

PLATAFORMA GITHUB: ESTUDO DA REDE DE COLABORAÇÃO ENTRE USUÁRIOS DOS PROJETOS DAS LINGUAGENS DE PROGRAMAÇÃO JAVASCRIPT, PYTHON E GO

Osmir Custódio Mariano¹, Tanilson Dias dos Santos²

¹Graduando em Ciência da Computação – UFT. E-mail: osmirmariano@uft.edu.br

²Professor do Curso de Ciência da Computação- UFT. E-mail: tanilson.dias@uft.edu.br

Resumo: Entende-se por rede de colaboração a interação existente entre usuários, projetos e entidades com competências e interesses distintos no mesmo ambiente de acolhimento. Esta pesquisa estudou especificamente uma rede de colaboração da plataforma Github composta por usuários dos projetos das linguagens de programação Javascript, Python e Go, visando entender como os utentes de cada projeto interagem. Para a obtenção dos dados, foram realizados acessos à plataforma estudada, a fim de verificar a quantidade de usuários e como acontecem as interações entre os mesmos. Com isto, foi possível identificar dois tipos de usuários, os efetivos e contribuintes. Os usuários efetivos podem criar novos repositórios e adicionar novas pessoas aos projetos e os usuários contribuintes são os que não fazem parte desses, porém podem realizar contribuições, visto que todos eles são de código aberto. Cada projeto foi explorado separadamente e foram observadas suas características estruturais, tais como a quantidade de usuários e repositórios.

Palavras-chave: rede de colaboração, projetos, github.

1. INTRODUÇÃO

A internet é um dos principais meios de informações que o homem já desenvolveu, ela traz diversos conceitos, entre esses, estão os ligados a associação e networking. Para [3], as redes de computadores estão alterando os paradigmas de comunicação e associação, isso estabelecendo relacionamentos com um contingente maior de pessoas independentemente das restrições territoriais. Destarte, novas formas de colaborações foram desenvolvidas com este intuito, entre elas, as redes de colaborações, sendo uma delas o Github, um site de codificação social, escrita em Ruby on Rails pelos programadores da Logical Awesome, Chirs Wanstrath, PJ Hyett e Tom Preston. Lançado em 2008 e é utilizado por desenvolvedores do mundo todo, sendo considerada uma das principais redes de compartilhamentos de códigos e mais popular repositório de código fonte aberto[5].

Em razão do site de repositório Github ser uma rede de colaboração entre programadores, o objetivo desse estudo é compreender como acontecem os relacionamentos das colaborações entre os usuários dos projetos das linguagens de programação Python, Javascript e Go.

2. MATERIAL E MÉTODOS

A teoria dos grafos estuda as relações entre os objetos de um determinado conjunto, desse modo o presente artigo faz uma abordagem direta a essa conjectura, visto que realizou um estudo das relações de colaborações entre os projetos das linguagens de programação Python, Javascript e Go. Essa teoria busca analisar propriedades estruturais dos grafos através de representações matriciais e seus espectros, ou seja, dos autovalores das matrizes associadas a eles[6]. Os autovalores de uma matriz correspondem ao valor próprio ou característico da matriz, isto é, dada uma matriz $A_{n \times n}$ dizemos que um escalar λ é chamado de autovalor de A, se existir um vetor não-nulo x pertencente ao resultado da matriz, tal que $Ax = \lambda x$.

Este trabalho foi realizado por meio da pesquisa de natureza descritiva e observatória, e utilizou como procedimento técnico a pesquisa bibliográfica. O estudo foi desenvolvido junto ao site de repositório Github, sendo estudados 3 projetos de linguagens de programação Python, Javascript e Go, desse modo identificando os usuários em cada, bem como as relações de colaborações existentes entre os mesmos.

Com base nos estudos realizados aos projetos, pode-se constatar que durante o período de 26 de maio a 06 de junho de 2016, época em que foram realizadas as coletas dos dados, foram identificadas as quantidades de usuários e repositórios, sendo a linguagem Javascript composta por 29 repositórios e 17 usuários, a Go com 39 repositórios e 31 usuários e Python com 28 repositórios e 42 usuários. Como esses projetos tem uma característica de serem dinamicamente mutáveis, é provável que com o passar do tempo o cenário do ambiente analisado tenha sido alterado.

Para as identificações das relações de colaborações existentes entre os usuários nos projetos, foram efetuados os reconhecimentos das alterações nos códigos fontes dos projetos pelos usuários, isto através da própria plataforma que disponibiliza tais informações. Posteriormente, com estes dados foram gerados os grafos de colaborações para cada projeto, bem como realizados os cálculos dos autovalores e graus dos vértices.

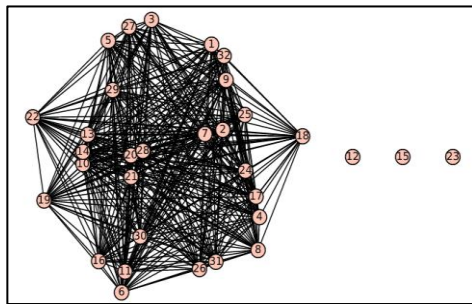
Para o desenvolvimento dos experimentos foram utilizadas as ferramentas *Numpy* usado para computação científica, específico para realização de cálculos e *Sagemath* como IDE de desenvolvimento, ao qual foram implementados todos os testes juntamente com a linguagem de programação Python.

3. RESULTADOS E DISCUSSÃO

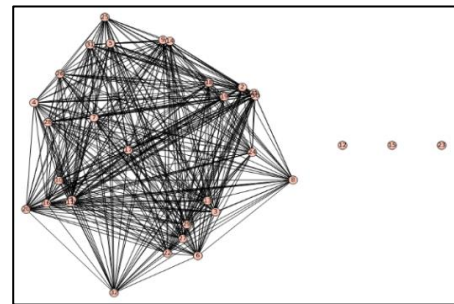
Considerando como a pesquisa tinha o objetivo de estudar a rede de colaboração na plataforma Github em relação aos 3 projetos Javascript, Python e Go, foram identificados os usuários que realizaram alterações no código fonte e submeteram ao projeto, pois a partir dessas identificações foi possível apontar os usuários que se relacionaram no projeto e assim gerar os grafos de colaborações para cada, sendo gerados em 30 de maio às 15h:30min, com o auxílio da ferramenta *Sagemath*. Os vértices foram considerados como utentes e os relacionamentos entre eles as arestas. Através dessas representações foi demonstrado detalhadamente todos os relacionamentos que cada usuário apresentou no projeto, Figura 1 (a), (b) e (c). Observando na Figura 1 (a) e tomando os vértices 5 e 22 como exemplo, percebe-se que estes dois usuários estão se relacionando, visto que os vértices 5 e 22 apresentam ligações entre si, enquanto que o 18 e 12 não dispõe desse relacionamento.

Assim, o presente grafo é caracterizado como um grafo que possuem 4 componentes, sendo que 3 das componentes apresentam graus igual a zero e uma apresentam 3 vértices possuem grau zero, e 29 vértices possuem graus maiores que zero. Os vértices que apresentam grau zero, corresponde aos utentes que não dispõem de relacionamentos e os 29 vértices são os que representam relacionamentos.

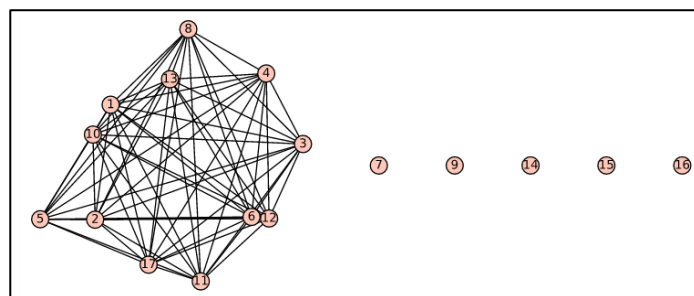
Em relação ao grau dos vértices, isto é, o número de usuários que estão se relacionando. A quantidade de vizinhos é muito importante para a caracterização da rede de colaboração, pois se, por um lado, um vértice não tem vizinhos, este não trocar informações com outros vértices é o que acontece aos grafos dos projetos como se pode observar na Figura 1, existem vértices que não recebem nenhum tipo de ligação, como é o caso do vértice 18 citado.



(a) Grafo de colaboração do projeto Go



(b) Grafo de colaboração do projeto Python



(c) Grafo de colaboração do projeto Javascript

Figura 1 – Grafos de colaborações dos projetos Go, Python e Javascript

Fonte: MARIANO, SANTOS, 2016

Em relação aos cálculos dos autovalores e graus dos vértices a Figura 2 representa a relação entre estes dois valores do projeto Go.

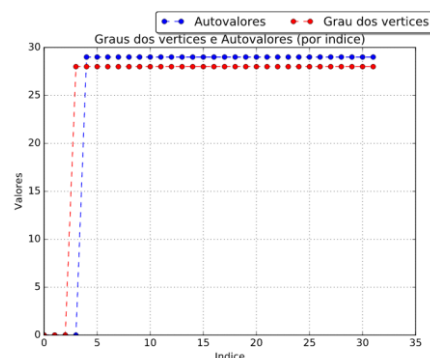


Figura 2 – Gráfico dos autovalores e graus dos vértices do projeto Go

Fonte: MARIANO, SANTOS, 2016

Devido a matriz laplaciana ser semidefinida positiva em todos os projetos, ou seja, todos seus autovalores são maiores ou iguais a zero, os dados obtidos nos 3 projetos demonstram os resultados dos autovalores que são iguais ou maiores que zero e devido os grafos possuírem vértices conexos, estes serão sempre iguais bem como para os graus. E como a matriz laplaciana forma um conjunto linearmente dependente, existem pelo menos um autovalor da laplaciana que é nulo, assim corroborado nos cálculos dos autovalores. Para o presente grafo estão listados os valores em ordem decrescente dos autovalores e graus dos vértices.

[illegible]

Observando as relações entre autovalores, graus dos vértices e número de componentes conexas, percebe-se tanto os graus como os autovalores maiores que zero que eles apresentam valores iguais, isso devido ao grafo apresentar componentes conexas. Atentando para os graus dos vértices e os autovalores, os maiores valores correspondem aos usuários que se relacionam ao projeto e os iguais a zero, são os que não estabelece relacionamento.

6. CONCLUSÕES

De acordo com o estudo realizado entre os projetos das linguagens de programação Python, Javascript e Go, constatou-se que a rede de colaboração dos projetos apresentou características semelhantes, visto que apesar de alguns usuários fazerem parte do projeto, estes não efetivaram contribuições na forma de *upload* de códigos, dessa forma sendo estes usuários representados por vértices desconexos da componente conexa principal dos grafos estudados. A maioria dos usuários que fazem parte de um projeto geralmente estão dispostos a trabalhar e dar contribuições nos projetos.

REFERÊNCIAS

- [1] LIMA, A.; ROSSI, L. and MUSOLESI, M. **Coding Together at Scale: GitHub as a Collaborative Social Network** In **Proceedings of 8th AAI International Conference on Weblogs and Social Media (ICWSM)**. 2014.
- [2] **Assessing the value of cooperation in Wikipedia**, **First Monday**, volume 12, number 4, (2007). At. Disponível em: <<http://rstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1763/1643>>. Acesso em: 23 jun 2016.
- [3] CASTELLS, M. **Sociedade em rede**. 8ª ed. Vol.1. São Paulo: Paz e Terra. (2005)
- [4] CHITUC, C. M.; AZEVEDO A.L. **Multi-Perspective Challenges on Collaborative Networks Business Environment. Collaborative Networks and their Breeding Environments**, New York Springer. 2005.
- [5] FINLEY, K. **GitHub has surpassed Sourceforge and Google Code in popularity**. Disponível em <<http://readwrite.com/> 2011/06/02/github-has-passed-sourceforge>. Acesso em: 10 jun 2016.
- [6] ABREU N. R.; DEL-VECCHIO, V.; TREVISAN, C. V. **Teoria Espectral de Grafos – Uma Introdução**. III Colóquio de Matemática da Região Sul. (2014)