

Sosyal Ağlar, Analiz Yöntemleri ve Araçları

Doktora atölyesi-16-20 Aralık 2011

Istanbul Bilgi Üniversitesi

Mehmet Gençer

mehmetgencer@yahoo.com

Mehmet.gencer@ieu.edu.tr

<http://mgencer.com>



Ağ nedir?

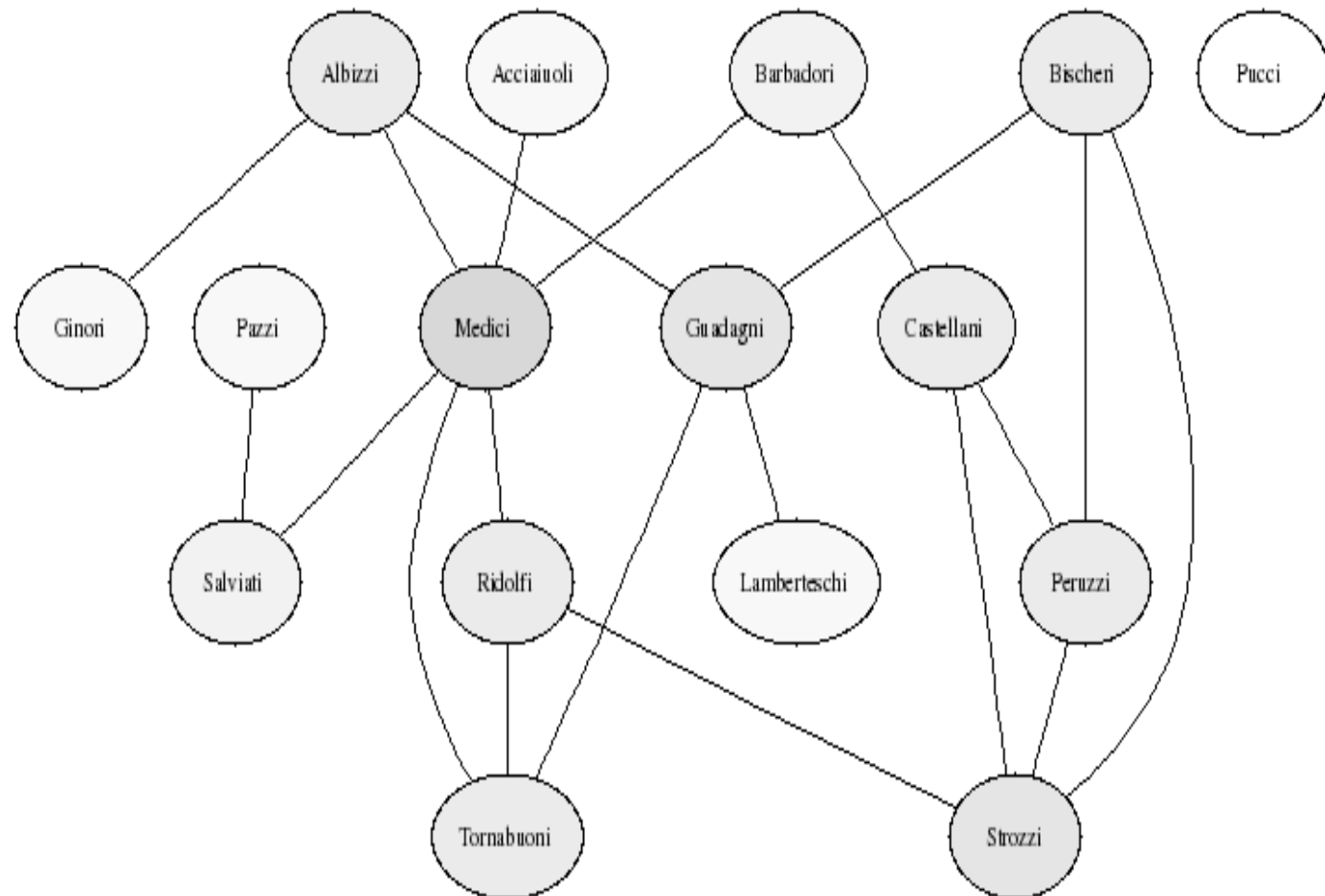
- Bir nesne seti, N , ve herbiri iki yada daha fazla nesneye dokunan bağların, B , toplamı ağı oluşturur:

- $A=(N,B)$

$$N=\{n_1, n_2, \dots, n_N\}$$

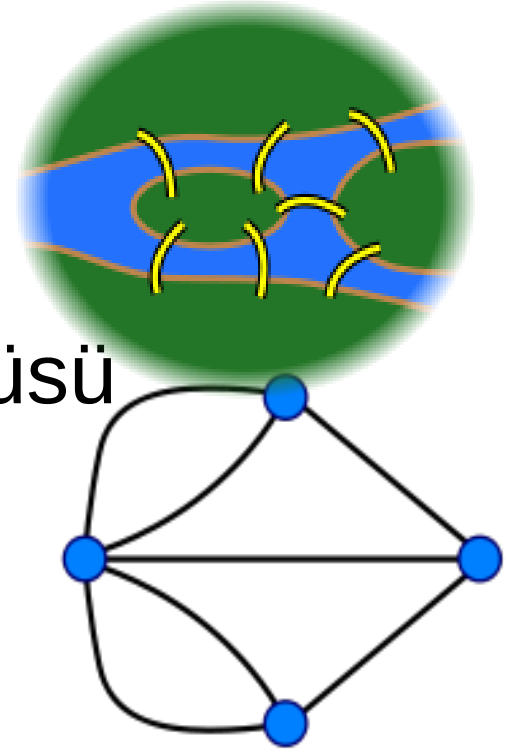
$$B=\{$$
$$(n_1, n_2),$$
$$(n_2, n_4),$$

...

$$\}$$


Ağlar

- Coğrafya/Topoloji:
Euler (1735): Königsberg'in yedi köprüsü
Yol ağı
- Tasarlanmış:
Elektronik devre bağlantıları
Yazılım paketleri arasındaki bağımlılık ilişkileri
- Doğal:
Hücre proteinleri arasındaki tetikleme ilişkileri
Ekosistem etkileşim ilişkileri



Ağlar ...

- Kişiler arası ağlar:
Sınıfta kim kimi seviyor (Moreno, 1930lar)
Seks ilişkileri (AIDS yayılımı)
Evlilik yoluyla aileler arası kurulan bağlar
Aynı şirketin yönetim kurulunda bulunmaktan
doğan ilişkiler ağı
- Organizasyonlar arası ağlar:
Terör hücreleri arası ilişki kurucular
board-interlocks
şirket ortak kurucuları (silikon vadisi)

Sosyal ağ nedir?

- Yaygın tanım: facebook, twitter, ..., nesnesi bireyler ve bağları onlar arasındaki iletişim olayları olan ağlar.
- örgüt/inovasyon-ekonomi/işletme araştırmacısı için olası bir tanım: ağdaki nesnelere VEYA ilişkilerin (üye/üreten/çalışan) *bireyler* olduğu ağlar. Örn.: e-posta grupları, tavsiye sistemleri, ..
- Communicatively Constituted Organizations(CCO)

Ağ veri seti

$$A=(\mathbf{N},B)$$

- Tanıdık veri seti: nesne **nitelikleri**

	Yaş	Cinsiyet	Gelir	Eğitim
Örnek 1				
Örnek 2				

Anket yöntemiyle toplanabilir.

Örneklemin nasıl seçileceği, ve popülasyonu temsil gücü konusunda güçlü istatistiksel kriterler ve literatür mevcut.

Analiz yöntemleri ve araçları yerleşmiş.

Ağ veri seti ..

$A=(N, \mathbf{B})$

- **İlişkisel** veriler:

Anket yöntemiyle toplanamaz. Belge/arşiv çalışması gerektirir.

Örnekleme meselesi son derece sorunludur.

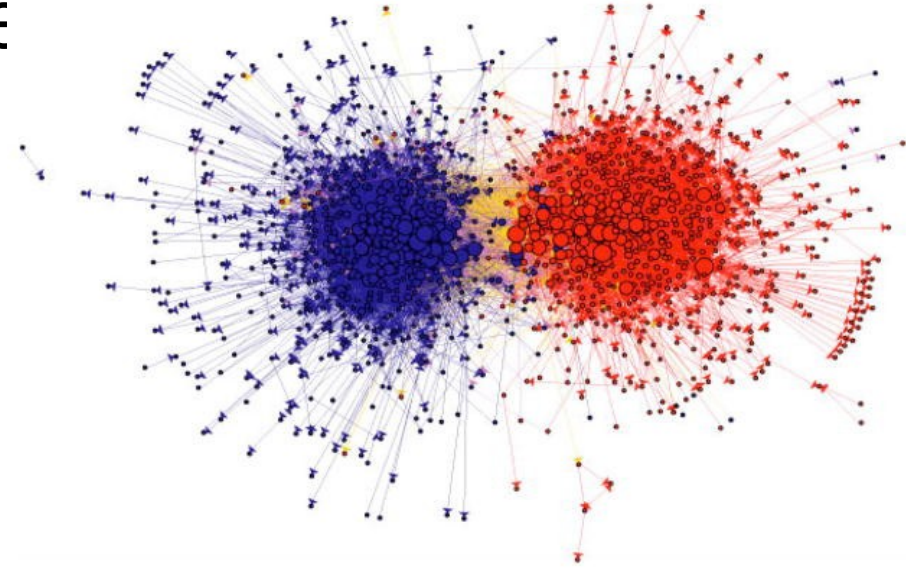
Analiz yöntem ve araçları çeşitli, dağınık, çoğu zaman araçlar zayıf ve/veya birbiriyle uyumsuz.

Bağlar aynı türden ilişkiyi temsil etmelidir. İki veya daha fazla türden ilişki içiren bir ağ sözkonusu olduğunda “ağsal” ölçümler yapılabilir ama bunları yorumlayamayız.

Bağların varlığının yanısıra gücü, varolduğu zaman aralığı, vb., ekstra veriler toplanabilir.

Ağ veri seti ...

- Ağ analizi ilişkilerin analizidir. Ancak nesne nitelikleri bilindiğinde ve birlikte kullanıldığında daha güçlü yorum imkanı doğar. (örnek: cumhuriyetçi ve demokrat bloglar)
- Alternatif olarak aynı nesne seti için iki farklı ilişkiden doğan iki ağın karşılaştırması da yorum imkanı verir.



İlişkisel verilerin temsili - sosyomatris

- İlk çalışmalar: sosyomatris (Moreno, 1938)

	n1	n2	n3	n4
n1	0	1	0	1
n2	1	0	0	0
n3	1	1	0	0
n4	1	0	0	0

Örneklem sorunu görülüyor

Avantajları: Matris matematiği gelişmiş, ölçüm ve analizleri tanımlamak/gerçekleştirmek kolay.

Dezavantajları: (1) Örneklem boyutu (N) artınca çok fazla büyüyor (N^2). (2) Bağların zaman aralığı, vb., verileri temsil etme imkanı vermiyor.

Notlar: (1) bağlar tek yönlü ise matris simetrik, (2) bağların gücü birbirinden farklı olabilir.

İlişkisel ölçümler - sosyomatris

- Moreno: “the whole configuration is built up by its single elements”. Gestalt yöntemlerinden ilişkisel sosyolojiye doğru bir ayrılma.

Veri seti: (1) sınıf öğrencilerinin birbirini sevme bilgisi, asimetrik matris (sevme tek yönlü bir ilişki), (2) aynı nesne seti için rastgele sayılarla doldurulmuş kontrol verisi.

Ölçümler:

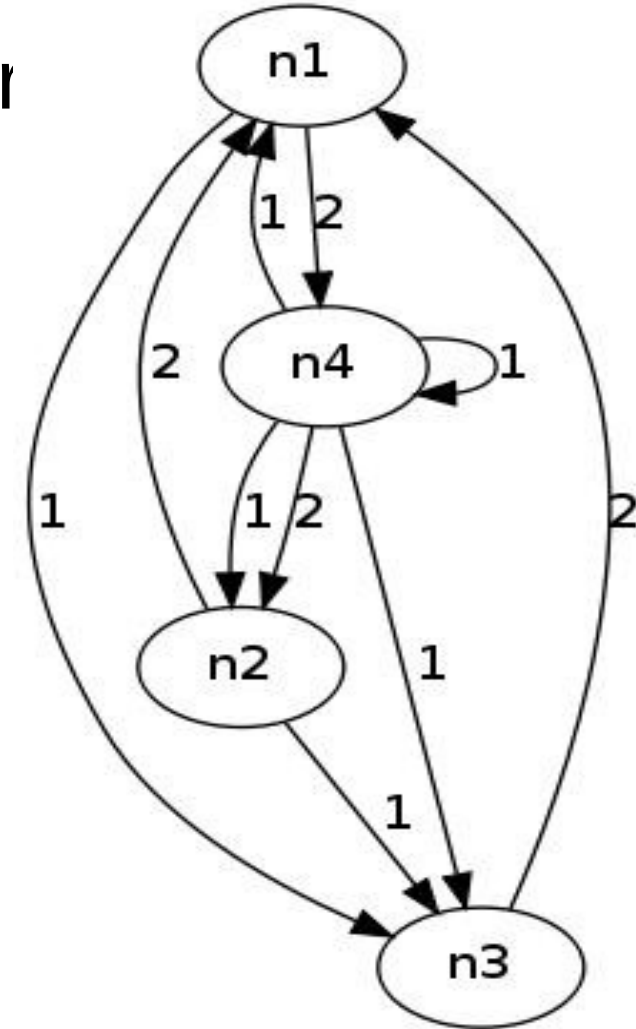
- nesne ne kadar seviliyor: matriste satır toplamı, nesne bazlı ölçü
- Nesne çiftlerinde (dyad) karşılıklılık: simetrik hücre bilgisi, 'dyadic' ölçü

Analiz: karşılıklılık rastlantısal mıdır? İki veri setinin karşılaştırması. Sonuç: gerçek veride karşılıklılık kontrol verisine⁰ (ve kuramsal olasılık hesaplarına) göre %213 daha fazla.

İlişkisel verilerin temsili - çizge

- Matematikte kullanılan 'çizge' (graph) ile temsil (Euler, 1700s). Her 'düğüm' için onu diğer düğümlere bağlayan bağların (ve bağ gücünün) listesi tutulur:

n1: (n3,1), (n4,2)
n2: (n1,2), (n3,1)
n3: (n1,2)
n4: (n1,1), (n2,2),
(n2,1), (n3,1), (n4,1)



İlişkisel verilerin temsili -çizge..

- Çizge temsili sosyomatrix'in kısıtlamalarını kaldırır. Temsilin boyutu $(N+B)$ bağ sayısı ile orantılıdır.
- Bağların gücü, zaman aralığı, vb., niteliklerini saklamaya uygundur.
- Dezavantajları: Çizge matematiği matrix matematiğine göre daha az olgunlaşmıştır. Bilgisayarda çizge temsili için çok sayıda birbiriyle yarışan standart var. Analiz ve görselleştirme için birbiriyle çoğu zaman uyumsuz çok sayıda yazılım aracı var.

İlişkisel verilerin temsili -çizge...

Çizge temsil standartlarının çeşitliliği veri setlerinin çeşitliliğinden kaynaklanıyor:

- Tek-kümeli/iki-kümeli ağlar
- Yönlü/yönsüz bağlar
- Bağ gücü bilgisinin olup olmayışı
- Zaman bilgisinin olup/olmayışı
- Bağ niteliklerinin varlığı

ÖNERİ: Excel/CSV kullanın

Ağ ölçütleri

Ağ düzeyi ölçütler

- Yoğunluk (density) = $2 / N*(N-1)$: olası bağların ne kadarı mevcut? Büyüklüğü veya doğası farklı ağlar arasında karşılaştırılmaz.
- Clustering coefficient

Nesne düzeyi ölçütler

- Ego density: “arkadaşlarım ne kadar birbirini tanıyor”
- Derece Merkeziliği (degree centrality): nesneye bağlı bağların sayısı.

İkili düzeyi, grup düzeyi, ...

Ağlarda akış

- Granovetter: “The strength of weak ties”, iş bulma konusunda yakın çevreden ziyade uzak bağlantıların etkisi.

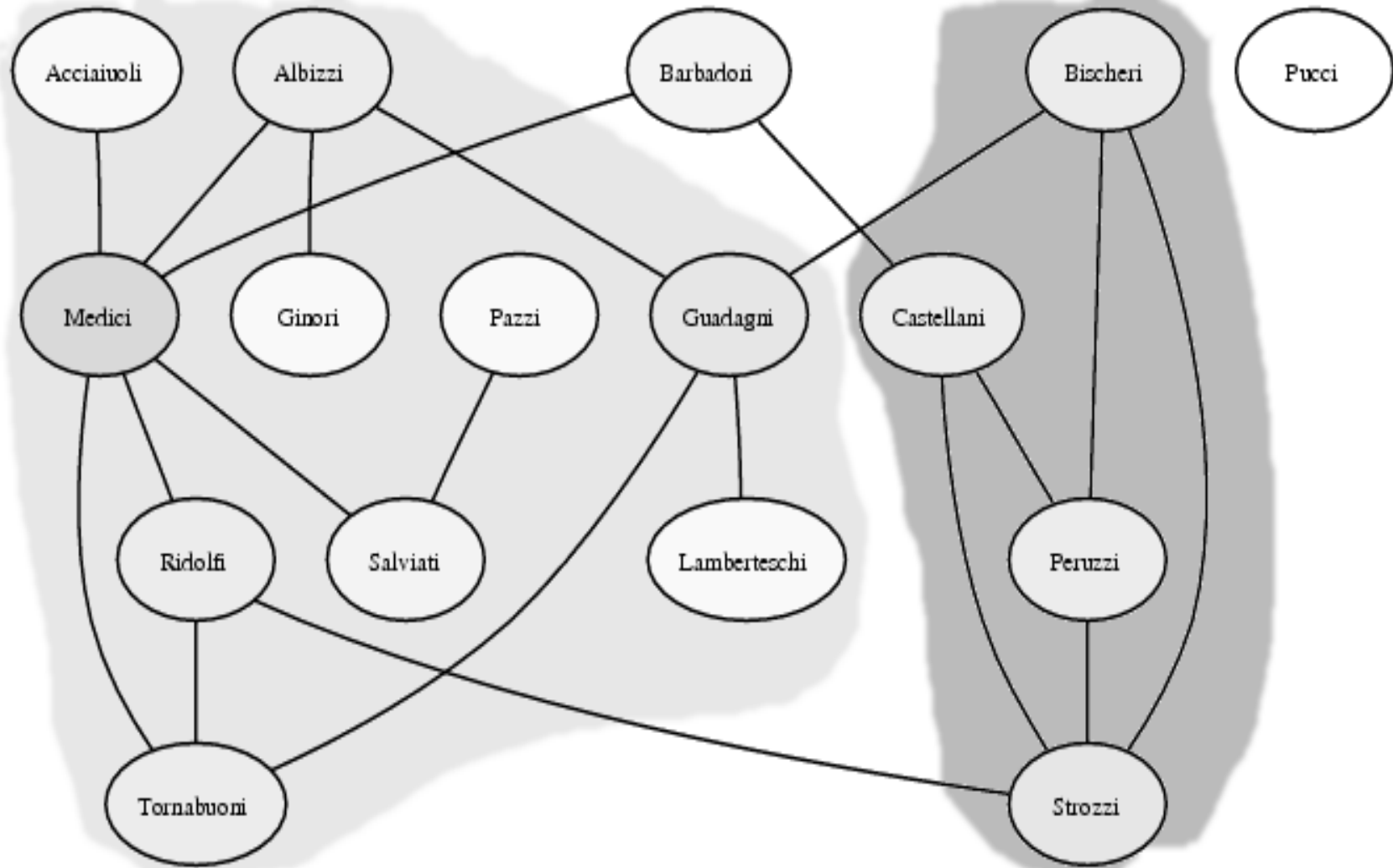
Kohezyon ve Difüzyon ikiliği

- Burt “Brokerage and Closure”,

Örneklerle ağ tipi ve ölçütleri -1

- Ortaçağ Floransa'sında ailelerarası evlilik ve iş bağları:
- Politik güç ↔ **betweenness centrality**
“bir nesne tüm diğer nesnelere birleştiren tüm yolların ne kadarı üzerinde yer alıyor”. “degree centrality” yerel bir ölçüt iken bu **ağ genelini hesaba katan bir nesne ölçütü.**
- Sonuç: Medici ailesi tüm iktisadi/politik ağa erişimin kavşağında

Floransa Aileleri



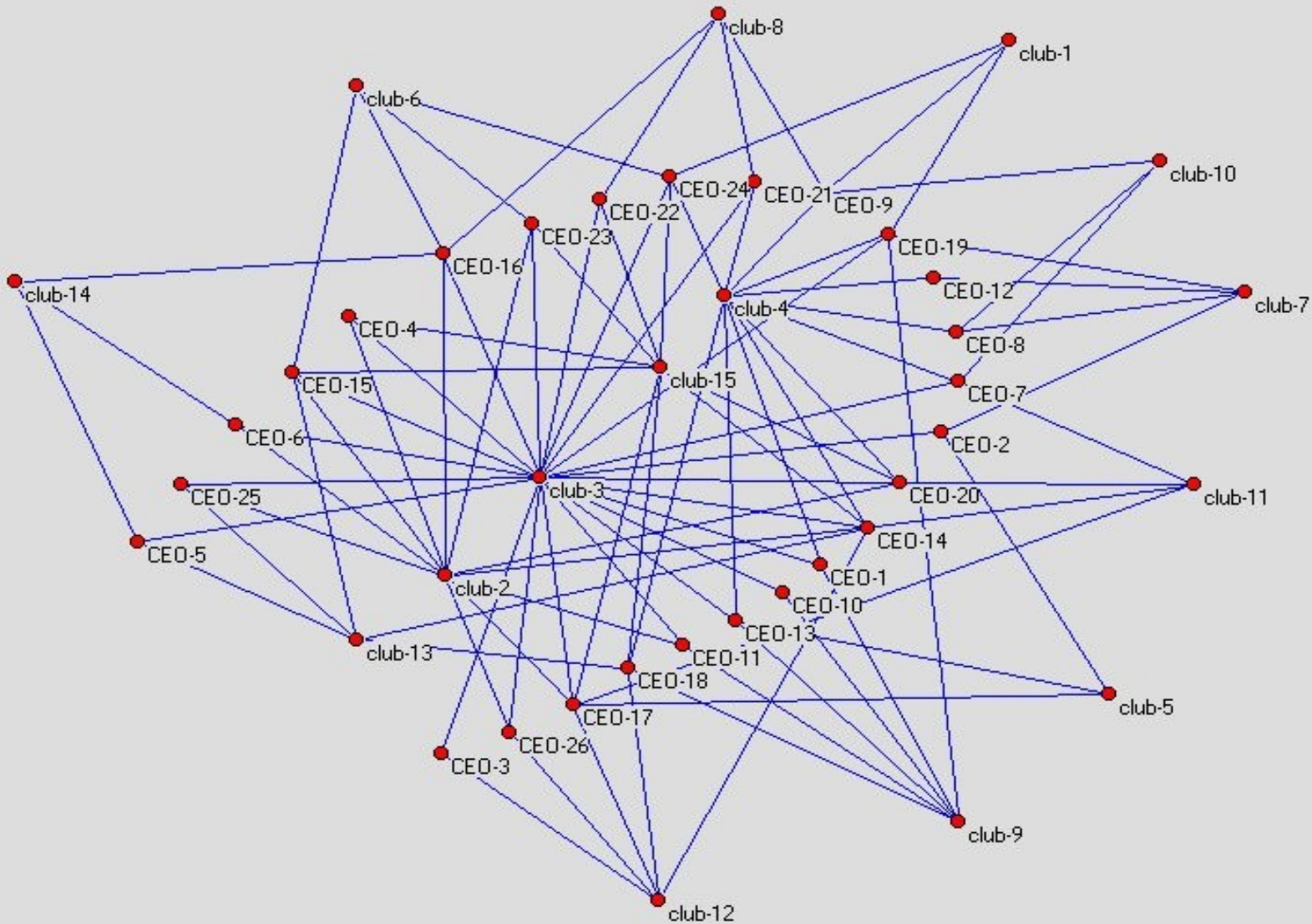
Örneklerle ağ tipi ve ölçütleri -2

- Galaskiewicz: CEO'ların sosyal klüp üyelikleri.
- İki-kümelili ağ!

	Klüp 1	Klüp 2	Klüp 3	Klüp 4
Kişi 1	1		1	
Kişi 2	1			1
Kişi 3	1			1
Kişi 4		1	1	
Kişi 5		1		

- İki şekilde tek kümelili ağa çevrilebilir
- Ölçütler: closeness centrality, information centrality

CEO Klüp üyelikleri



Farklı uygulamalar

- Google: 'pagerank'
- Webometri, scientometri

Daha çok ölçüt

- Burt: Yapısal gedikler ve otonomi.
- Watts ve Strogatz: “**Small Worlds**”
kümelenme katsayısı (clustering coefficient):
yereller ne kadar bağlı
Uzaklık: yerel kümelerin ortalama uzaklığı
- Ağ çapı: bağlı bir ağda nesne çiftleri arası en kısa uzaklıkların en uzununu (“six degrees of separation”)

Daha çok yöntem

Grupların bulunması:

- Bileşen: hepsi birbirine bağlı nesnelerin oluşturduğu grup
- K-klik: aralarındaki bağlantı en fazla k nesneden geçen grup
- Küme: gevşek bir tanım, aralarındaki bağlantı yoğunluğu -görelî- yüksek nesne seti

Daha çok yöntem ..

Ağ yapısı, nesne rolleri:

- Yapısal denklik: ilişkileri aynı olan nesneler, örneğin sosyomatrix satırlarının benzeşimi kullanılarak
- Blok-model: bağ yapıları denk/yakın nesneleri kümeleyince ortaya çıkan yapı (örn: cumhuriyetçi-demokrat blogları)

Daha daha çok ...

- Olasılık yöntemleri (Ikeda ve Strauss): üçlü bağların (triad, “arkadaşımın arkadaşı arkadaşımıdır”) raslantısal olup olmadığı, ve nesne özelliklerinin kombinasyonu ile ne derece ilgili olduğu.
- Fizik alanında: biyolojik/doğal/mühendislik problemleriyle ilgili olarak, rasgele ağlar (Erdős).

Önemli eksikler

- Dinamik/zamansal modeller yok (fizikte, çok sınırlı)
- Hızlı değişen (twitter) ağlar
- Araçlar-arası entegrasyon, standartlar

Araçlar

- Eski: UCINET
- Fizik etkisi ağırlıklı: Pajek
- Sosyal ağ araştırmaları kökenli: Gephi
- İstatistik entegrasyonu: R/Statnet

CCO: Communicatively Constituted Organizations

- Ortaya çıkan bir akademik alana takılmış yeni bir isim, self-organization etkileri, iş örgütlenmesinin aşağıdan yukarıya oluşumu.
- Esasları:
 - (1) İletişim olaylarını inceler
 - (2) Her türlü örgütsel iletişimi içerir
 - (3) İletişimde mütekabiliyeti esas alır
 - (4) Kim ve ne oluşu sürekli şekillenen bir olgu sayar
 - (5) iletişim olaylarının dışına çıkmaz
 - (6) Ne örgüt ne örgütleniş'e ağırlık verir

Kaynaklar

- John Scott, 1991, "Social Network Analysis: An introduction", Sage
- J.T.L. Moreno, H.H.Jennings, 1938. Statistics of Social Configurations, Sociometry 1/3.
- Mark S. Granovetter, "The Strength of Weak Ties", American journal of sociology, 1973
- Cooren et al., Communication, Organizing and Organization, Organization Studies September 2011 vol. 32 no. 9 1149-1170