Um Algoritmo para Encontrar a Cobertura de k-Cliques em Redes Sociais

Luís Cavique [1], Armando B. Mendes [2], Jorge M.A. Santos [3]

[1] Universidade Aberta, lcavique@univ-ab.pt

[2] Universidade dos Açores, amendes@uac.pt

[3] Universidade de Évora, jmas@uevora.pt

Resumo. Na análise de redes sociais, uma k-clique é a relaxação de uma clique, i.e., uma k-clique é um quase sub-grafo completo. Um k-clique num grafo é um sub-grafo onde a distancia entre quaisquer par de vértices não é maior que k. A visualização de um pequeno número de vértices é fácil de obter. Contudo, quando o número de vértices aumenta a visualização torna-se incompreensível. Nesta comunicação, propomos uma nova abordagem na extracção de conhecimento em grafos, utilizando k-cliques. O conceito que clique relaxado é estendido para todo o grafo, de forma a ter uma visão geral, ao cobrir a rede com k-cliques. Sequências de coberturas de k-cliques são apresentadas combinando o conceito dos "pequenos mundos" com estruturas com coesão. Resultados computacionais e exemplos são apresentados.

Palavras-chave: extracção de conhecimento em dados, extracção de conhecimento em grafos, redes sociais

Abstract. In social network analysis, a k-clique is a relaxed clique, i.e., a k-clique is a quasi-complete sub-graph. A k-clique in a graph is a sub-graph where the distance between any two vertices is no greater than k. The visualization of a small number of vertices can be easy to perform in a graph. However, when the number of vertices and edges increases the visualization becomes incomprehensible. In this paper, we propose a new graph mining approach based on k-cliques. The concept of relaxed clique is extended to whole the graph, to achieve a general view, by covering the network with k-cliques. The sequence of k-clique covers is presented, combining small world concepts with community structure components. Computational results and examples are presented.

Keywords: data mining, graph mining, social networks

1 Introdução

Depois da comunicação de Tim Berners-Lee [2] na conferência internacional de World Wide Web WWW2006 sobre as três idades da Web, as redes sociais associados à Web 2.0, suscitaram uma explosão de interesse, na tentativa de melhorar a socialização e na criação de novos modelos de gestão dos conhecimentos.

O termo "redes sociais" foi cunhado por Barnes em 1954, no entanto, a visualização através de um grafos, chamado sociogramas, foi apresentado por Moreno [18]. Esta área científica da sociologia tenta explicar como funciona difusão da inovação, porque é que as alianças e os conflitos são gerados em grupos, como é que a liderança emerge e como a estrutura do grupo afecta o sua eficácia.

Um importante desenvolvimento sobre a estrutura das redes sociais teve origem numa experiência realizada pelo psicólogo americano Stanley Milgram [16]. A experiência de Milgram consistiu em enviar cartas de pessoas no Nebraska, no Centro-Oeste, para serem recebidas em Boston, na costa leste dos EUA, com a seguinte particularidade – as pessoas eram instruídas a passar as cartas de mão em mão, a alguém que conheciam, até chegarem à origem. As cartas que chegaram ao destino foram passadas por cerca de seis pessoas. Milgram concluiu que, os americanos não estão a mais de seis passos entre si. Esta experiência deu origem ao conceito de "seis graus de separação" e ao conceito de "pequeno mundo".

Um exemplo interessante dos “pequenos mundos” é o "Número de Erdos" [12]. Erdos é um dos mais prolíficos matemáticos de todos os tempos, sendo autor de mais de 1500 artigos publicados com mais de 500 co-autores. Erdos é o número zero e os investigadores que trabalharam com ele são chamados de Erdos número 1. Os co-autores de Erdos número 1 são conhecidos como Erdos número 2, e assim por diante, construindo um dos mais antigos “pequenos mundos”.

O trabalho de Erdos e Renyi [8] apresenta interessantes propriedades dos grafos aleatórios. O interesse foi reavivado recentemente com o modelo de Watts e Strogatz, publicado na revista Nature [22], que apresenta novas propriedades dos pequenos mundos.

Os analistas de redes sociais precisam de realizar inquéritos de cada pessoa sobre os seus amigos, pedir a sua aprovação para publicar os dados e acompanhar a população durante anos. Por outro lado, as redes sociais como LinkedIn MySpace podem fornecer os dados necessários sem este tipo de esforço.

A visualização de um pequeno número de vértices de um grafo pode ser desenhado na sua totalidade. No entanto, quando o número de vértices e arestas aumenta a visualização torna-se incompreensível. A grande quantidade de dados extraídos da Internet não é compatível com o desenho completo do grafo. Existe, portanto, uma necessidade premente para encontrar novas ferramentas para reconhecimento de padrões e métodos estatísticos para quantificar grandes grafos e prever o comportamento de sistemas de rede.

A extracção de conhecimento de grafos pode ser definida como a arte e a ciência de encontrar informação útil de grafos, como padrões e “outliers”, fornecidos respectivamente, por dados repetidos ou dados esporádicos existentes em grafos de grandes dimensões ou de redes complexas [9], [6].

Neste trabalho, propomos criar novas medidas com base na cobertura do grafo com k-cliques, com o objectivo de obter uma visão geral do grafo. Na secção 2, revemos os conceitos de redes sociais, detalhando as estruturas de subgrupos coesos. Na secção 3 apresentamos o algoritmo de duas fases que analisa em primeiro lugar para subgrupos coesos e, em segundo lugar descobre o conjunto mínimo de subgrupos coesos que cobrem todos os vértices. Na secção 4 são apresentados os resultados computacionais e exemplos. Finalmente, no capítulo 5, são apresentadas as conclusões.

2. Conceitos de Redes

A análise de redes Sociais é uma abordagem muito que surgiu na moderna sociologia, e estuda a interacção entre indivíduos, organizações e outros tipos de entidades. Em [19] e [21] podemos encontrar as bases teóricas e as principais técnicas de redes sociais.

A representação de uma rede social foi bastante influenciada pela teoria grafos. Nas redes sociais o conjunto de vértices (ou nós) correspondem aos "actores" (i.e., pessoas, empresas, agentes sociais) e do conjunto de arestas correspondem as "ligações" (isto é, relacionamentos, associações, links).

Uma das aplicações mais comuns da sociologia tem sido o tema da coesão social. Os subgrupos coesos podem incluir grupos como, grupos de trabalho, equipas desportivas, partidos políticos, cultos religiosos, ou estruturas secretas, como as organizações criminosas ou terroristas células. Nesta secção, detalha algumas técnicas de subgrupos coesos como as cliques e cliques relaxadas, como o k-clique, k-club/k-clã e k-plex.

2.1 Clique

Dado um grafo não orientado G= (V, E) onde V denota o conjunto de vértices e E o conjunto de arestas, o grafo G1= (V1, E1) é chamado um sub-grafo de G se V1⊆V, E1⊆E e para cada aresta (vi,vj)∈E1 os vértices vi, vj∈V1. Um sub-grafo G1 é completo se existe uma aresta para cada par de vértices. Um sub-grafo completo é também chamado de clique, ver figura 1. Uma clique é máxima, se não está contida numa outra clique. O número de clique de um grafo é igual à cardinalidade da maior clique de G e é obtido pela resolução do problema NP-difícil da máxima clique.

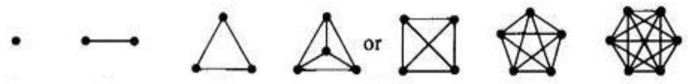

Fig. 1. Cliques com 1, 2, 3, 4, 5 e 6 vértices

A estrutura de clique, onde deve existir uma aresta para cada par de vértices, apresenta muitas restrições na modelação da vida real. Deste modo, abordagens alternativas têm sido sugeridas, com o fim de relaxar o conceito de clique, como o k-clique, k-clã/k-clube e k-plex.

2.2 k-clique

Luce [15] introduziu o modelo baseado na distância chamado k-clique, onde k é o comprimento máximo do caminho entre cada par de vértices. Uma k-clique é o subconjunto de C tal que para cada i, j∈C, a distância d(i, j) ≤ k. Uma 1-clique é idêntica a uma clique, pois a distância entre os vértices é igual a um. Uma 2-clique é o sub-grafo completo maximal com um caminho de uma ou duas arestas. O caminho de comprimentos dois pode ser exemplificado pelo "amigo do amigo" das relações sociais. Em sítios sociais como o LinkedIn, cada membro pode visualizar as suas ligações e também as ligações a dois e três graus de distância. O aumento do valor k corresponde a um relaxamento progressivo do conceito de clique. Veja figura 2.

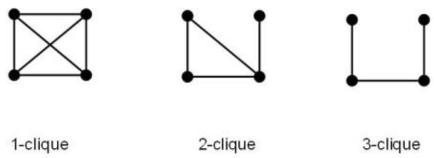

Fig. 2. Exemplos de k-cliques

2.3 k-clã e k-clube

Uma limitação do modelo k-clique é que alguns vértices podem ficar demasiado afastados do grupo. Para superar este inconveniente Alba [1] e Mokken [17] introduziram os modelos baseados no diâmetro, com o nome de k-clube e k-clã, respectivamente. O comprimento do caminho mais curto entre os vértices u e v em G é denotado por dG(u,v). O diâmetro de G é dado por diam(G)=max dG(u, v) para todos u,v∈ V. Para encontrar todas as k-clãs, em primeiro lugar todas as k-cliques são encontradas e, em seguida, é aplicada a restrição diam(k-clique)≤k, para remover as k-cliques indesejadas. Na figura 3, à esquerda, a 2-clique (1,2,3,4,5) foi removido porque o dG(4,5)=3. Outra abordagem possível, dos modelos baseados no diâmetro, é o k-clube que é definido como um subconjunto de vértices S tal que diam(G[S])≤ k. Na figura seguinte, duas 2-cliques: (1,2,3,4,5) e (2,3,4,5,6), um 2-clã: (2,3,4,5,6) e três 2-clubes: (1,2,3,4), (1,2,3,5) e (2,3,4,5,6) podem ser encontrados.

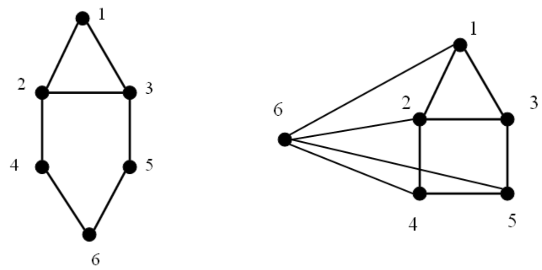

Fig.3. 2-clãs, 2-clubes (esquerda) e 3-plex (direita)

2.4 k-plex

Uma outra forma de relaxar uma clique é através do conceito de k-plex, que tem em consideração o grau vértices. O grau de um vértice de um grafo é o número de arestas incidentes no vértice, e é denotado por deg(v). O grau máximo de um grafo G é o grau máximo dos seus vértices e é denotado por $\Delta(G)$. Por outro lado, o grau mínimo é o grau mínimo dos seus vértices e é denotado por $\delta(G)$. Um subconjunto de vértices S é o chamado de k-plex se o grau mínimo no sub-grafo induzido for $\delta(G[S])\geq |S|$-k. Na figura 3, à direita, o grafo com 6 vértices, $|S|= 6$, e do grau de vértices 1, 3, 4 e 5 não exceda o valor 3. Assim, o grau mínimo, no sub-grafo induzido $\delta(G[S])= 3$.

3. OAlgoritmo de duas-fases

As medidas utilizadas na análise de Redes Complexas e "Graph Mining" são baseadas em procedimentos de baixa complexidade computacional, como o diâmetro do grafo, o grau de distribuição dos nós e a verificação da conectividade, subestimando o conhecimento da estrutura das componentes do grafo. Neste trabalho, propomos uma medida baseada na cobertura do grafo com k-cliques, com vista a compreender melhor a estrutura do grafo, combinando os conceitos dos "pequenos mundos" com a análise de subgrupo coesos.

Com o fim de encontrar a cobertura mínima de k-cliques é proposto um algoritmo de duas fases. Em primeiro lugar, todas as k-cliques maximais do grafo são geradas. Em segundo lugar, o subconjunto mínimo do k-cliques é escolhido para cobrir todos os vértices do grafo.

Para encontrar todas as k-cliques maximais do grafo, usamos uma simples transformação do grafo, de tal maneira que podemos reutilizar um algoritmo já estudado, o algoritmo da máxima clique.

Procedimento 1: Algoritmo de duas-fases para encontrar a cobertura de k-cliques
Entrada: distância k e grafo G
Saída: cobertura de k-cliques
1. Encontrar todas as k-clique maximais do grafo G
1.1. Transformação do grafo num k-grafo (Proc. 2)
1.2. Aplicar o algoritmo da clique máxima (Proc. 3)
2. Encontre a cobertura mínima de G com k-cliques
2.1. Aplicar o algoritmo da cobertura de conjuntos

3.1. Encontrar as k-cliques maximais no grafo G

Esta secção detalha o processo de transformação do grafo e a geração das cliques maximais.

3.1.1. Grafo Transformação

Vamos chamar de grafo k-G(V, E) à transformação do grafo G(V, E) num grafo, tal que para cada $i,j\in V$, a distância $d(i,j)\leq k$.

Para criar um grafo k-G(V,E), iremos reutilizar o Algoritmo Floyd [10], que considera todos os caminhos mais curtos num grafo, e, em seguida, para cada aresta inferior ou igual a k, será criada uma nova aresta no novo grafo.

Procedimento 2: Transformação do Grafo
Entrada: k, M[n,n] matriz de adjacência do grafo G
Saída: D[n,n] a matriz com k-distâncias do grafo G (ou k-Grafo)
1. D=M
2. Para cada (h,i,j) D[i,j] = min(D[i,j], D[i,h]+D[h,j])
3. Para cada (i,j) se (D[i,j] <=k) D[i,j]=1 senão D[i,j]=0
4. Retornar D;

3.1.2. Algoritmo da Clique Máxima
O problema da Clique Máxima é um problema NP-difícil que visa encontrar o maior sub-grafo completo em um determinado grafo. Nesta abordagem, temos a intenção de encontrar um limite inferior para o problema de maximização, com base na heurística proposta por Johnson [13] e na meta-heurística, que utiliza a Procura Tabu, desenvolvida por Soriano e Gendreau [20]. Parte do trabalho descrito nesta secção também pode ser encontrado em [4] e [3].

Vamos definir A(S) como o conjunto de vértices que estão adjacentes aos vértices da solução corrente S. Seja n=|S| a cardinalidade da clique S e $A^k(S)$ o subconjunto de vértices com k arcos incidentes em S. A(S) pode ser dividido em subconjuntos $A(S)= \cup\ A^k(S)$, k=1,..,n. A cardinalidade do conjunto de vértices |V| é igual à soma dos vértices adjacentes A(S) mais os vértices não adjacentes $A^0(S)$, resultando em $|V|= \Sigma|A^k(S)|+n$, k= 0,.., n.

Para uma determinada solução S, definimos uma estrutura de vizinhança N(S), se esta gera uma solução admissível S'. Neste trabalho vamos utilizar três estruturas de vizinhança.

Vamos considerar a seguinte notação:
$N^+ (S) = \{S': S'= S \cup\{v^i\}, v^i\in A^n(S)\}$
$N^- (S) = \{S': S'= S \setminus\{v^i\}, v^i\in S\}$
$N^0 (S) = \{S': S'= S \cup\{v^i\}\setminus\{v^k\}, v^i\in A^{n-1}(S), v^k\in S\}$
S – a solução corrente
S*– a clique maximal de maior cardinalidade
T– a lista tabu
N(S) – as estruturas de vizinhança

Procedimento 3: Heurística Tabu para o Problema da Clique Máxima
Entrada: k-Grafo, S
Saída: clique S*
1. iniciar T; S*=S;
2. enquanto não for condição de fim
2.1. se ($N^+(S)\setminus T \neq$ null) escolha a melhor S’
2.2. senão se ($N^0(S)\setminus T \neq$ null) escolha a melhor S’; actualizar T
2.2.1. senão escolha a melhor S’ em $N^-(S)$; actualizar T
2.3. actualizar S=S’
2.4. se (|S|>|S*|) S*=S;
3. fim ciclo enquanto;
4. retornar S*;

Encontrar uma clique maximal num k-grafo é o mesmo que encontrar uma k-clique maximal num grafo. Para gerar um grande conjunto de k-cliques maximais, um algoritmo multi-partida é utilizado, que chama a Heurística Tabu para resolver o Problema Clique Máxima.

3.2. Cobertura Mínima com k-cliques

Os dados de entrada para o algoritmo de cobertura com k-cliques é uma matriz onde as linhas correspondem aos vértices do grafo e cada coluna é uma k-clique que abrange um determinado número de vértices. Uma heurística para a cobertura com cliques foi proposta por Kellerman [14] e melhorada por Chvatal [5].

Na heurística construtiva em cada iteração, escolhe uma linha a ser coberta, em seguida escolhe a melhor coluna que cobre a linha, actualização a solução S e os restantes vértices R. A linha escolhida é, normalmente, a linha que é mais difícil de cobrir, ou seja, a linha a que correspondem menos colunas. Depois de concluir a cobertura do grafo, o segundo passo é eliminar a redundância; começando por ordenar a cobertura por ordem decrescente de custo, e de seguida verificar se cada k-clique é realmente essencial.

Esta heurística construtiva pode ser melhorada utilizando uma heurística de Procura Tabu que remove as colunas mais caras e reconstrói uma nova solução, tal como apresentado em [11].

O problema de partição de um conjunto é muito semelhante ao problema de cobertura de um conjunto. A partição de um conjunto é uma divisão sem sobreposição de peças do conjunto. Por outro lado, a cobertura de um conjunto permite sobreposições ou sobre-coberturas. Em ambos os problemas sub-coberturas não são permitidos.

4. Resultados Computacionais

Para a implementação computacional do algoritmo algumas escolhas têm de ser feitas, como o ambiente computacional, os ficheiros dos grafos e as medidas de “graph mining”.

Os programas foram escritos em linguagem C e o compilador Dev-C++ foi utilizado. Os resultados computacionais foram obtidos a partir de um processador 2.53GHz Intel Core 2Duo com 4,00 GB de memória principal, a funcionar sob o sistema operativo Windows Vista.

Tabela 1. Sequência de cobertura de k-clique

grafo	nr nós	diâmetro	cardinalidade da cobertura de k-cliques									
			k=1	k=2	k=3	k=4	k=5	k=6	k=7	k=9	k=18	k=40
Teste	18	6	7	4	3	2	2	1	--	--	--	--
Erdos-97-1	472	6	9	8	7	7	4	1	--	--	--	--
Erdos-98-1	485	7	8	8	7	5	1	1	1	--	--	--
Erdos-99-1	492	7	8	8	7	7	1	1	1	--	--	--
brock200_1	200	2	24	1	--	--	--	--	--	--	--	--
brock200_2	200	2	26	1	--	--	--	--	--	--	--	--
brock400_1	400	2	26	1	--	--	--	--	--	--	--	--
brock400_2	400	2	23	1	--	--	--	--	--	--	--	--
c-fat200-1	200	18	16	11	9	8	7	7	6	5	1	--
c-fat200-2	200	9	9	7	5	4	3	3	3	1	--	--
c-fat500-1	500	40	16	12	9	7	7	6	6	4	3	1

Para validar o algoritmo de duas-fases, dois grupos de conjuntos de grafos foram utilizados, os grafos Erdos e alguns ficheiros do concurso da clique DIMACS [7]. Nos grafos Erdos cada nó corresponde a um investigador, e dois nós são adjacentes se os investigadores publicaram juntos. Os grafos são denominadas "Erdos-xy", onde "x" representa os dois últimos dígitos do ano em que foi criado o grafo, e "y" é a distância máxima de Erdos para cada vértice no grafo. O segundo grupo de grafos inclui algumas instâncias de cliques DIMACS. Este grupo contém exemplos da família de grafos "brock" e da família de grafos “c-fat”.

Neste trabalho, propomos uma medida de “graph mining” baseada em k-cliques, para aprofunda o conhecimento na estrutura do grafo, combinando os “pequenos mundos” com a análise de subgrupo coesos. A medida para os “pequenos mundos” utiliza o diâmetro do grafo, ou seja, o maior caminho mais curto no grafo, como já foi dito. Para os subgrupos coesos estudo, usamos o conceito de k-clique.

Para a análise de cada grafo, consideramos o número de nós, o diâmetro e a cardinalidade do conjunto de k-cliques que cobrem todos os nós, variando k de 1 até ao valor do diâmetro. O "teste" grafo com 18 nós, tem o mais longo caminho igual a 6, para cobrir o grafo sete 1-clique são necessárias, ou quatro 2-cliques são necessárias, e assim por diante, até que uma 6-clique é necessária. Os grafos Erdos-98-1 e Erdos-99-1, com diâmetro 7, são cobertas apenas com um 5-clique. Estes valores exemplificam a diferença entre o k-clique e o k-clã; estes grafos são de 5-clique, mas não 5-clãs porque o diâmetro é igual a sete. O grafos "brock", conhecidos como esconder grafos, têm diâmetro igual a 2 e uma cobertura de um 2-clique é suficiente. A maior parte dos casos DIMACS apresentar este perfil. Por outro lado, os grafos "c-fat", têm diâmetros maiores que 7, gerando longas sequências de coberturas de k-cliques. Na medida proposta, a sequência de cobertura de k-cliques, identifica famílias de grafos, e parece ser bastante promissora.

5. Conclusões

Dada a grande quantidade de dados, fornecidos pela Web 2.0, existe uma necessidade premente de criar novas medidas para melhor compreender a estrutura das redes, o modo como os seus componentes estão organizados e o modo como evoluem ao longo do tempo.

As medidas na análise de Redes Complexas são essencialmente baseadas em procedimentos de baixa complexidade computacional, como o diâmetro do grafo, a distribuição do grau dos nós e a verificação da conectividade, subestimando o conhecimento da estrutura das componentes do grafo.

Neste trabalho o conceito de clique relaxada é alargado a todo o grafo, para obter uma visão geral, através da cobertura da rede com k-cliques. A sequência de cobertura de k-clique é apresentada, combinando conceitos dos “pequenos mundos” com estrutura das componentes coesas. A análise das sequências identifica diferentes tipos de grafos, mostrando que famílias dos grafos têm estruturas diferentes.

Existem ainda um conjunto de características, não mencionadas no presente documento, mas que podem ser obtidas, como a sobre-coberturas dos nós, a cardinalidade das k-cliques e a composição das k-cliques.

Os trabalhos com redes sociais não excedem as dezenas de nós. Neste trabalho, a proposta do algoritmo de duas-fases trabalha com grafos com centenas de nós, com um tempo de execução de alguns segundos. Em trabalhos futuros, gostaríamos de alargar esta abordagem ao maior de dados, com milhares de nós.

Referências

1. Alba, R. D.: A graph-theoretic definition of a sociometric clique, Journal of Mathematical Sociology, 3, 113-126 (1973)
2. Berners-Lee, T.: The Next Wave of the Web: Plenary Panel, 15th International World Wide Web Conference, WWW2006, Edinburgh, Scotland (2006)
3. Cavique L., C. Luz: A Heuristic for the Stability Number of a Graph based on Convex Quadratic Programming and Tabu Search, special issue of the Journal of Mathematical Sciences, Aveiro Seminar on Control Optimization and Graph Theory, Second Series (to appear 2009)
4. Cavique L., C. Rego and I. Themido: A Scatter Search Algorithm for the Maximum Clique Problem, In: Essays and Surveys in Metaheuristics, C. Ribeiro e P. Hansen (Eds), Kluwer Academic Publishers, 227-244 (2002),
5. Chvatal V.: A greedy heuristic for the set-covering problem, Math. Oper. Res. 4, 233–235 (1979)
6. Cook D.J., L.B. Holder, Editors: Mining Graph Data, John Wiley & Sons, New Jersey (2007)
7. DIMACS: Maximum clique, graph coloring, and satisfiability, Second DIMACS implementation challenge, URL http://dimacs.rutgers.edu/Challenges/, accessed April 2009 (1995)
8. Erdos, P., Renyi, A.: On Random Graphs. I., Publicationes Mathematicae 6, 290–297 (1959)
9. Faloutsos M., P. Faloutsos, C. Faloutsos: On power-law relationships of the Internet topology, In: SIGCOMM, 251–262 (1999)
10. Floyd, Robert W.: Algorithm 97: Shortest Path, Communications of the ACM, 5(6), 345 (1962)
11. Gomes M., L. Cavique, I. Themido: The Crew Time Tabling Problem: an extension of the Crew Scheduling Problem, Annals of Operations Research, volume Optimization in transportation 144(1), 111-132 (2006)
12. Grossman J., P. Ion, R. D. Castro: The Erdos number Project, URL http://www.oakland.edu/enp/, accessed April 2009 (2007)
13. Johnson D.S.: Approximation algorithms for combinatorial problems, Journal of Computer and System Science, 9, 256-278 (1974)
14. Kellerman E.: Determination of keyword conflict, IBM Technical Disclosure Bulletin, 16(2), 544–546 (1973)
15. Luce, R. D.: Connectivity and generalized cliques in sociometric group structure, Psychometrika, 15, 159-190 (1950)
16. Milgram, S.: The Small World Problem, Psychology Today, 1(1), 60-67 (1967)
17. Mokken, R. J.: Cliques, clubs and clans, Quality and Quantity, 13, 161-173 (1979)
18. Moreno, J. L.: Who Shall Survive?, Nervous and Mental Disease Publishing Company, Washington DC (1934)
19. Scott J.: Social Network Analysis - A Handbook, Sage Publications, London (2000)
20. Soriano P., Gendreau M.: Tabu search algorithms for the maximum clique, In: Clique, Coloring and Satisfiability, Second Implementation Challenge DIMACS, Johnson D.S., Trick M.A. (Eds.), 221-242 (1996)
21. Wasserman, S., K. Faust: Social Network Analysis: Methods and Applications, Cambridge University Press (1994)
22. Watts, D.J., Strogatz, S.H.: Collective dynamics of small-world networks, Nature 393(6684), 409–10 (1998)