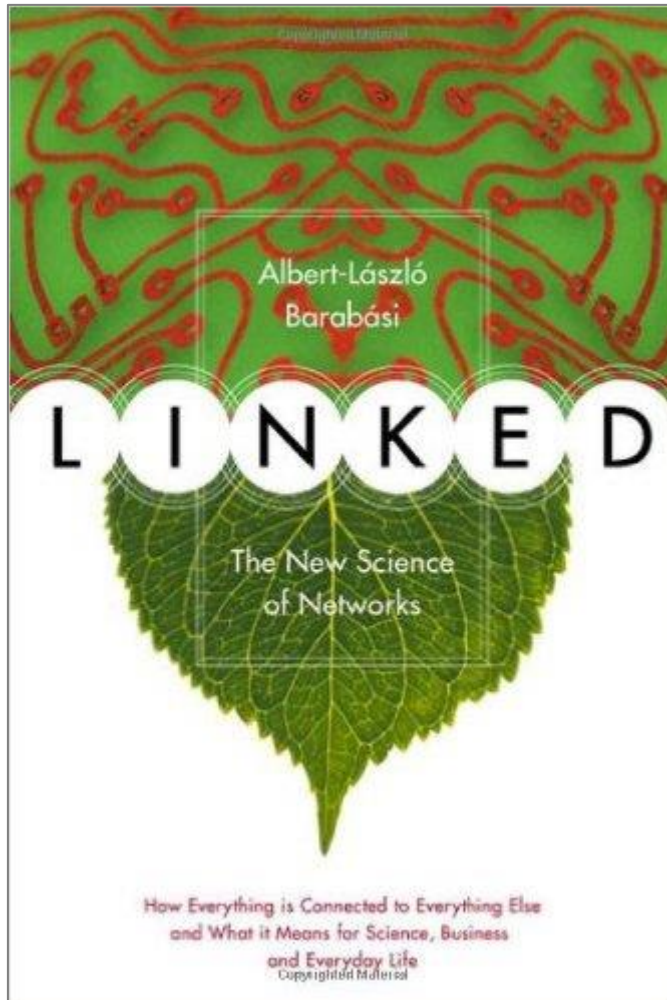
GRAPH ANALYTICS



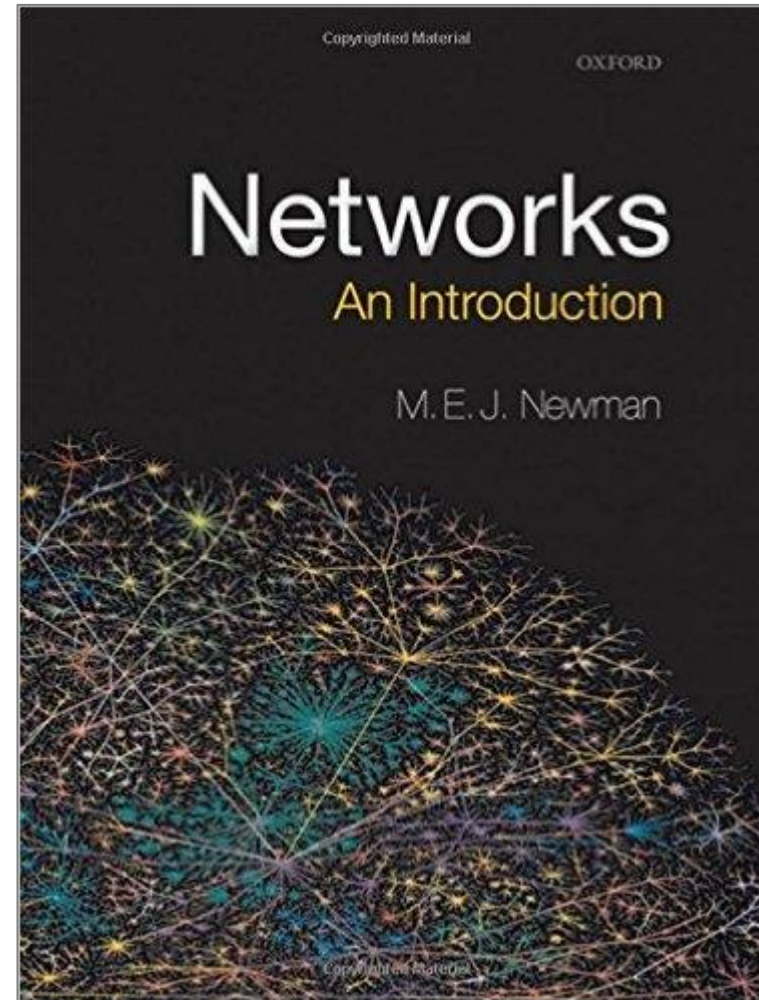
bancos de dados

ferramentas de análise

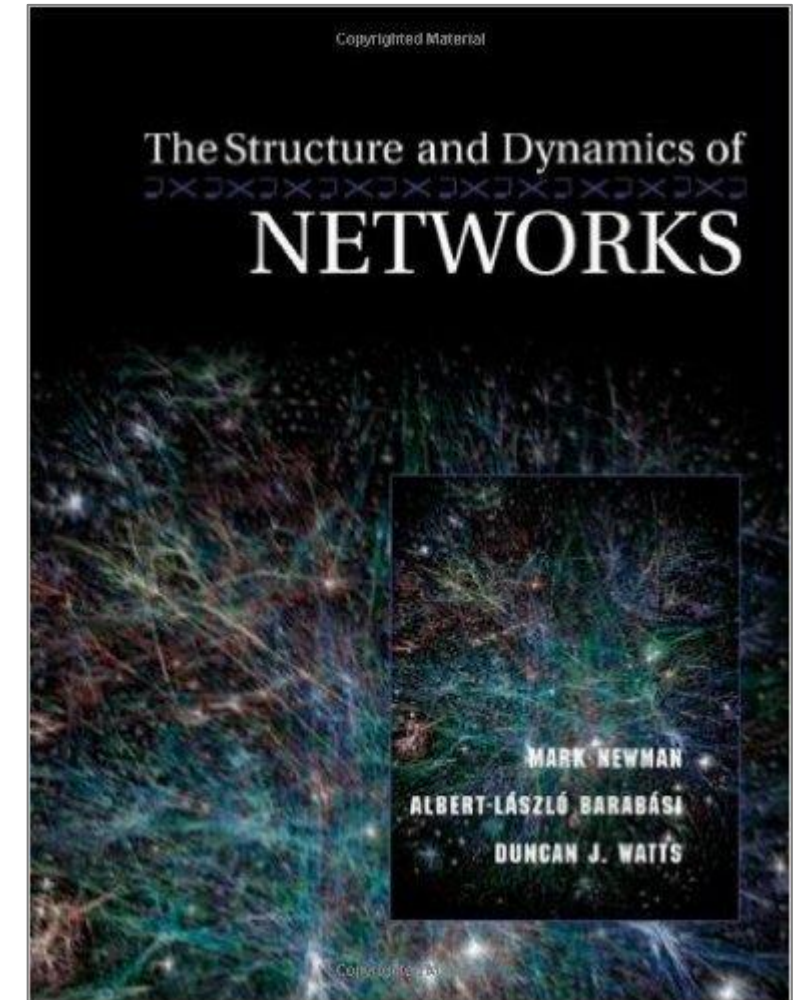
visualização



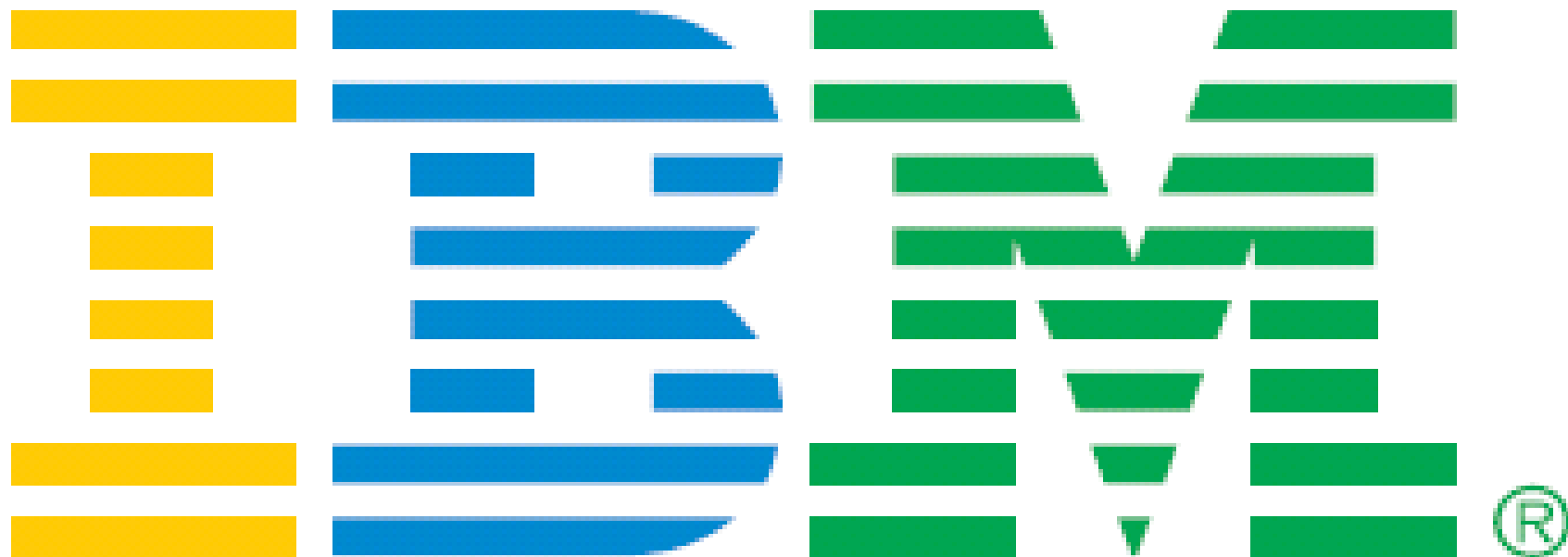
introdução



livro-texto



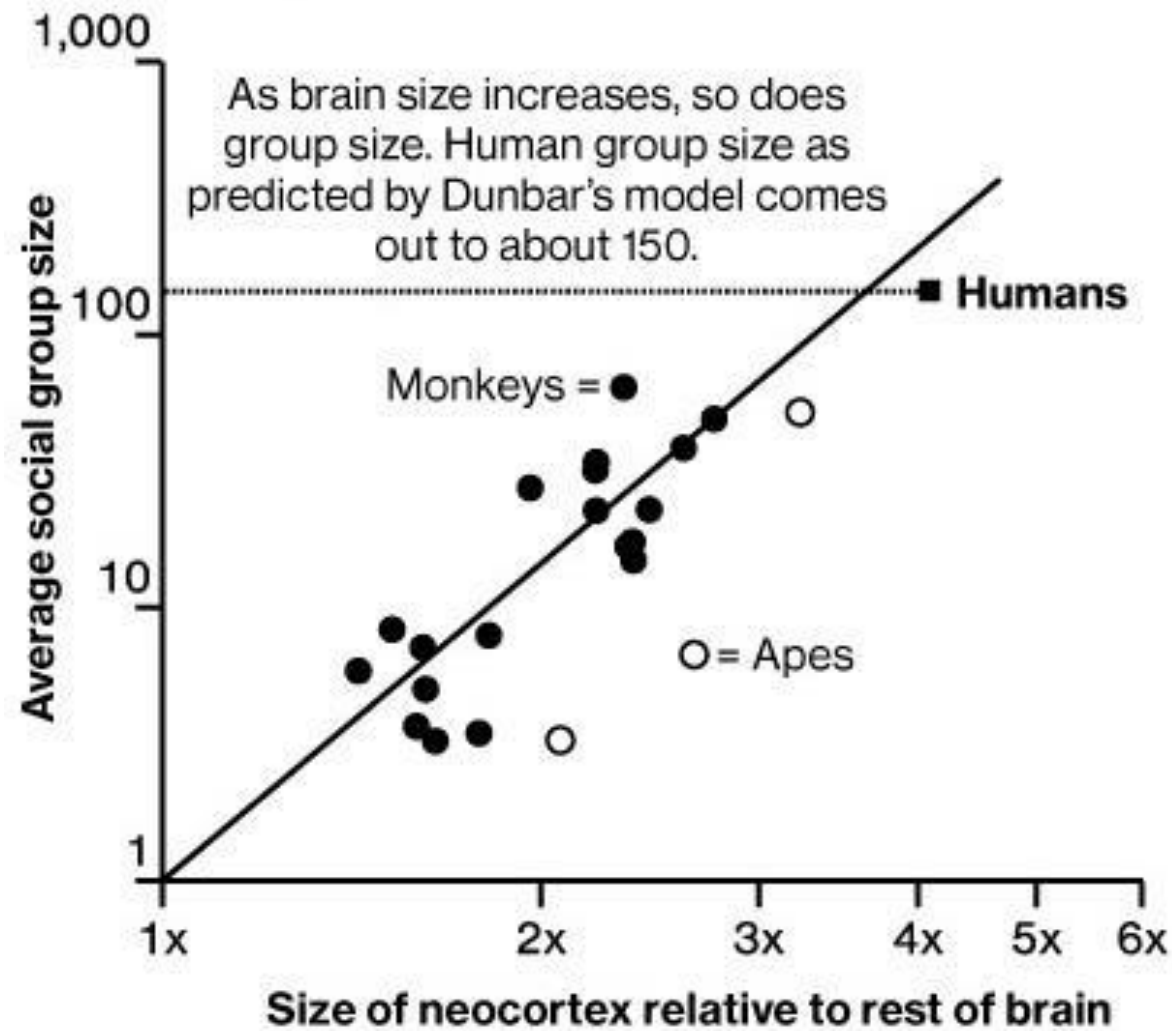
avanzado



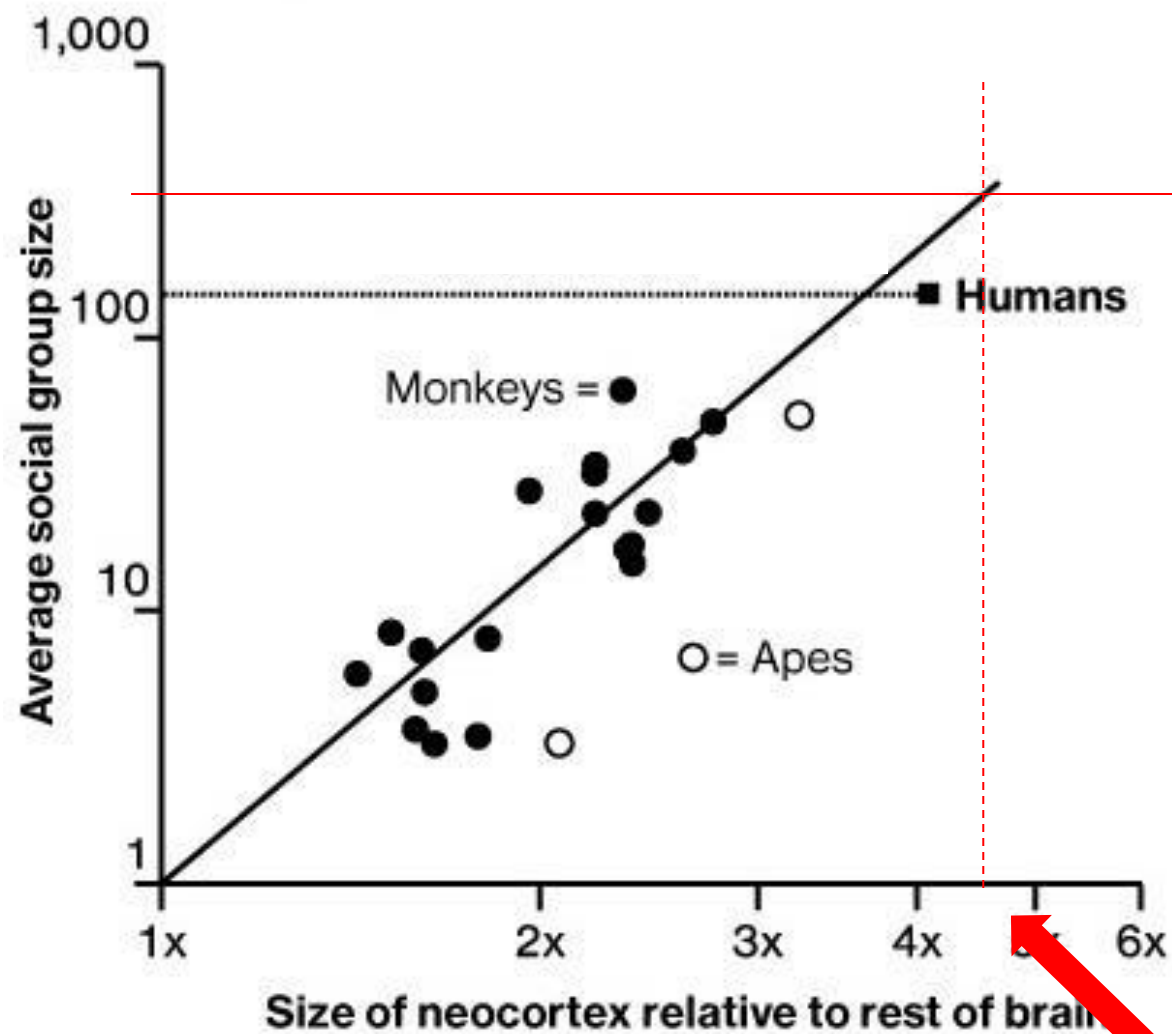
IBM Research – Brasil
csantosp@br.ibm.com, @cinhanez

BACKUP

Dunbar's Number

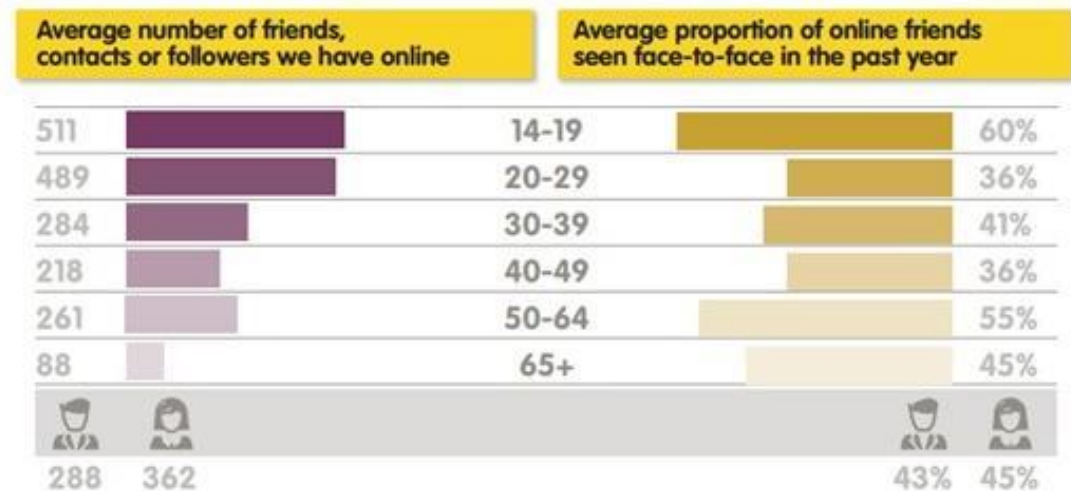


Homus Interneticus: social media aumentando o cérebro



DATA: THE SOCIAL BRAIN HYPOTHESIS, DUNBAR 1998

Source: <http://digitalmarketinglab.com.au/index.php/2014/10/02/australian-digital-stats-compilation-2014-30-fast-facts/>



neocortex cresceu 20% com mídias sociais

Studies about the relations in Health Professional Networks using SNA Techniques

- Understand the structure, characteristics and how they function in professional health networks (nurses, physicians, directors, etc) and its impact into care is important.
- Meltzer et al (2010) demonstrate how SNA can assist in the design of effective quality improvement teams. They use degree, betweenness and density in order to choose health professionals to compose a team care.
- Using degree, betweenness and density in physician networks, some recent works related these metrics with costs of hospitalization and readmission rate.
- High degree scores in physician collaboration network demonstrate that physicians collaborate with each other.
- A big fluctuation among betweenness centrality scores in physician collaboration network reflects an uneven collaborate and communicate with the member of the network.

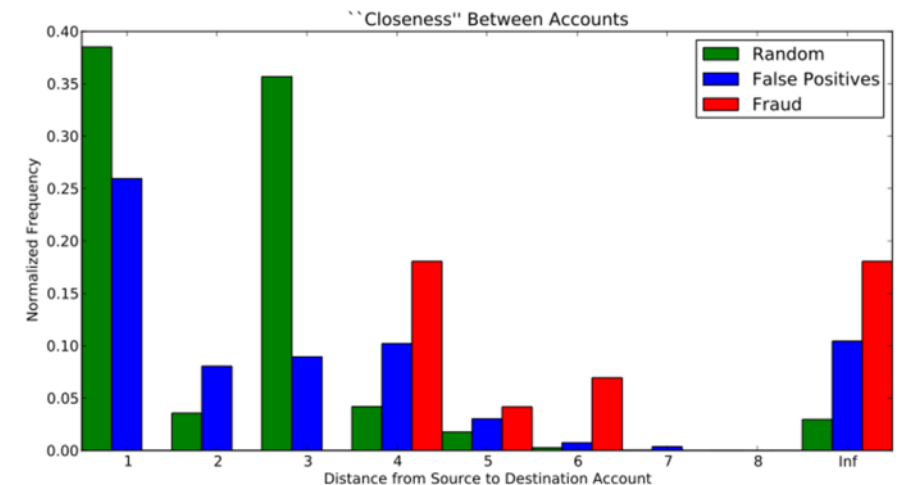
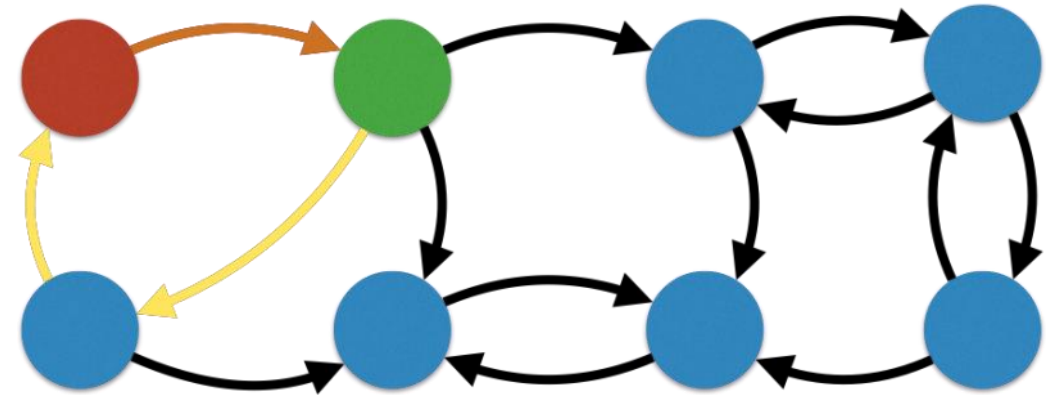
- Cunningham FC et al. “Health professional networks as a vector for improving healthcare quality and safety: a systematic review”. BMJ Qual Saf, 2012;21:239-249
- Meltzer D et al. “Exploring the use of social network methods in designing healthcare quality improvement teams”. Soc Sci Med. 2010;71(6):1119-1130
- Uddin S, Hossain L, Effects of Physician Collaboration Network on Hospital Outcomes. In: Proceedings of the fifth Australian Workshop on Health Informatics and Knowledge Management (HIKM 2012). Melbourne:Australia. 2012:67-73.
- Pollack CE, Weissman GE, Lemke KW, Hussey PS, Weiner JP, Patient Sharing Among Physician and Cost of Care: A Network Analytic Approach of Care Coordination Using Claims Data. J Gen Intern Med, 2012;28(3):459-65.
- Abdelzaher AF, Ghosh P, Mussawi AF, Jackson DS, Palesis JA, Deshazo JP, Care Coordination Metrics of Patient Sharing among Physicians: A Social Network Analytic Approach. In: Proceedings of BCB’14. Association of Computing Machinery (ACM). Newport Beach:USA, 2014:635-636.
- Uddin S, Kelaheer, M, Piraveenan, M, Impact of Physician Community Structure on Healthcare Outcomes. Stud Health Technol Inform. 2015;214:152-8.

- ***Hypothesis: Transaction networks define a ‘community’ for each account where ‘normal’ transactions are done***
- ***Notions of community:***
 - Entities an account transacts with
 - Accounts which are within “six degrees of separation” (shortest path)
 - Strongly connected component (paths exist in both directions)
 - Clusters of accounts with same pattern
 - Structure of egonets
 - Accounts to which there is a “reasonable” amount of flow

Example Analytic: Degree of separation



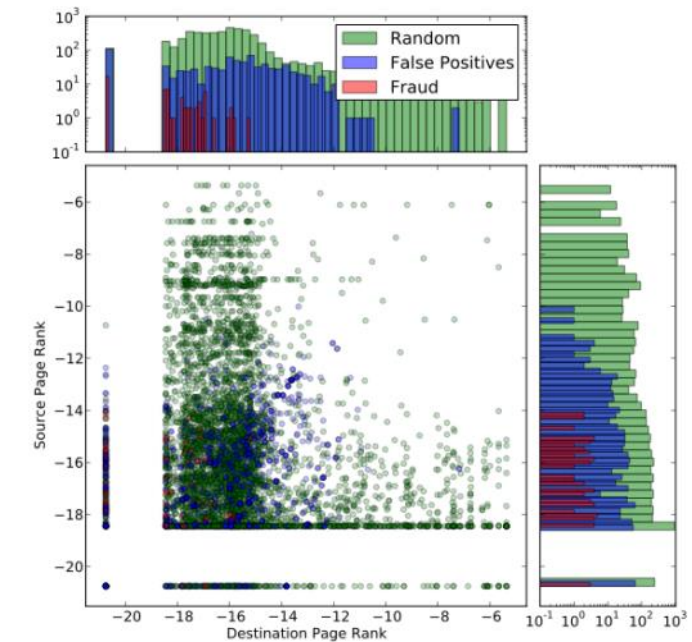
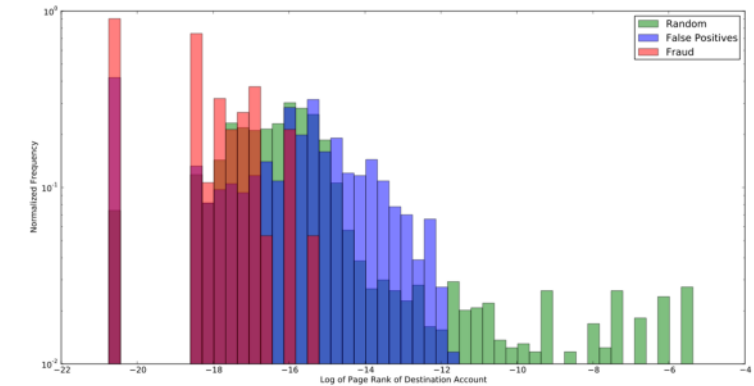
- Hypothesis: Trust can be a transitive property
- Measure the shortest path between any two accounts, the reverse path, and undirected path
- Shorter paths imply higher trust
- Some accounts have no connected path



Example Analytic: PageRank

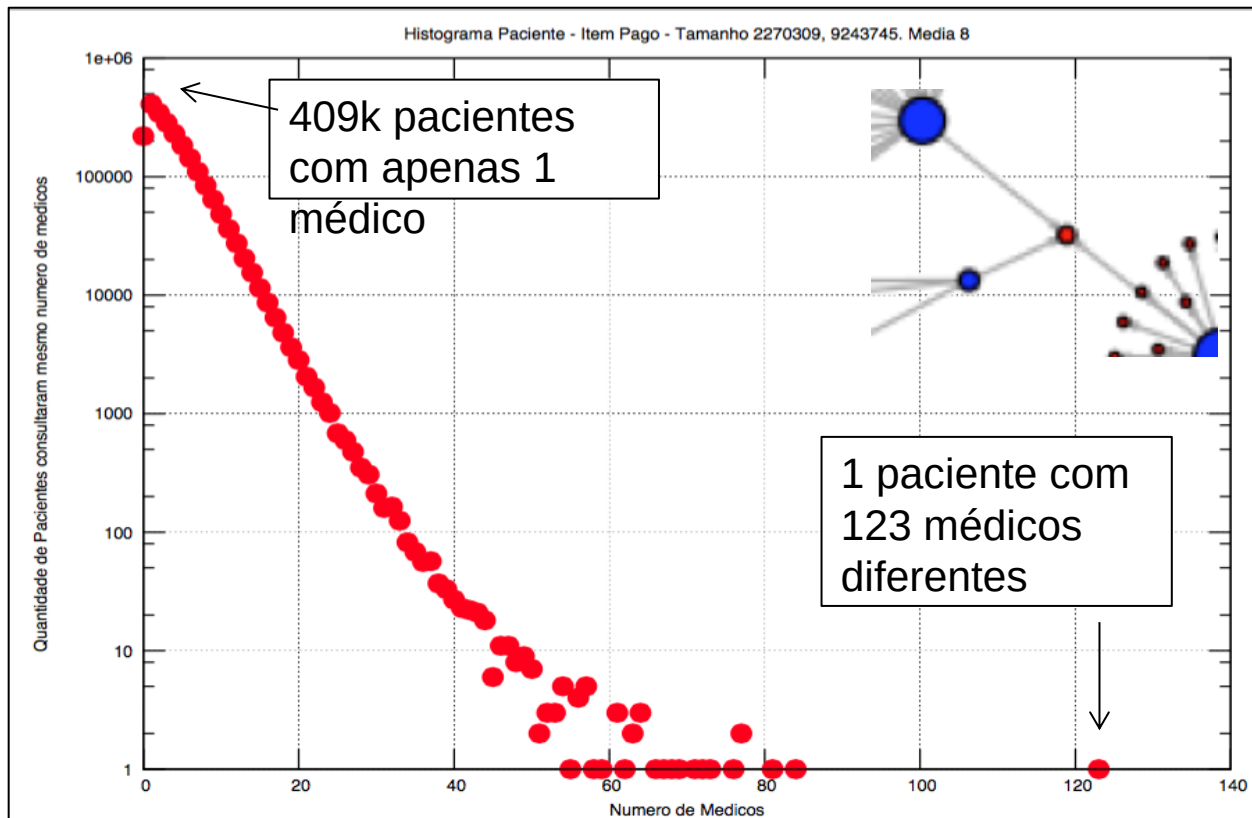


- Google's algorithm for measuring trust or importance in an account
- Recursively defined influence distribution
- Both unweighted and weighted (€) varieties
- “Reverse” PageRank for modeling source accounts
- Yields 5–50+% False Positive reduction and good AUC



Distribuição de Grau de Grafo de Médicos - Pacientes

histograma de pacientes



histograma de médicos

