Análise de mobilidade em comunidades de Software Livre

Eduardo F. Santos[1 2], **Fernanda Lima**[2]

[1]Lightbase Consultoria em Software Público
SCLN 309 Bl. B sala 213, Asa Norte, Brasília — DF, CEP 70.775-520

[2]CIC/UnB
Campus Universitário Darcy Ribeiro, Brasília/DF – Brasil, CEP 70910-900

eduardo.santos@lightbase.com.br,ferlima@cic.unb.br

Abstract. *In order to identify new leaderships through the concepts of centrality and prestige inside software communities, complex networks study and analysis can help us understand if the collaboration activity is capable of producing knowledge inside the community. This article proposes the usage of a mobility rate metric, calculated as the reason between the members who could actually learn through collaboration. The calculation is made by analyzing the creation of collaboration objects, and the result is exemplified applying the technique to Brazilian Public Software Portal communities' collaboration data.*

Resumo. *Dentro do estudo de comunidades de software em redes complexas, a identificação dos conceitos de centralidade e prestígio para descoberta de novas lideranças pode ajudar a entender o retorno da colaboração como instrumento para geração de conhecimento na comunidade. Este artigo propõe a utilização de uma métrica para avaliar a taxa de mobilidade, um indicativo sobre a razão dos membros que conseguiram efetivamente aprender através de troca de conhecimento. O cálculo é realizado pela análise da criação de objetos de colaboração na comunidade, exemplificado através da aplicação da técnica aos dados de colaboração do Portal do Software Público Brasileiro.*

1. Introdução

O Portal do Software Público Brasileiro é reconhecido mundialmente como uma iniciativa importante de compartilhamento e construção de conhecimento colaborativo. Um dos mais importantes grupos de estudo sobre o universo do software livre no mundo, o consórcio organizador do *Open World Forum*[1], escolheu, no ano de 2008, a iniciativa brasileira como uma das que faria parte do *FLOSS*[2] *2020 Roadmap* [Meffe and Freitas 2008].

Contudo, qual o valor real que a participação em um ambiente virtual como o do Portal traz para o cidadão? No Livro *A Riqueza Revolucionária* [Tofler and Tofler 2007][p.39], os autores citam o fenômeno da "terceira onda de riqueza" no momento em que o acúmulo de conhecimento se torna um ativo econômico

[1]O *Open World Forum* ou apenas OWF é um evento realizado anualmente em Paris que reúne as maiores autoridades do universo do Software Livre no mundo. Mais informações podem ser encontrados no Portal do evento: www.openworldforum.org

[2]Sigla para software livre e aberto como é conhecida principalmente na Europa: *Free, Libre and Open Source Software*.

importante. Se definirmos o conhecimento como o "acúmulo de técnicas para resolver problemas de determinada sociedade" e a aprendizagem como o acúmulo do conhecimento [Gaines 2013], verificar o impacto da colaboração na aprendizagem pode apontar um benefício real obtido através de um ambiente virtual. O objetivo deste trabalho é encontrar relações entre a criação de objetos de colaboração nas comunidades do Portal do Software Público Brasileiro e o acúmulo de conhecimento, propondo uma métrica que seja capaz de estimar a probabilidade de uma aprendizagem efetiva acontecer na comunidade.

2. Referencial Teórico

No Portal do Software Público Brasileiro, o Governo tenta agir como indutor da inovação, mas instrumentos de avaliação que permitam entender a complexidade dos agentes envolvidos e o retorno de cada um são necessários, levando em conta o caráter da colaboração. A mesma preocupação é apresentada pela Rede Telecentros.BR ao analisar as redes sociais de agentes de inclusão digital [Silva et al. 2013a] apontando a importância da formação de novos agentes e a necessidade de encontrar indicadores de desempenho que melhor se apliquem à realidade das redes sociais. Uma proposta de métrica é apresentada com base em indicadores da Análise de Redes Sociais, tais como centralidade, prestígio, densidade da rede e coeficiente de agrupamento [Silva et al. 2013b].

2.1. Análise de Redes Sociais

Para medir a intensidade da troca de conhecimentos e suas implicações a teoria da Análise de Redes Sociais permite mapear as relações construindo uma espécie de "árvore" indicando possíveis pontos de acumulação do conhecimento. Em [Foulonneau et al. 2013] os autores apresentam uma proposta de estudo de topologia de redes complexas [Boccaletti et al. 2006] baseados na colaboração em projetos de Software Livre. Do ponto de vista de quem adiciona conhecimento em uma contribuição iniciada por outro, a conclusão é que a estrutura da colaboração "não muda significativamente com o tempo, não constituindo um indicador válido sobre o ciclo de vida da comunidade".

Em [Durugbo et al. 2011] os autores fazem uma revisão sistemática sobre colaboração em redes complexas e propõem organizações das topologias em grupos e métricas para avaliação da colaboração em cada uma. O objetivo é "analisar como a estrutura de colaboração e seu comportamento podem ser utilizados para determinar e rever os fatores que influenciam o fluxo da informação". Já em [Ngonmang and Viennet 2013] os autores apresentam uma proposta de predição em redes complexas, tentando mostrar como a análise dos grafos pode prever o comportamento da comunidade. Apontam, contudo, que o método é capaz de isolar "fatias" do tempo, mas não possui métricas diretas para analisar as possibilidades de crescimento e morte. Para resolver a limitação identificada é necessário utilizar teorias que levem em consideração a variação das relações com o tempo.

2.2. Comportamento e Predição em Redes Complexas

O método CPM [Derényi et al. 2005], ou *k-clique Percolation Method*, fornece subsídios para organizar redes aleatórias em grafos, cujo nascimento e morte obedecem a critérios não intrinsecamente definidos. Imaginemos a função $f(t)$ onde t é o intervalo de tempo

que definimos para observar a rede em recortes. A ideia é traçar o grafo no tempo $t = 0$, caminhar uma unidade ($t = 1$) e desenhar o grafo novamente. Podemos então construir a função de autocorrelação $C(t)$ para quantificar a sobreposição relativa entre dois estados da mesma comunidade $A(t)$ em um número de intervalos de tempo t:

$C(t) \equiv \frac{|A(t_0) \bigcap A(t_0+t)|}{|A(t_0) \bigcup A(t_0+t)|}$, onde $|A(t_0) \bigcap A(t_0 + t)|$ é o número de nós em comum entre $A(t_0)$ e $A(t_0 + t)$; $|A(t_0) \bigcup A(t_0 + t)|$ é o número de nós na união entre $A(t_0)$ e $A(t_0 + t)$. Em redes muito dinâmicas o coeficiente decai mais rápido, ou seja, as relações mudam com muita frequência. Se o coeficiente muda pouco com o tempo, temos que a rede é mais estática.

3. Análise de mobilidade

A seguir são apresentados dados gerais sobre o Portal do Software Público Brasileiro e suas comunidades extraídos no dia **02 de fevereiro de 2011**:

- **50** comunidades de software;
- **96678** usuários válidos;
- **33032** mensagens de fórum;
- **15536** páginas Wiki.

3.1. Dados sobre os objetos de colaboração

Objeto é, para fins de pesquisa, a inserção de qualquer dado no banco, podendo significar conteúdos interativos ou não. Contudo, como o objetivo é analisar a colaboração utilizaremos o conceito de "redes onde grupos/indivíduos são conectados" [Durugbo et al. 2011], e as comunidades serão analisados na lógica das redes complexas [Boccaletti et al. 2006]. A primeira extração de objetos do banco de dados traz os dados da Tabela 3.1.

n_objects	object_type
300653	content_item
219005	dotlrn_member_rel
217264	notification_request
188360	content_folder
130740	as_item_data
127134	as_session_results
93420	calendar
92469	dotlrn_student_profile_rel
28992	forums_message
24351	as_section_data
11990	events_registration
11277	::xowiki::Page
8886	file_storage_object
	(...)
	(114 registros)

Tabela 1. Lista de objetos mais criados

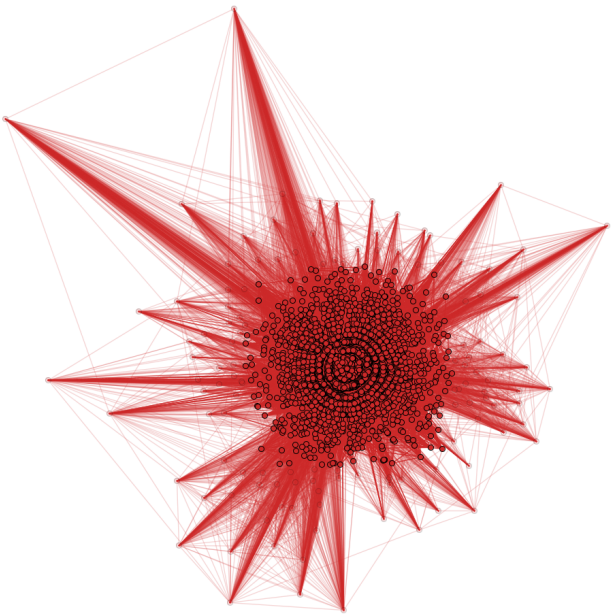

Figura 1. Grafo de colaboração nos fóruns

O framework OpenACS utilizado no Portal do Software Público utiliza um modelo de dados orientado a objetos [Hernández et al. 2005], classificando-os de acordo

com o tipo. Na lista de objetos da Tabela 3.1 sabemos que os objetos que permitem a análise de interação são mensagens de fórum (`forums_messages`) e as páginas Wiki (`::Xowiki::Page`), sendo que o primeiro representa mais que o dobro do segundo.

3.2. Grafos de colaboração

A Figura 1 mostra o grafo de colaborações relativo à troca de mensagens de fórum dentro do Portal do Software Público traçado com auxílio da ferramenta PExGraph [Paulovich et al. 2007]. Selecionamos para a amostra as comunidades que possuem participação mais significativa no Portal. É possível notar a presença de pequenos pontos mais escuros em forma de círculo organizados ao redor do gráfico, elementos que centralizam as colaborações. Ao detalhar a presença de cada um dos nós identificamos tratar-se das figuras dos Coordenadores de Comunidade, pessoas que têm relação formal com o software. Ainda que não sejam remunerados para realizar a troca de mensagens, responder as dúvidas é uma atribuição informal.

3.3. Cálculo do coeficiente de autocorrelação

O dado mais esclarecedor sobre o comportamento das comunidades é o coeficiente de autocorrelação, apresentado na Figura 2. De maneira geral os valores são bastante baixos, o que significa que entre dois períodos de tempo consecutivos as relações mudam bastante. As comunidades do Portal do Software Público Brasileiro são todas voltadas para o compartilhamento de software. O procedimento padrão é o cadastro seguido do *download* da ferramenta desejada e consequente entrada na comunidade. Para obter o software os usuários devem necessariamente entrar primeiro no grupo desejado, pois só assim terão permissão de acesso ao objeto. Em seguida, normalmente acontece a instalação/configuração inicial, fase em que o surgimento de dúvidas é natural. Assim o fórum passa a ser o caminho pelo qual passam os usuários menos experientes.

Uma vez solucionados os problemas de instalação/configuração, a tendência natural é que os usuários parem de acessar regularmente o mecanismo, retornando somente no caso de surgirem outras dúvidas. Exceção à regra é feita para os usuários mais avançados, que podem ser classificados como desenvolvedores. Em pesquisa realizada no ano de 2008 com os usuários do Portal do Software Público Brasileiro [Freitas 2008], a maior parte identificou seu principal perfil como usuário: 59,03%. Os desenvolvedores representam uma parcela significativa, 18,67%, porém bastante inferior. É comum então que o

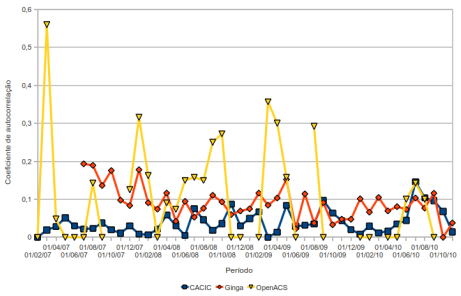

Figura 2. Coeficiente de autocorrelação para comunidades selecionadas

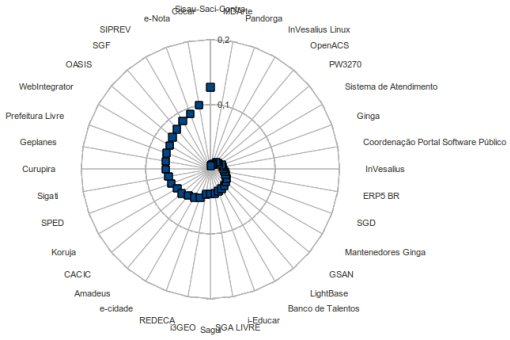

Figura 3. Taxa de Mobilidade das comunidades do Portal do Software Público Brasileiro

contato com a comunidade através do fórum desapareça por um breve período, retornando apenas em alguma ocasião específica.

3.4. Cálculo da taxa de mobilidade

Em uma comunidade aberta de software, os papéis na relação de colaboração podem se alterar entre os membros. Se a construção de conhecimento na comunidade atinge de forma decisiva os envolvidos poderia haver algum tipo de **mobilidade** entre eles. Normalmente inicia-se em uma comunidade perguntando ou expondo suas dúvidas, mas com o passar do tempo e o acúmulo de conhecimento sobre o software é possível responder mais que perguntar. Podemos identificar o fenômeno verificando o percentual de membros presentes em um lado da relação, sendo receptores de conhecimento, que vai para o outro, tornando-se geradores de conhecimento. Alterar o lado da relação no contexto apresentado significa que um usuário que costumava somente perguntar passa também a responder.

Utilizamos $t = 0$ como o período inicial observado. O conjunto $p(t)$ contém todos os usuários que iniciam uma discussão no fórum, ou receptores de conhecimento, enquanto o conjunto $r(t)$ representa os usuários que responderam às perguntas na comunidade – geradores de conhecimento. Na primeira interação ($t = 1$) verificamos quais membros que estavam no conjunto $p(t)$ que aparecem no conjunto $r(t + 1)$, ou seja, $p(t) \bigcap r(t + 1)$, definindo se houve **mobilidade** no período $0 < t <= 1$. Como a aprendizagem não ocorre necessariamente no período seguinte ao surgimento da dúvida, podemos definir os conjuntos $P(t) = \sum_0^k p(t)$ e $R(t) = \sum_0^k r(t)$ para um intervalo $0 < t <= k$. Encontramos assim a relação geral da taxa de mobilidade: $m(t) = \frac{P(t) \bigcap R(t)}{R(t)}$

Para exemplificar a aplicação da taxa proposta, aplicamos a fórmula às comunidades do Portal do Software Público, como apresentado na Figura 3.

4. Considerações finais

Vale a pena notar que o caráter das comunidades afeta a **mobilidade** entre os usuários: softwares cuja utilização é mais simples, dentre os quais estão incluídos os aplicativos, possuem uma probabilidade maior de mobilidade. As causas podem ser analisadas na teoria da Sociologia pelo conceito de capital tecnológico-informacional [Freitas and Meffe 2008]. Em comunidades cujo software é mais fácil de usar, a colaboração acontece de maneira quase natural, à medida que novos membros adquirem conhecimento suficiente e decidem também colaborar. Os dados permitem concluir que em geral as relações entre as comunidades mudam bastante, e que a construção do conhecimento através da mudança do papel do usuário nos fóruns, que passa a responder ao invés de perguntar, é observada com mais frequência naquelas cuja facilidade de instalação e/ou utilização é maior.

A observação sobre a taxa de mobilidade (Figura 3) valida a hipótese de que a facilidade de uso do software favorece a aprendizagem. Contudo, até mesmo o acesso ao Portal é limitado pelos fatores econômicos: pesquisa publicada no próprio Portal [Freitas 2008] aponta que somente 9,3% dos usuários possuem renda inferior a R$ 1.064,00, e somente 7,0% não cursam ou cursaram um curso superior. A taxa de mobilidade poderia ainda ser utilizada em contextos que não o das redes de colaboração de software. O mesmo fenômeno pode ocorrer em escolas e universidades se as redes sociais

forem adicionadas ao processo de ensino-aprendizagem. Também é possível observar a mobilidade em redes corporativas, na promoção da gestão do conhecimento, enfim, em qualquer ambiente onde as relações possuam um caráter de produção colaborativa.

Referências

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics reports*, 424(4):175–308.

Derényi, I., Palla, G., and Vicsek, T. (2005). Clique percolation in random networks. *Phys. Rev. Lett.*, 94(16):160202.

Durugbo, C., Hutabarat, W., Tiwari, A., and Alcock, J. R. (2011). Modelling collaboration using complex networks. *Information Sciences*, 181(15):3143–3161.

Foulonneau, M., Pawelzik, R., Grégoir, B., and Donak, O. (2013). Analyzing the open source communities' lifecycle with communication data. In *MEDES 2013*.

Freitas, C. (2008). Pesquisa sobre o impacto do software público brasileiro. Disponível em: `http://www.softwarepublico.gov.br/result_pesquisa` Acesso em 14 de janeiro de 2011.

Freitas, C. and Meffe, C. (2008). FLOSS em um mundo livre: Inovações e melhores práticas do Brasil. *2020 FLOSS Road Map*.

Gaines, B. R. (2013). Knowledge acquisition: past, present and future. *IJHCS – International Journal of Human-Computer Studies*, 71:135–156.

Hernández, R., Grumet, A., and Team, O. C. (2005). Openacs: robust web development framework. In *12th Annual Tck/Tk Conference*.

Meffe, C. and Freitas, C. (2008). Floss in an open world: Innovations and best practices from brasil. *2020 FLOSS Roadmap*.

Ngonmang, B. and Viennet, E. (2013). Toward community dynamic through interactions prediction in complex networks. In *Signal-Image Technology & Internet-Based Systems (SITIS), 2013 International Conference on*, pages 462–469. IEEE.

Paulovich, F. V., Oliveira, M. C. F., and Minghim, R. (2007). The projection explorer: A flexible tool for projection-based multidimensional visualization. In *Proceedings of the XX Brazilian Symposium on Computer Graphics and Image Processing - SIBGRAPI*.

Silva, A. S., Brito, S. R., Vijaykumar, N. L., Rocha, C. J. A., Costa, J. C. W. A., and Francês, C. R. L. (2013a). Análise de redes sociais para avaliação e monitoramento de programas de treinamento em larga escala baseados no uso de ambientes de aprendizagem e redes sociais online. Disponível em `http://migre.me/jDSlx` Acessado em 05/06/2014.

Silva, A. S., Brito, S. R., Vijaykumar, N. L., Rocha, C. J. A., Costa, J. C. W. A., and Francês, C. R. L. (2013b). Employing online social networks to monitor and evaluate training of digital inclusion agents. *Social Network Analysis and Mining*.

Tofler, A. and Tofler, H. (2007). *Riqueza Revolucionária - O Significado da Riqueza no Futuro*. Futura.