



DETECÇÃO DE COMUNIDADES NUMA REDE DE RELACIONAMENTO DE ALUNOS COMMUNITY DETECTION PROBLEM FOR AN ACADEMIC NETWORK

Leandro Farias Maia^a; Heytor Bruno Nobre^a; Claudia M. Justel^a; Camila C.F. Oliveira^b

^a Instituto Militar de Engenharia (IME) - Rio de Janeiro, RJ, Brasil – Departamento de Engenharia de Computação

^b Universidade Federal do Rio de Janeiro (COPPE-UFRJ) - Rio de Janeiro, RJ, Brasil – Departamento de Engenharia de Produção

Resumo

Este trabalho consiste em uma aplicação do problema de detecção de comunidades em redes sociais num grafo de relacionamento de alunos de uma instituição de ensino superior. Executamos um experimento utilizando dados reais obtidos a partir de informações de alunos do segundo ano do curso básico da instituição. Foram modelados dois grafos de relacionamento a partir informações fornecidas pelos alunos durante o ano 2011. Com esses grafos, analisamos a formação de grupos utilizando o algoritmo de Clauset, Newman e Moore (CNM). Posteriormente, comparamos os grupos obtidos pelo algoritmo com os grupos reais formados a partir da escolha de especialidade de engenharia, que determina a distribuição dos alunos nas distintas especialidades oferecidas a partir do terceiro ano do curso. Para automatizar o experimento, foi criado um site que permite gerar grafos a partir de informações reais inseridas via um formulário de opinião. O grafo produzido foi criado respeitando as condições da entrada do algoritmo de CNM. Finalmente, são apresentadas as conclusões sobre a modelagem do problema e os resultados obtidos.

Palavras-chave: grafos, comunidades, algoritmos.

Abstract

This work is an application of the community detection problem in social networks for a real relationship graph. We observe the relations between students of a higher education institution and perform an experiment using real data. Here we model two graphs with information provided by the students during 2011. We used the Clauset, Newman and Moore algorithm (CNM) to obtain the communities. After that, we compared the groups obtained by the algorithm with the real groups created by the choice of engineering specialty offered at the institution in the beginning of the third year of the course. In order to execute the experiment, a website was created to help in the construction of the graph. The graph produced with the website was created obeying the input conditions of the CNM algorithm. Finally, some conclusions about the modeling of the problem and the results obtained are suggested.

Keywords: graphs, communities, algorithms.

1. INTRODUÇÃO

As redes sociais podem ser constituídas a partir das relações interpessoais entre um grupo de indivíduos e o estudo dos grupos pode mostrar como estes afetam as decisões e o comportamento de um indivíduo dentro dele. A detecção de comunidades por algoritmos é uma das linhas de pesquisa em redes (Fortunato, 2010). Do ponto de vista da modelagem do problema, uma das grandes dificuldades é identificar a melhor representação da relação entre os indivíduos a ser descrita pelas arestas no grafo, ou seja, qual aspecto dos relacionamentos interpessoais deve

ser considerado para encontrar resultados significativos na detecção de comunidades e o que essas comunidades detectadas irão representar.

O problema da detecção de comunidades recai na divisão da rede em partições do conjunto de vértices densamente conectados. Existem vários tipos de algoritmos de detecção de comunidades: os algoritmos divisivos detectam arestas intercomunitárias e as removem (Girvan *et al.*, 2002); os algoritmos aglomerativos unem nós/comunidades recursivamente e os métodos de otimização que são baseados na maximização de uma função objetivo adequada (Clauset *et al.*, 2004; Blondel *et al.*, 2008).



Diferentes abordagens foram propostas para determinar comunidades, cada uma das quais utiliza uma função específica para descrever as características dos grupos a serem construídos. Portanto, escolher o algoritmo mais apropriado para cada tipo de problema determinará a corretude e maior exatidão possível dos resultados obtidos.

Um estudo prévio envolvendo algoritmos de particionamento de grafos em comunidades, sendo estes redes sociais, pode ser encontrado em (Barbosa *et al.*, 2011).

O objetivo desse trabalho é aplicar um algoritmo de detecção de comunidades em uma rede de relacionamentos de alunos baseada em dados reais, para estudar seu comportamento e como a formação de diferentes grupos dentro de uma instituição de ensino superior é determinada por essa estrutura. Com uma análise dos resultados obtidos após a aplicação dos algoritmos, espera-se obter uma previsão do comportamento dos membros da rede em uma situação real.

O trabalho está organizado da maneira seguinte. Na Seção 2, são apresentadas as definições básicas para o entendimento do trabalho. A Seção 3 apresenta a descrição do experimento realizado. Os resultados obtidos são apresentados na Seção 4, assim como a descrição do site usado para gerar os grafos. Finalmente, na Seção 5, são relatadas as conclusões do trabalho e, na Seção 6, a bibliografia utilizada.

2. GRAFOS, REDES SOCIAIS E COMUNIDADES

Um **grafo não orientado** é uma estrutura $G = (V, E)$ formada por um conjunto finito e não vazio de vértices, $V = \{v_1, \dots, v_n\}$, e um conjunto de pares não ordenados de vértices denominados arestas, $E = \{(v_i, v_j), \text{ em que } v_i, v_j \in V, 1 \leq i, j \leq n\}$. Consideramos o grafo sem laços e sem arestas múltiplas.

O **grau de um vértice**, $grau(v)$, é igual à quantidade de arestas cuja extremidade é o vértice v .

Um **corte** num grafo $G = (V, E)$ é o particionamento do conjunto de vértices em dois conjuntos disjuntos, S e $V - S$. Denota-se o corte pelo par $(S, V - S)$.

Um **grafo completo** é um grafo em que existe uma aresta entre quaisquer pares de vértices.

Um **k-clique** em um grafo é um subconjunto de k vértices do grafo que induz um subgrafo completo.

Outras informações sobre grafos podem ser obtidas em (Szwarcfiter, 1988).

Uma **rede** pode ser representada por um grafo no qual os vértices representam os objetos e as arestas as conexões entre os objetos. Existem exemplos de redes em diversos domínios, como por exemplo: redes de comunicação, redes sociais, redes biológicas, redes de informação. Uma **rede social**, entretanto, pode ser entendida como uma rede na qual o conjunto de objetos é formado por pessoas e as conexões entre elas são relações sociais. Entende-se por **comunidade** um grupo de objetos que possuem propriedades comuns e/ou desempenham uma função similar dentro de uma rede.

A ideia de comunidades em uma rede pode ser formulada como um particionamento de grafos (*graph clustering*). Clauset, Newman e Moore propuseram um algoritmo para detecção de comunidades (denotado como CNM) e apresentado em Clauset *et al.* (2004). A complexidade deste algoritmo é $O(m \cdot \log^2(n))$, em que n e m são, respectivamente, o número de vértices e arestas do grafo que representa a rede. Em (Kwak *et al.*, 2009) pode ser achada uma descrição em pseudo-linguagem deste algoritmo.

O **algoritmo CNM** procura um particionamento do grafo em comunidades de maneira tal que seja maximizado um parâmetro chamado modularidade. O conceito de **modularidade** (Q) representa a qualidade de um particionamento de um grafo em comunidades. Dado um grafo $G = (V, E)$ com $|V| = n$, $|E| = m$, definimos um particionamento de acordo com a Equação 1.

$$\mathcal{C} = \cup C_i \quad (1)$$

Em que,

C_i : subconjunto de V , com $C_i \neq \emptyset$ e $C_i \cap C_j = \emptyset, i \neq j$.

Deste modo, define-se a modularidade pela Equação 2.

$$Q(\mathcal{C}) = \sum_i (e_{i,i} - a_i^2) \quad (2)$$

Em que,

$e_{i,i}$: quociente entre o número de arestas entre vértices pertencentes à comunidade i e o número de arestas do grafo (m)

a_i : o quociente entre o número de todas as arestas que atravessam a fronteira da comunidade i e o número total de arestas do grafo (m).

Considerando todas as possíveis partições C , o valor da modularidade, $Q(C)$, varia de -1 a 1. O valor $Q = 0$ implica que o número de arestas dentro de uma comunidade não é melhor do que um número obtido aleatoriamente. O valor $Q = 1$ é o máximo possível, mas, na prática, nunca é atingido.



A abordagem utilizada pelo algoritmo CNM para obter o resultado desejado é um algoritmo guloso. A partir de um conjunto de n comunidades unitárias, estas são unidas iterativamente tentando aumentar o valor da modularidade em cada iteração.

Outros conceitos sobre redes complexas podem ser achados nos livros de Kleinberg *et al.* (2010) e Newman (2010).

3. DESCRIÇÃO DO EXPERIMENTO

Nesta seção, será apresentado o método usado para gerar os grafos de relacionamento baseados em dados reais de alunos do segundo ano de 2011 do Instituto Militar de Engenharia. Posteriormente, apresentamos dois grafos correspondentes a duas redes de relacionamento. O primeiro deles, denominado **Grafo Questionário**, foi criado com base em um formulário de opinião formado por perguntas relativas aos relacionamentos entre alunos. O segundo grafo, denominado **Grafo Seção**, foi criado a partir de dados da ficha de escolha de especialidade de engenharia preenchida pelos alunos no final do segundo ano do curso.

Após a criação dos grafos, foi utilizada uma implementação do algoritmo CNM para detectar as comunidades, extraída da biblioteca de programas desenvolvida pelo grupo de Análise de Redes Sociais da Universidade de Stanford (*SNAP*). O algoritmo CNM foi escolhido devido a sua ampla utilização no contexto de redes sociais e pelo fato de ser uma solução aproximada ao problema de maximização da modularidade que é NP-Completo (Brandes *et al.*, 2008). Finalmente, foi utilizado o software *GRAPHVIZ* para visualizar os grafos e as comunidades obtidas. As subseções 3.1 e 3.2 descrevem como foram criados cada um dos grafos.

3.1. O Grafo Questionário

Visando construir uma rede com base nos relacionamentos existentes entre os alunos do 2º ano de 2011 do curso de graduação do Instituto Militar de Engenharia (IME), foi aplicado um formulário de opinião formado por duas questões. Uma delas com o objetivo de identificar as relações de amizade existentes entre os alunos. A segunda questão do formulário tinha como objetivo destacar as três disciplinas preferidas dos alunos entrevistados dentre as disciplinas oferecidas nos dois primeiros anos do curso.

Deste modo, foi criado um grafo em que os vértices representam alunos e grupos de disciplinas, e as arestas representam relações de amizade entre alunos e de preferência dos mesmos pelas disciplinas. Observando que, se duas ou três disciplinas escolhidas por um aluno pertencerem ao mesmo grupo, o grafo terá uma única aresta. Os vértices foram numerados de 1 até 105 para os alunos e

os números 300, 302, 303, 304, 305 e 306 para grupos de disciplinas definidos por seções de ensino (departamentos) que oferecem as especialidades de engenharia existente no IME e ministram as disciplinas.

Esta modelagem foi obtida após algumas tentativas, nas quais as comunidades obtidas pelo algoritmo CNM eram muito dispersas. Inicialmente, o formulário de opinião continha três questões, as duas mencionadas anteriormente mais uma questão sobre afinidade entre alunos, que foi eliminada. A modelagem do grafo com um vértice para cada disciplina produziu um grafo muito esparso, portanto foi adotado um critério de agrupação de disciplinas. Após diferentes tentativas, as disciplinas foram agrupadas segundo mostra a Tabela 1.

Tabela 1. Grupos de disciplinas/vértices do Grafo Questionário

Grupo de disciplinas	Vértice
Álgebra Linear Cálculo Proba. e Estatística	300
Introdução à Computação Cálculo Numérico	303
Eletricidade Técnica Física Laboratório de Física	302
Resistência dos Materiais Mecânica Desenho Técnico	304
Intr. Ciência dos Materiais Química	305
Projeto de Inglês Redação Científica	306

Fonte: Maia *et al.* (2012)

O Grafo Questionário foi elaborado a partir da ferramenta desenvolvida para melhorar a transparência do processo de construção do grafo. Esta ferramenta será detalhada na próxima seção. Foram utilizadas duas perguntas do formulário de opinião: uma sobre relacionamento de alunos e outra sobre disciplinas de maior identificação dentre as disciplinas do curso básico (2 primeiros anos do curso de engenharia).

A Figura 1 apresenta o **Grafo Questionário**. Após executar o algoritmo CNM, foram identificadas 7 comunidades que são mostradas na mesma figura. Os vértices com a mesma cor pertencem à mesma comunidade obtida pelo algoritmo CNM.

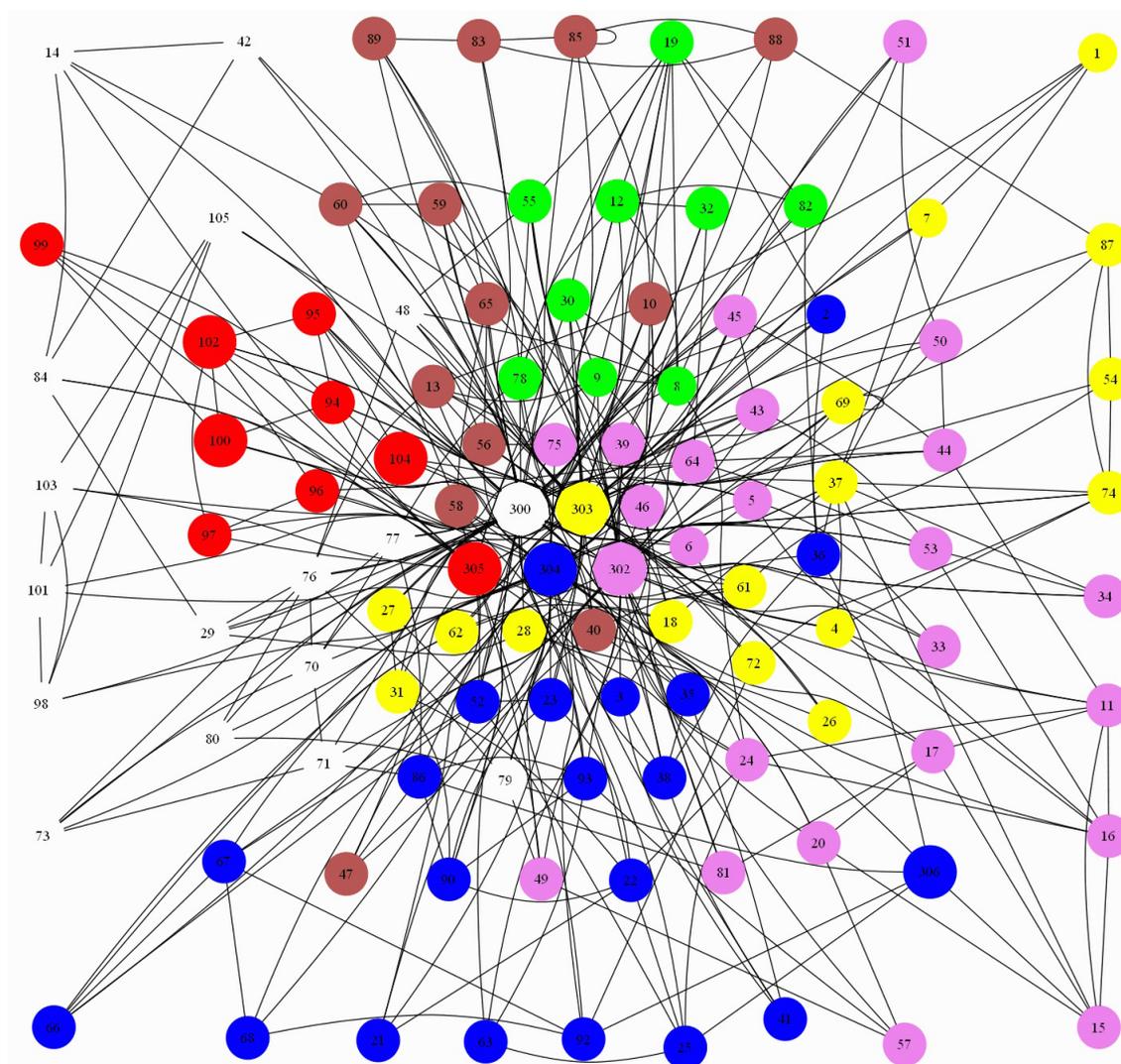


Figura 1. Grafo Questionário com as 7 comunidades definidas pelo Algoritmo CNM.

Fonte: Maia et Nobre (2012)

3.2. O Grafo Seção

Buscando estabelecer uma nova agrupação dos os alunos do segundo ano do curso de graduação de 2011, foi criado o Grafo Seção. Sua construção foi baseada na ficha de escolha de especialidade, preenchida pelos mesmos alunos ao final do segundo semestre letivo do segundo ano. Nessa ficha, cada aluno ordena as especialidades de engenharia de acordo com a sua preferência, considerando as 10 especialidades oferecidas no IME. No início do terceiro ano do curso, os alunos são distribuídos nas especialidades, utilizando a ficha de escolha preenchida pelos alunos e alguns critérios de desempenho e vagas por especialidade.

Durante a montagem do Grafo Seção, surgiram alguns problemas relacionados com a coleta de dados. Primeiro, foi observado que alguns alunos não preencheram o formulário de opinião. Portanto, o número de vértices representando alunos do **Grafo Seção** não foi considerado

igual ao número de vértices representando alunos do **Grafo Questionário**. Além disso, foi realizada uma adaptação para não comprometer a comparação entre as comunidades existentes com as determinadas pela aplicação do algoritmo CNM para o Grafo Questionário: as especialidades de engenharia foram agrupadas segundo as seções de ensino (ou departamentos) que oferecem as mesmas, como mostra a Tabela 2.

Cada vértice do Grafo Seção corresponde a um aluno ou a um grupo formado por algumas das dez especialidades existentes no curso de engenharia do IME. As arestas são ligações entre alunos e especialidades, em que cada aluno está conectado a cada uma das três primeiras especialidades que especificou na ficha descrita no parágrafo anterior. Observando que, se as 2 ou 3 das primeiras especialidades escolhidas por um aluno pertencerem à mesma seção, o grafo terá uma única aresta.



Tabela 2. Grupos de especialidades/vértices do Grafo Seção

Especialidade	Seção	Vértice
Construção	SE/2	106
Eletricidade Eletrônica Comunicações	SE/3	107
Materiais Mecânica aut. Mecânica arm.	SE/4	108
Química	SE/5	109
Cartografia	SE/6	110
Computação	SE/8	111

Fonte: Maia et Nobre (2012)

A Figura 2 ilustra o **Grafo Seção**, em que os vértices representam os alunos (numerados de 1 até 105) e as especialidades agrupadas segundo a Tabela 2. O resultado da execução do algoritmo CNM determinou 4 comunidades. Os vértices com a mesma cor na figura pertencem à mesma comunidade obtida pelo algoritmo CNM.

4. RESULTADOS OBTIDOS

Esta seção apresenta os resultados obtidos neste trabalho. A subseção 4.1 compara os resultados obtidos pelo algoritmo CNM para os dois grafos descritos nas subseções 3.1 e 3.2, com a divisão em comunidades segundo os grupos de alunos por especialidades no início do terceiro ano do curso (2012). A subseção 4.2 apresenta uma breve descrição da página criada para inserir os dados reais que permitem obter grafos de relacionamento.

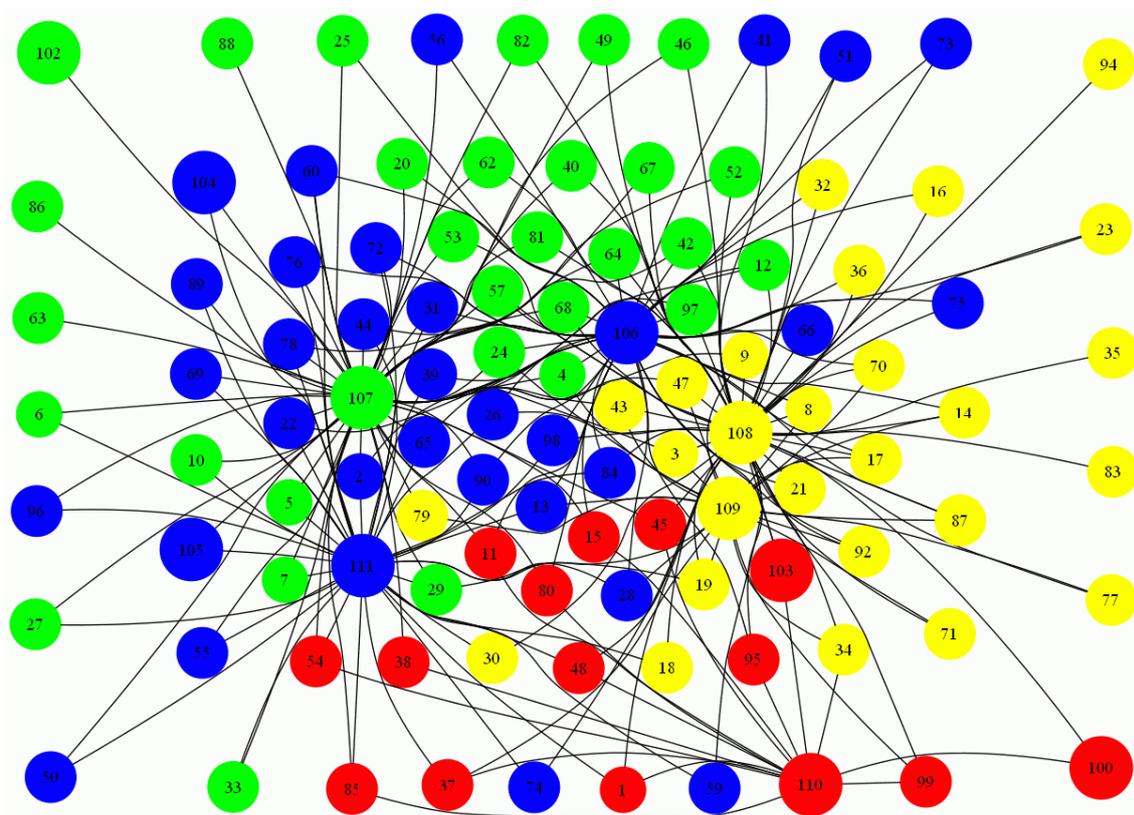


Figura 2. Grafo Seção e as 4 comunidades definidas pelo algoritmo CNM
Fonte: Maia et Nobre (2012)

4. RESULTADOS OBTIDOS

Esta seção apresenta os resultados obtidos neste trabalho. A subseção 4.1 compara os resultados obtidos pelo algoritmo CNM para os dois grafos descritos nas subseções 3.1 e 3.2, com a divisão em comunidades segundo os grupos de alunos por especialidades no início do terceiro ano do curso (2012). A subseção 4.2 apresenta uma breve descrição da página criada para inserir os dados reais que permitem obter grafos de relacionamento.

4.1. Comparação das comunidades obtidas para o Grafo Questionário e para o Grafo Seção

A fim de comparar as comunidades obtidas para os grafos apresentados na seção anterior com as turmas reais formadas no início do terceiro ano letivo, foi utilizado o Grafo Seção para colorir os vértices segundo o critério a seguir. Neste caso, **não** foi utilizado o algoritmo CNM para determinar as comunidades. Os vértices do **Grafo Seção** foram coloridos de maneira tal que alunos que ficaram na



mesma seção no início do terceiro ano, foram coloridos com a mesma cor do vértice que representa a seção. Os vértices seguem a numeração de 1 até 105 para os alunos e de 106 até 111 para as 6 seções de ensino.

Na Figura 3, os vértices com a cor rosa representam os alunos da SE/2 (junto com o vértice 106); a cor verde

representa os da SE/3 (e o vértice 107); a cor amarela representa os da SE/4 (e o vértice 108); a cor marrom representa os da SE/5 (e o vértice 109); a cor vermelha representa os da SE/6 (e o vértice 110); a cor azul representa os da SE/8 (e o vértice 111); e a cor branca representa os alunos que trancaram a matrícula no início do terceiro ano do curso.

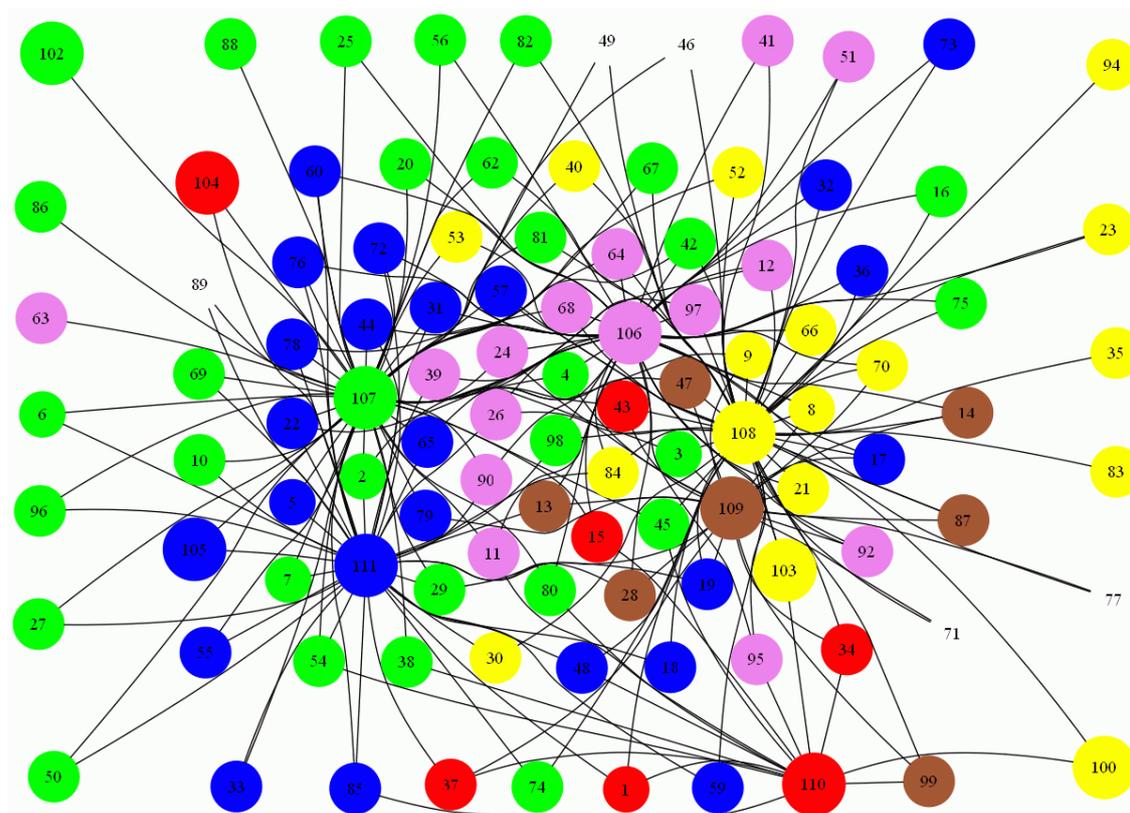


Figura 3. Grafo Seção mostrando as 6 comunidades definidas pelas escolhas reais de especialidades, no início terceiro ano do curso.

Fonte: Maia et Nobre (2012)

Comparando os resultados reais das comunidades apresentadas no grafo da Figura 3 (6 comunidades) com as comunidades obtidas pelo algoritmo CNM, apresentadas na Figura 2 (4 comunidades), foram identificadas algumas interseções, apesar desta segunda divisão em comunidades possuir menos comunidades. Nas 4 comunidades obtidas pelo algoritmo CNM no grafo da Figura 2, observamos que boa parte das informações correspondentes às seções SE/2 e SE/8 formam uma única comunidade (na cor azul). O mesmo acontece com as seções SE/4 e SE/5 (na cor amarela). Além disso, o algoritmo CNM aplicado no Grafo Seção produziu uma comunidade com 50% de acertos ao respeito da comunidade correspondente à SE/3 no caso real.

Com base nestes resultados, podemos concluir que a aplicação do algoritmo CNM apresentou um bom resultado para o Grafo Seção. As diferenças identificadas podem ter sido geradas por falta de informações vinculadas ao

significado do grafo, como a ordem de preferência das disciplinas escolhidas por um determinado aluno.

As comunidades determinadas pelo algoritmo CNM no Grafo Questionário também foram comparadas aos dados reais de distribuição dos alunos nas especialidades no início do terceiro ano do curso. Foi realizada uma análise para identificar os vértices comuns nas comunidades obtidas nesses dois casos.

Dentre os resultados obtidos nessa comparação, podemos destacar que, das 7 comunidades detectadas pelo algoritmo CNM no Grafo Questionário, uma única comunidade não incluiu os vértices que representam disciplinas (esta comunidade aparece colorida com a cor verde na Figura 1). Além disso, em todas as comparações realizadas entre os grupos formados no início do terceiro ano letivo (Figura 3) e as comunidades obtidas pelo algoritmo CNM no Grafo



Questionário, a percentagem de vértices que representam alunos coincidentes é menor que 25% dos alunos das comunidades correspondentes aos grupos no início do terceiro ano letivo.

Com base nestes resultados, podemos concluir que a aplicação do algoritmo CNM no Grafo Questionário não apresentou bons resultados. Esses resultados pouco satisfatórios podem ter sido gerados pelo conteúdo das perguntas do formulário de opinião, que pode ser melhorado.

4.2. O Site

Após os testes realizados com dados dos alunos do segundo ano do curso básico em engenharia, percebeu-se a necessidade de automatizar a criação dos grafos. Os grafos utilizados neste trabalho contêm aproximadamente 100 nós e entre 300 e 700 arestas.

O grafo de entrada do algoritmo CNM é representado por um arquivo de texto que contém a lista de todas as arestas. Para obter uma interface que permita gerar o grafo

no formato de entrada do algoritmo CNM, foi criada uma página *html* que permite cadastrar os integrantes da turma a ser pesquisada e, posteriormente, aplicar o formulário com perguntas direcionadas aos alunos daquela turma. Esse processo de inserção de dados é dividido em duas fases. Na primeira, o administrador cadastra a turma de alunos. Cada aluno será representado por um vértice no grafo de relacionamentos. Depois disso, o administrador pode convocar os alunos a responderem um formulário previamente criado. Nessa fase, o aluno escolhe seu nome dentre a lista de nomes de alunos cadastrados e responde ao formulário (de afinidade, ou a lista de disciplinas (ou grupos de disciplinas), ou a lista de especialidades pretendidas) e a página adiciona as arestas no grafo correspondente. Como resultado deste processo, é construído um arquivo em formato *txt* com as informações dos vértices e arestas do grafo de relacionamento que será usado como entrada no algoritmo de detecção de comunidades.

A Figura 4 mostra o processo de cadastro, criação do grafo e determinação de comunidades com a posterior visualização.

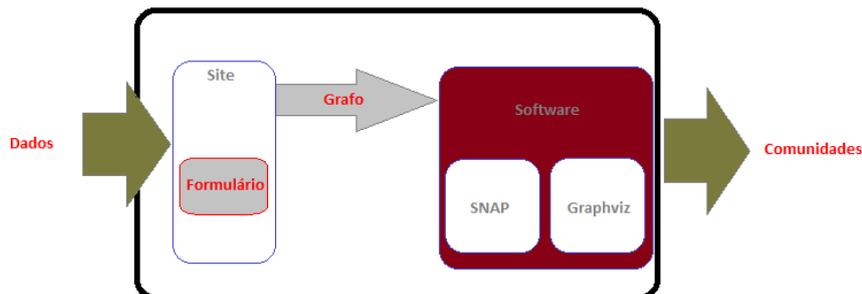


Figura 4. Processo completo de criação do grafo utilizando a ferramenta desenvolvida
Fonte: Maia et Nobre (2012)

5. CONCLUSÕES

Após uma pesquisa sobre algoritmos de determinação de comunidades, foi escolhido o algoritmo proposto por Clauset *et al.*(2004) para determinar comunidades nos grafos de relacionamento produzidos neste trabalho.

A partir da modelagem apresentada na Seção 3, foram criados dois grafos – o Grafo Questionário e o Grafo Seção – utilizando dados reais fornecidos pelos alunos através, respectivamente, de um formulário de opinião e da ficha de escolha de especialidade. O objetivo desta pesquisa foi comparar as comunidades detectadas nestes grafos por meio do algoritmo CNM, com as comunidades reais formadas pelos alunos no início do terceiro ano do curso. Essas comunidades reais são determinadas pela formação das turmas correspondentes às diferentes seções de engenharia do IME, nas quais os alunos continuam o curso durante a especialização.

A análise realizada sobre o Grafo Seção produziu resultados bastante favoráveis. Quase 50% dos vértices das comunidades determinadas pelo algoritmo representam alunos que realmente cursam a mesma especialidade no terceiro ano do curso. Portanto, os resultados obtidos com o algoritmo CNM para o Grafo Seção podem ser usados como uma estimativa das preferências dos alunos pelas especialidades. Dessa forma, o número de vagas para cada especialidade e a alocação de salas de aula poderiam ser planejadas antecipadamente usando a informação das comunidades fornecida pelo algoritmo.

No caso do Grafo Seção, a modelagem ainda pode ser modificada. Adicionar pesos nas arestas permitirá incluir a prioridade de escolha do aluno pela especialidade. Uma variante do problema seria considerar um limite superior para o número de vértices nas comunidades.



Já em relação ao Grafo Questionário, o processo foi mais complexo. Foi necessária uma reestruturação do mesmo ao longo dos trabalhos, gerando diversas versões. Por isso, foi considerada a necessidade da criação de uma ferramenta de geração automática do grafo, respeitando o formato de entrada do algoritmo CNM. Para tal fim, foi criado o *site* descrito na Seção 4 que facilitou o cadastro dos alunos, assim como a geração do arquivo texto que representa o grafo.

A última versão do Grafo Questionário gerado pela ferramenta automatizada produziu resultados mais condizentes do que os obtidos com as versões anteriores, quando comparados aos grupos de alunos que cursam a mesma especialidade no terceiro ano do curso. Mas ainda não são satisfatórios. Podemos observar que, como todos os alunos analisados frequentam o curso de engenharia, algumas disciplinas ministradas nos dois primeiros anos do curso que apresentam maior afinidade podem não dar boa informação sobre as preferências das especialidades. Deste modo, sugerimos como trabalhos futuros analisar a conveniência de utilizar as disciplinas no formulário de opinião, ou utilizar um conjunto reduzido delas, que seja mais apropriado. Também seria interessante verificar se a pergunta sobre relacionamento de alunos no formulário de opinião é relevante neste contexto. Portanto, sugerimos realizar novos experimentos com outras turmas para melhorar a modelagem do Grafo Questionário.

6. AGRADECIMENTOS

Este trabalho foi desenvolvido como parte do Projeto de Final de curso de graduação em Engenharia de Computação. Agradecemos aos professores integrantes da banca de avaliação pelas sugestões dadas, que permitiram melhorar a versão final do trabalho. A pesquisa foi parcialmente financiada pelo Projeto Universal – CNPq Nro. 305516/2010-8 e pelo Projeto Institucional PIBITI-CNPq.

7. BIBLIOGRAFIA

Barbosa, D. A. B. L., Avelino, L. D., Souza, R. F., Oliveira, C. C. G. F. e Justel, C. (2011), “Medidas de centralidade e detecção de comunidades em rede de co-autoria”, Anais do XLIII Simpósio Brasileiro de Pesquisa Operacional, pp. 1-10.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R. e Lefebvre, E. (2008), “Fast unfolding of communities in large networks”, *J. Stat. Mech.*, P10008.

Brandes, U., Delling, D., Gaertler, M., Gorke, R., Hofer, M., Nikoloski, Z., Wagner, D. (2008), “On Modularity Clustering”, *IEEE Transactions on Knowledge and Data Engineering*, 20(2), pp.172-188.

Clauset, A., Newman, M.E.J. e Moore, C. (2004), “Finding community structure in very large networks”, *Physical Review E*, 70:066111.

Fortunato, S. (2010) “Community detection in graphs”. *Physics Reports* 486, 75-174 (2010) arxiv 0906.0612.

Girvan, M., Newman, M.E.J. (2002), “Community structure in social and biological networks”, *Proceedings of the National Academy of Sciences of the United States of America*, 99(12), pp.7821–7826.

GRAPHVIZ , disponível em: www.graphviz.org (Acesso em: março de 2011).

Kleinberg, J et Easley, D. (2010), *Networks, crowds and markets*, Cambridge University Press.

Kwak, H., Jeong, H., Choi, Y., Moon, S. e Eom, Y.-H. (2009), “Mining communities in networks: A solution for consistency and its evaluation”. *Proceedings of the 9th ACM SIGCOMM Conference on Internet Measurement Conference (IMC’09)*, pp. 301 -314.

Maia, L.F., Nobre, H.B. (2012). *Deteção de comunidades em rede de relacionamento de alunos*, Monografia em Engenharia de Computação, Instituto Militar de Engenharia, Rio de Janeiro, RJ.

Newman, M. (2010), *Networks: an introduction*, Oxford University Press.

SNAP, disponível em: snap.stanford.edu (Acesso em: março de 2011).

Szwarcfiter, J.L. (1988), *Grafos e algoritmos computacionais*, Ed Campus.