

UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE FÍSICA DE SÃO CARLOS

ALEXANDRE MARTINS PASTORE

Influência das características de conectividade na execução distribuída de tarefas
em redes complexas

São Carlos
2022

ALEXANDRE MARTINS PASTORE

Influência das características de conectividade na execução distribuída de tarefas
em redes complexas

Dissertação apresentada ao Programa de Pós-Graduação em Física do Instituto de Física de São Carlos da Universidade de São Paulo, para obtenção do título de Mestre em Ciências.

Área de concentração: Física Computacional
Orientador: Prof. Dr. Gonzalo Travieso

Versão Corrigida
(versão original disponível na Unidade que aloja o Programa)

São Carlos
2022

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTE TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Martins Pastore, Alexandre
Influência das características de conectividade na execução distribuída de tarefas em redes complexas / Alexandre Martins Pastore; orientador Gonzalo Travieso - versão corrigida -- São Carlos, 2022.
52 p.

Dissertação (Mestrado - Programa de Pós-Graduação em Física Computacional) -- Instituto de Física de São Carlos, Universidade de São Paulo, 2022.

1. Redes complexas. 2. Barabasi-Albert. 3. Esdos-Renyi. 4. Holme-Kim. 5. Lancichinetti-Fortunato-Radicchi. I. Travieso, Gonzalo, orient. II. Título.

todos que me apoiaram e incentivaram a continuar nessa caminhada, minha eterna gratidão pela compreensão e apoio ao longo período de elaboração deste trabalho.

AGRADECIMENTOS

Agradeço a todos que contribuíram e me apoiaram nestes anos de pesquisa, em momentos difíceis e atípicos que todos estamos passando.

Em particular queria agradecer ao meu orientador, Prof. Dr. Gonzalo, por sempre acreditar na conclusão deste trabalho, sempre preocupado como andavam os experimentos e disposto a discutirmos os resultados.

Também gostaria de agradecer minha mãe e minha esposa que sempre estiveram me incentivando e torcendo para dar tudo certo.

A todos os amigos e conhecidos que fiz na cidade de São Carlos, muitos não estão na cidade, mas mesmo longe sempre estavam me incentivando.

O presente trabalho foi realizado com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) Código de Financiamento 133601/2019-6.

RESUMO

PASTORE, A. M. **Influência das características de conectividade na execução distribuída de tarefas em redes complexas**. 2022. 52 p. Dissertação (Mestrado em Ciências) - Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2022.

Redes complexas consistem em uma representação em forma de grafo que apresenta propriedades topográficas bastante particulares, não encontrado em grafos simples. Um exemplo comum desse tipo de representação é a rede Web do Google. Haja vista a importância dos grafos, as redes complexas podem ser extremamente úteis para distribuição de tarefas em uma rede. Nesse contexto, este trabalho teve como finalidade passar por algumas características consecutivas na execução de distribuição de tarefas para saber qual a influência dessas características, denominadas características de distribuição da rede, efeito do grau médio, efeito heterogeneidade de graus, efeito de coeficiente de aglomeração e Efeito de estrutura de comunidade. Para tal, considerou-se a distribuição de redes conhecidas, como redes Barabási-Albert, Erdős-Rényi, Holme-Kim e Lancichinetti-Fortunato-Radicchi. Essas características foram estudadas através de simulações observando o parâmetro do tempo médio de execução de tarefas, tempo médio na espera da fila e o tempo médio ocioso dos nós. Os resultados aqui apresentados, demonstram que é possível identificar a influência na distribuição da rede atribuída à execução de tarefas, mas não exclusivamente o motivo são essas características, o que possibilitará em aplicações futuras uma nova abordagem para descobrir quanto essa característica influencia e assim outras características podem influenciar.

Palavras-chave: Redes complexas. Barabási-Albert. Erdős-Rényi. Holme-Kim. Lancichinetti-Fortunato-Radicchi.

ABSTRACT

PASTORE, A. M. **Influence of connectivity characteristics in the distributed execution of tasks in complex networks**. 2022. 52 p. Dissertation (Master of Science) - São Carlos Institute of Physics, University of São Paulo, São Carlos, 2022.

Complex networks consist of very particular topographical properties graphical representation, not commonly found in simple graphs. A standard example is the Google web network. Given the importance of graphs, complex networks are extremely useful for distributing tasks in a network. In this context, this work aimed to go through some consecutive characteristics in the execution of task distribution to know the influence of these particular characteristics, named as network characteristics distribution, average degree effect, degree heterogeneity effect, agglomeration coefficient effect, and community effect. Thus, we consider the distribution of known networks, such as Barabási-Albert, Erdős-Rényi, Holme-Kim, and Lancichinetti-Fortunato-Radicchi networks. These characteristics were studied through simulations observing the average time of execution of tasks, average time in the queue, and the average idle time of the nodes. These results demonstrated the feasibility to identify the influence on the distribution of the network attributed to the execution of tasks, but not exclusively the reason these characteristics, which will allow in future applications a new approach to discovering how much this characteristic and other characteristics may influence the network.

Keywords: Complex networks. Barabási-Albert. Erdős-Rényi. Holme-Kim. Lancichinetti-Fortunato-Radicchi.

LISTA DE ABREVIATURAS E SIGLAS

BA	Barabási-Albert
ER	Erdős-Rényi
FT	formação de tríades
HK	Holme-Kim
LFR	Lancichinetti-Fortunato-Radicchi
WS	Watts e Strogatz

SUMÁRIO

1	Introdução	17
2	Desenvolvimento	19
2.1.	Conceitos básicos	19
2.2.	Distribuição de graus	19
2.3.	Caminho mais curto	20
2.4.	Agrupamento	20
2.5.	Modelos de redes	21
2.5.1	Erdős-Rényi	21
2.5.2	Barabási-albert	22
2.5.3	Home-Kim	23
2.5.4	Lancichinetti-Fortunato-Radicchi	25
3	Modelo de execução colaborativa de tarefas	29
4	Resultados	33
4.1.	Efeito de distribuição de rede	33
4.2.	Efeito do grau médio	37
4.3.	Efeito Heterogeneidade de graus	41
4.4.	Efeito de coeficiente de aglomeração	43
4.5.	Efeito de estrutura de comunidade	46
5	Conclusão	51
	Referências	53

1 INTRODUÇÃO

O estudo de redes complexas é um tema interdisciplinar que abrange diversas áreas de conhecimento, tais como a matemática, biologia, física, sociologia e ciência da computação. Redes complexas se referem a um grafo que apresenta uma estrutura topológica não trivial, o que de forma abreviada seria composta por um conjunto de vértices (nós) que são interligados por meio de arestas.¹ A primeira vez em que foi proposto o estudo de grafos foi na resolução do problema das pontes de *Königsberg* proposto por Leonard Euler.

O sucesso de novas áreas científicas pode ser avaliado pelo potencial em contribuir para novas abordagens teóricas e em aplicações a problemas do mundo real. Redes complexas têm se saído extremamente bem em ambos os aspectos, com sua sólida base teórica sendo desenvolvida ao longo dos anos e com uma variedade de aplicações.²

No mundo real diversos aspectos podem ser representados por meio de uma rede complexa partindo da analogia de problemas específicos, como por exemplo, a computação em nuvem em que um problema que influencia fortemente o desempenho de tais sistemas é a escolha de nós como servidores responsáveis pela execução das tarefas dos clientes.³

Com o avanço tecnológico de informações e a disponibilidade de computadores e redes de comunicação mais amplas, foi permitida a análise de dados em grande escala, causando uma mudança significativa na área.¹ As pesquisas, que antes tinham seu foco em redes consideradas pequenas e nas propriedades dos vértices individuais ou arestas, passaram a levar em consideração propriedades estatísticas em larga escala.¹ Nos dias atuais é comum um estudo com redes envolvendo milhões de vértices, as quais anteriormente eram compostas apenas por algumas dezenas ou, em casos específicos, centenas de vértices. Esta mudança de paradigma mostrou que as redes reais diferem das redes aleatórias, o que por muito tempo foi considerado como o principal modelo de redes. Foi descoberto que a topologia e a evolução das redes do mundo real apresentam propriedades organizacionais bem robustas e distintas das redes aleatórias. Essa é a principal razão pela qual as redes passam a ser chamadas de redes complexas.

As características topológicas tem impactado novos estudos em diversas áreas.⁴ Por um lado, os cientistas têm que lidar com questões estruturais, como caracterizar a topologia de uma arquitetura de fiação complexa, revelar os princípios unificadores que estão na base das redes reais e desenvolver modelos para imitar o crescimento de uma rede e reproduzir suas propriedades estruturais. Por outro lado, muitas questões relevantes surgem ao estudar a dinâmica de redes complexas, como aprender como um grande conjunto de sistemas

dinâmicos que interagem através de uma topologia de fiação complexa pode se comportar coletivamente.

Um modelo de execução colaborativa de tarefas pode ser apresentado como um trabalho que é realizado em conjunto de duas ou mais pessoas para realização de objetivos comuns. Um exemplo como pode ser apresentado para exemplificar como seria a execução de tarefas de modo colaborativas seriam em um escritório onde temos varias pessoas de um departamento e apenas uma pessoa tem autorização para carimbar documentos para serem repassados para a chefia, por exemplo, com apenas uma pessoa podendo fazer essa tarefa ela acaba sendo sobrecarregada e levando muito tempo para conseguir carimbar todas as folhas que estão na pilha de sua mesa. Ao repassar essa tarefa para outras pessoas do departamento, fazendo assim uma distribuição da tarefa, a pilha em sua mesa é repartida e a realização da tarefa é realizada de modo mais rápido.

O problema de execução colaborativa de tarefas seria que as tarefas que são geradas e executadas por agentes podem ficar muito tempo parada em um mesmo agente criando uma fila de espera em um determinado agente enquanto outro agente pode estar ocioso por não ter uma tarefa para ser executada, com a execução colaborativa isso pode ser resolvida mas se for escolhida de forma equivocada uma tarefa pode levar muito mais tempo de ser executada do que se não tivesse sido distribuída.

Com os estudos que são feitos neste trabalho se espera que as características da rede vão afetar o desempenho na execução colaborativa de tarefas, sendo que com este estudo pode-se distribuir melhor as tarefas dentro de uma rede gerando um término mais rápido de execução.

Alguns trabalhos já estudaram alguns casos de distribuição de tarefas, que seriam “On the effects of geographical constraints on task execution in complex networks”⁵, “Predicting efficiency in master-slave computing system,”⁶ “Effective networks for real-time distributed processing,”⁷ “Complex grid computing.”⁸

Este trabalho se diferencia dos anteriores pois apresenta pela primeira vez a abordagem na execução de tarefas de modo em que sua distribuição não seja centralizada.

2 DESENVOLVIMENTO

2.1 Conceitos básicos

Uma rede é uma abstração com a qual se pode codificar um relacionamento entre pares de objetos.⁹ O conjunto de objetos a rede será denotado por V e o número de de objetos por $n = |V|$ e daremos o nome para estes objetos de vértices ou nós da rede. Ao existir um relacionamento entre dois objetos, esse relacionamento será apresentado por um par não-ordenado. De tal forma, se $i, j \in V$ estão relacionados, será representado pelo par não-ordenado (i, j) . Este par ordenado que representa que existe o relacionamento, será chamado de aresta. O conjunto de todos os relacionamentos existentes entre os objetos em V , ou seja, o conjunto de todas as arestas da rede, será denotado por E . Com isso, temos que $E = \{(i, j) | i, j \in V, i \text{ está relacionado com } j\}$. O número de arestas da rede, m , é dado pelo número dos pares não-ordenados no conjunto E , ou seja, $m = |E|$. Logo, uma rede é definida por estes dois conjuntos, $R = (V, E)$, o que na computação é equivalente a definição de um grafo

Os conceitos básicos de uma rede complexa seriam:¹⁰

- a) Vértice: A unidade fundamental da rede.
- b) Aresta: A linha que conecta dois vértices.
- c) Direção/sem direção: Uma aresta é dita direcionada se ela tiver apenas uma direção e sem direção se correr em ambas as direções.
- d) Grau: O número de arestas conectadas a um vértice. Observe que o grau não é necessariamente igual ao número de vértices adjacentes a um vértice, sendo assim pode haver mais de uma aresta entre quaisquer dois vértices, mas neste trabalho iremos apenas usar grafos simples e nem ligações de um vértice com ele mesmo.

2.2 Distribuição de graus

O grau de um vértice em uma rede é o número de arestas incidentes sobre esse vértice.¹¹ Equivalentemente, definimos p_k como a probabilidade de que um vértice escolhido uniformemente ao acaso tenha grau k . Podemos estimar p_k para qualquer rede plotando um histograma dos graus dos vértices. Este histograma é a distribuição de graus para a rede.

Uma forma alternativa de apresentar os dados de grau e fazer um gráfico da função de distribuição cumulativa

$$P_k = \sum_{k'=k}^{\infty} p_{k'} \quad (1)$$

que é a probabilidade de que o grau seja maior ou igual a k .

2.3 Caminho mais curto

Em muitos ramos da ciências muitos sistemas podem ser modelados como grafos aleatórios que podem compartilhar muitas propriedades geométricas, alguns exemplos dessas características seriam o crescimento logarítmico do caminho mais curto, sendo este caminho na minimização do custo de travessia de uma rede entre dois nós, l .¹²

2.4 Agrupamento

O agrupamento, como no modelo *Watts-Strogatz* (WS),¹³ foi introduzido o coeficiente de agrupamento $\gamma \equiv \langle \gamma \rangle$.¹² O coeficiente de agrupamento local γ_v do vértice é definido da seguinte maneira: suponha que o vértice v tenha k_v vizinhos, sendo que k_v é chamado como o grau médio do vértice v , um vizinho é um vértice separado por exatamente uma aresta). Para esses vizinhos k_v , pode existir no máximo.

$$\binom{k_v}{2} = \frac{k_v(k_v-1)}{2} \quad (2)$$

de arestas conectando dois dos vértices k_v . Se define $|\varepsilon(\Gamma_v)|$, sendo ε o conjunto de arestas e Γ o agrupamento local, como o número de arestas reais existentes na rede conectando aqueles vizinhos, o coeficiente de agrupamento local é escrito como

$$\gamma_v \equiv \frac{|\varepsilon(\Gamma_v)|}{\frac{k_v(k_v-1)}{2}} \quad (3)$$

fica claro que γ é uma medida do número relativo de triádes(subgrafos totalmente conectados de três vértices). é observado também que γ está estritamente no intervalo de $[0,1]$.

2.5 Modelos de redes

Uma rede aleatória é uma rede onde as conexões entre os nós são escolhidas por um processo estocástico e a seguir serão apresentados alguns modelos usando alguns desses modelos.

2.5.1 Erdős-Rényi

As redes complexas¹⁴ podem ter diferentes modelos e o mais útil e simples a ser apresentado primeiro seria o grafo aleatório que é construído com arestas não direcionadas que são colocadas aleatoriamente entre um número fixo n de vértices para criar uma rede na qual cada aresta tem $\frac{1}{2}n(n-1)$ possíveis ser independentemente presente com alguma possibilidade, e o número de arestas ligadas a cada vértice - o grau do vértice - é distribuído de acordo com uma distribuição binomial ou uma distribuição Poisson no limite de grande n .

No modelo de ER ,¹⁴ constrói-se uma rede aleatória tendo N vértices e m arestas. As m arestas são escolhidas uniformemente e aleatoriamente dentre o conjunto de todas as $0 \leq m \leq \frac{N(N-1)}{2}$ arestas equiprováveis possíveis. Um exemplo desta rede com esse modelo pode ser vista na Figura 1

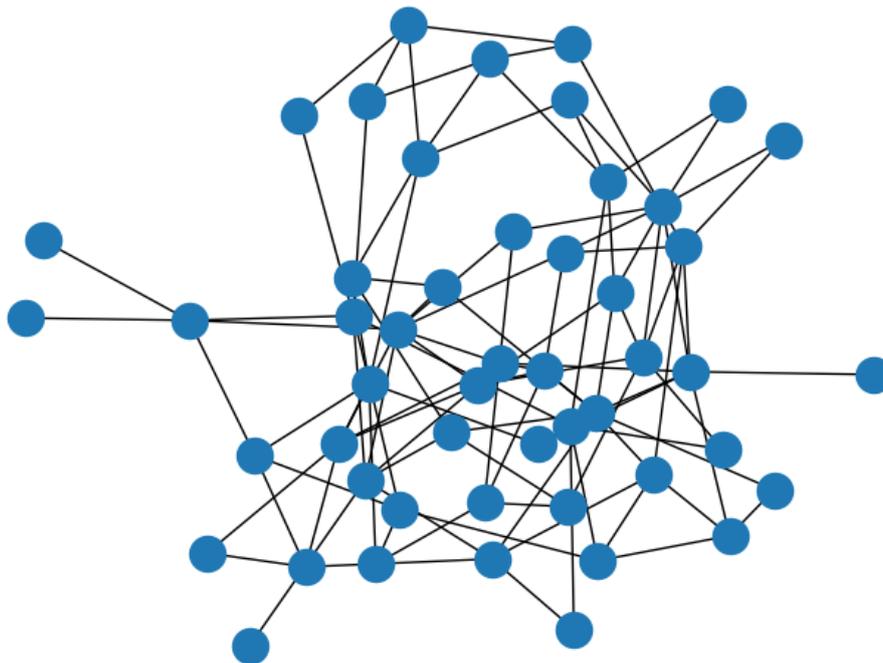


Figura 1 - Representação de uma rede complexa no modelo ER .
Fonte: Elaborada pelo autor

2.5.2 Barabási-Albert

No modelo de BA^{15} é composto por ter seus vértices que são adicionados um a um à rede com grau k , sendo a outra extremidade de cada aresta anexada outro vértice preexistente com probabilidade proporcional ao grau desse vértice. Neste modelo as arestas não são direcionadas, portanto não há distinção entre grau de entrada e saída.

Este modelo pode ser descrito pelos seguintes passos:

- Condição inicial: Para começar, a rede consiste em v_0 vértices
- Crescimento: um vértice v com m arestas é adicionado em cada passo de tempo, o tempo t é identificado como o número de passos de tempo
- Conexão preferencial: Cada aresta de v é então anexada a um vértice existente com a probabilidade proporcional ao seu grau, ou seja, a probabilidade para um vértice w ser anexada a v é $P_w = \frac{k_w}{\sum k_v}$

Um exemplo de como será uma rede complexa com o modelo BA, pode ser visto na figura 2.

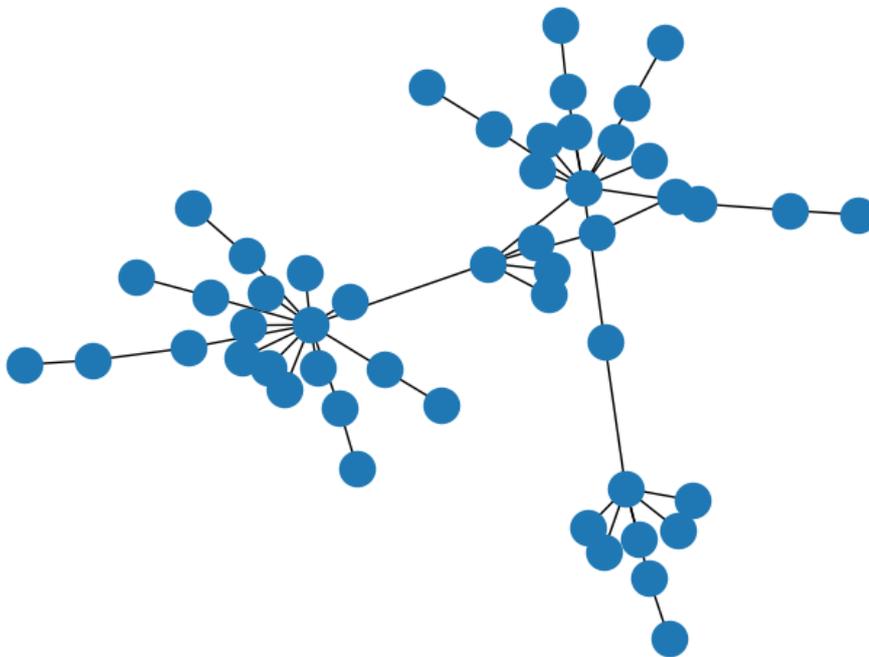


Figura 2 - Representação de uma rede no modelo BA.
Fonte: Elaborada pelo autor

2.5.3 Holme e Kim

Embora os modelos ER e BA tenham um l logaritmicamente crescente com o tamanho da rede, cada modelo não tem a propriedade que o outro modelo apresenta, no caso do modelo WS¹³, mostra um alto agrupamento, mas sem a distribuição de grau de lei de potência, enquanto o modelo BA com a natureza sem escala não possui a alta aglomeração.

O modelo BA representa com sucesso a natureza livre de escala de muitas redes, mas não descreve corretamente as redes com alto agrupamento. o modelo HK¹² adiciona ao modelo BA uma etapa em sua definição que seria a formação de tríades, mas para explicar esse passo é preciso relembrar os passos da definição do modelo BA, que seriam:

no modelo BA, o passo de crescimento é então iterado $N = |V|$ vezes, e para cada passo de crescimento o passo anexo preferencial é iterado m vezes para m arestas do vértice v recém-adicionado.

- Formação de tríades (TF): Se uma aresta entre v e w foi adicionada na etapa anexo preferencial anterior, então adicione mais uma aresta de v a um vizinho escolhido aleatoriamente de w . Se não houver nenhum par para conectar, ou seja, se todos os vizinhos de w já estiverem conectados a v , faça um passo PA.

Quando um vértice v e com aresta m é adicionado à rede existente, primeiro executamos uma etapa Anexo preferencial e, em seguida, realizamos uma etapa formação de tríades com probabilidade P_t ou uma etapa Anexo preferencial com probabilidade $1 - P_t$, que tomamos como parâmetro de controle no modelo.

O modelo de rede livre de escala padrão não apenas gera redes com certas propriedades geométricas, mas sugere um mecanismo para o surgimento de distribuições de graus de lei de potência em redes em evolução.

O algoritmo de rede sem escala em cluster definido acima fornece a mesma distribuição de grau que a rede sem escala padrão, pelo menos em cada etapa de TF seguindo uma etapa de PA.

Um exemplo de como seria uma rede complexa apresentando o modelo HK pode ser vista na figura 3.

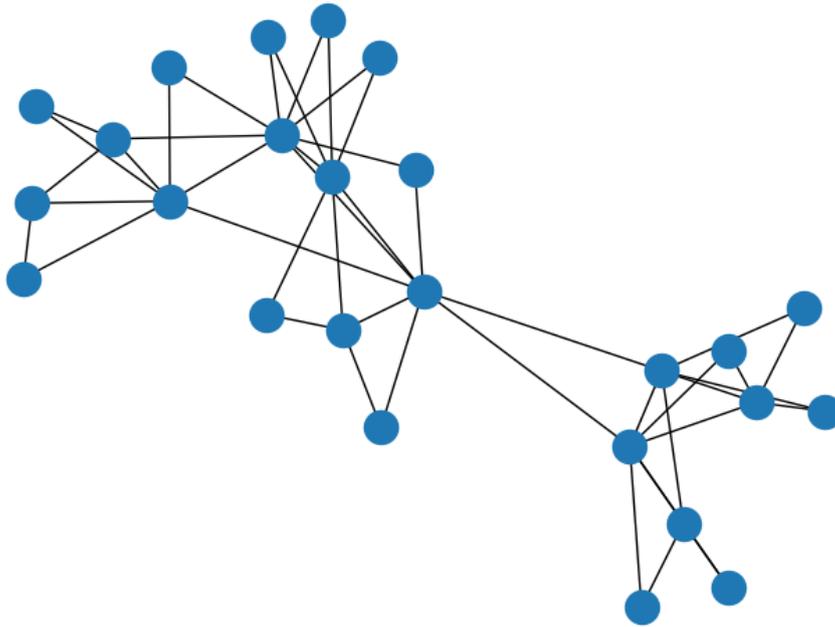


Figura 3 - Representação de uma rede no modelo *HK*.
Fonte: Elaborada pelo autor

2.5.4 Lancichinetti-Fortunato-Radicchi

Muitos sistemas complexos exibem uma combinação de compartimentos que são bastantes independentes umas das outras, na representação por grafos de um sistema complexo onde as unidades elementares de um sistema são descritos como nós e suas interações mútuas como arestas, tal estrutura modular é revelada pela existência de grupos de nós, chamados de comunidades ou módulos, com muitas ligações conectando nós dos mesmo grupo e comparativamente poucas arestas unindo nós de diferentes grupos.

Comunidades revelam uma interação não trivial na organização final de uma rede, e permitem inferir relações entre os nós, que podem não ser facilmente acessíveis a partir de testes empíricos diretos.

O modelo LFR¹⁷ cita que tanto a distribuição de graus quando de tamanho da comunidade são leis de potências, com expoentes α e β , respectivamente. Assumindo que N é o número de nó e k o grau médio.

Na construção do modelo LFR são realizados os seguintes procedimentos:

- 1) Cada nó recebe um grau obtido de uma lei de potência com expoente α . Os extremos da distribuição k_{min} e k_{max} são escolhidos de modo que o grau médio seja o modelo de configuração usado para conectar o nó para manter sua sequência de graus.

- 2) Cada nó compartilha uma fração $1 - \mu$ de suas arestas com outros nós de sua comunidade e fração μ com os demais nós da rede; μ é o parâmetro de mistura.
- 3) Os tamanhos das comunidades são tiradas da lei da distribuição de potência com expoente α , tal que a soma de todos os tamanhos é igual ao número n de nós do grafo. Os tamanhos de comunidade mínimo e máximo s_{min} e s_{max} são escolhidos para respeitar as restrições impostas pela nossa definição de comunidade: $s_{min} > k_{min}$ e $s_{max} > k_{max}$. Isso garante que um nó de qualquer grau pode ser incluído em pelo menos uma comunidade.
- 4) No início, todos os nós não estão atreladas a nenhuma comunidade. Na primeira iteração, um nó é atribuído a uma comunidade escolhida aleatória; se o tamanho da comunidade exceder o grau interno do nó, ou seja, o número de seus vizinhos dentro da comunidade, o nó entra na comunidade, caso contrário fica sem comunidade. Em interações sucessivas, colocamos um nó sem comunidade, colocamos em uma comunidade escolhida; se esta estiver completa, lançamos um nó selecionado aleatoriamente da comunidade, que fica sem comunidade. O procedimento pára quando não há mais nós sem comunidade.
- 5) Para impor a condição sobre a fração de vizinhos expressos pelo parâmetro μ , várias etapas de religação são executadas, de modo que os graus de todos os nós permanecem os mesmo e somente a divisão entre grau interno e externo seja afetada, quando necessário. Desta forma a proporção entre grau externo e interno de cada nó em sua comunidade será definida para o compartilhamento desejado com boa aproximação.

Uma rede apresentando o modelo LFR pode ser vista na figura 4.

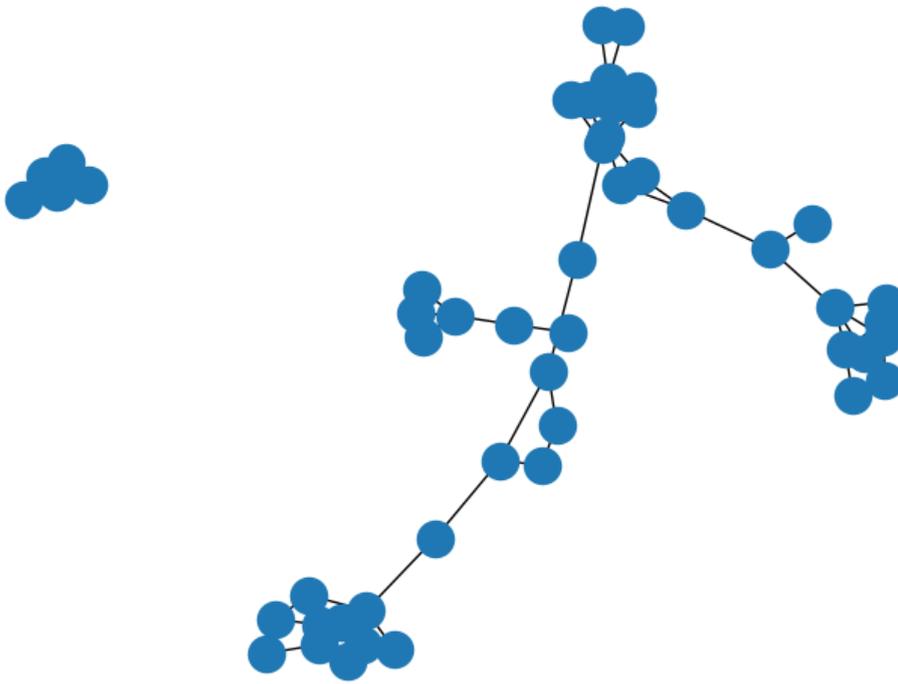


Figura 4 - Representação de uma rede no modelo *LFR*.
Fonte: Elaborada pelo autor

3 Modelo de execução colaborativa de tarefas

Neste modelo, existe um conjunto de agentes, que podem ser exemplificados como pessoas em um empreendimento cooperativo ou computadores em uma rede. Os agentes são responsáveis pela execução de tarefas e estas são geradas para um agente em um dado instante de tempo e com uma certa duração, sendo que nesta simulação todos os tempos considerados são discretos e a centralização é simultaneamente em todos os assentos.

Uma tarefa ao ser gerada para um agente pode ser executada por esse mesmo agente ou ser entregue para outro agente para execução, quando uma tarefa chega para um agente, se ele estiver desocupado, começa imediatamente a execução da tarefa e fica ocupado até seu término, pois as tarefas não podem ser interrompidas. Caso uma tarefa chegue quando o agente está ocupado, ela vai para uma fila de espera associada ao agente.

Quando o agente termina a execução de uma tarefa, ele pega uma nova tarefa da sua fila de espera para executar ou se essa fila estiver vazia, fica desocupado até a chegada de uma nova tarefa. Cada agente tem o conhecimento de alguns outros agentes, seus vizinhos. A relação de vizinhança entre os agentes forma uma rede de colaboração.

A cada instante de tempo, o agente verifica se ele está sobrecarregado (muitas tarefas para executar). Se estiver, escolhe uma das tarefas em sua fila de espera e a envia para um de seus vizinhos. A tarefa é colocada na fila de espera do vizinho.

Para melhor exemplificação de como este modelo é na prática vamos usar a Figura 5

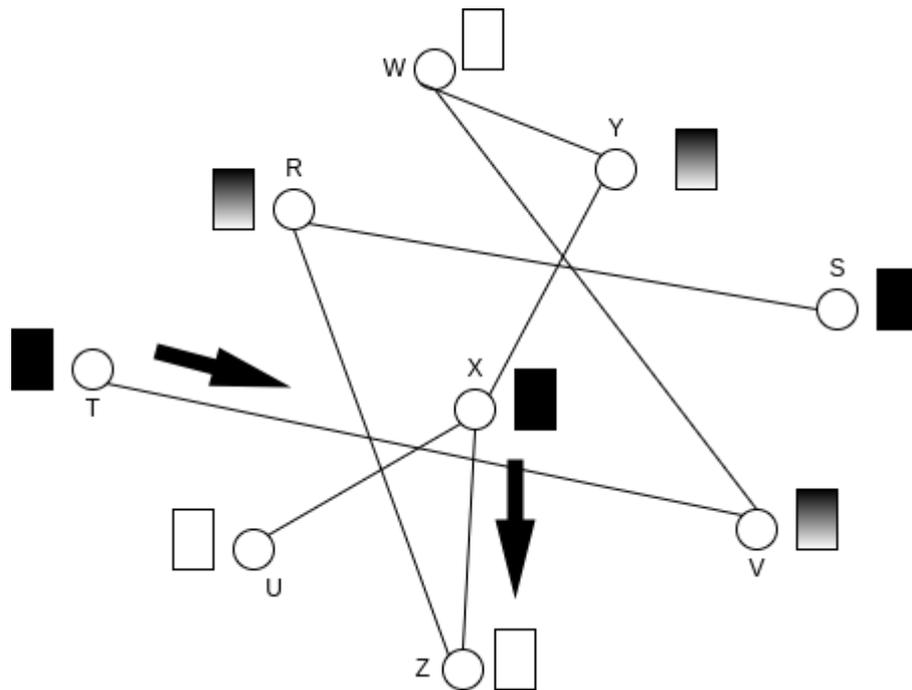


Figura 5 - Representação de um modelo de execução de tarefas.
 Fonte: Elaborada pelo autor

Pela Figura 5 vemos em um instante da criação de uma tarefa no nó T e X, como estes nós estão com suas respectivas filas com mais tarefas do que o limiar, cada tarefa criada no nó e redistribuída para um nó vizinha, no exemplo mostrado na Figura 5 a tarefa criada no nó T é transmitida para o nó V enquanto a tarefa criada no nó X é repassada para o nó Z. Para ambos os casos os nós que receberam as novas tarefas não estão com suas filas com mais tarefas do que o limiar, então com isso as novas tarefas serão adicionadas as filas como última a ser executada.

No caso do nó Z a tarefa que chega nela já será executada pois a fila se encontra vazia, já a tarefa que foi repassada para o nó V irá esperar para ser executada.

Após a descrição do modelo, algumas decisões precisam ser tomadas para serem feitas as simulações e essas decisões foram tomadas em cima de algumas perguntas e elas seriam:

- Quando são geradas as tarefas?

Uma tarefa pode ser gerada em cada agente, em cada instante de tempo, com uma probabilidade fixa. Foi decidido isto por causa que esta é a hipótese mais simples de geração de tarefas, com independência entre os instantes de geração de cada tarefa e entre os diversos agentes. Isto implica uma geração a taxa fixa de tarefas. Geração de tarefas em rajadas (muitas tarefas em um curto espaço de tempo, seguido de períodos maiores sem tarefas) pode ser um estudo interessante para o futuro, mas não faz parte desse trabalho.

- Qual o tamanho das tarefas?
O tamanho é fixo é igual para todas as tarefas, pelo fato de que é a situação mais simples, onde temos apenas um parâmetro para o tamanho das tarefas. Neste trabalho foi usado o tamanho das tarefas de 20. Estudos sobre o efeito de variação no tamanho de tarefas devem ser feitos no futuro, mas não são parte deste trabalho.
- Quando um agente retira uma tarefa da fila de espera para executar, que tarefa retirar?
Ele retira a tarefa que foi inserida na fila a mais tempo para evitar que tarefas fiquem na fila de espera por muito tempo, retirando as que lá estão a mais tempo, e não estabelecer nenhuma prioridade entre as tarefas. Sistemas com prioridades, por exemplo, baseado no tamanho das tarefas ou se as tarefas foram geradas localmente ou recebidas de vizinhos, podem ser interessantes, mas não fazem parte do estudo deste trabalho.
- Quando um agente decide que ele está sobrecarregado?
A decisão é baseada no tamanho da fila de espera, se tiver mais do que 4 tarefas, o agente se considera sobrecarregado. Justificativa para esta decisão é o fato de que olhar apenas o tamanho da fila é uma decisão puramente local, onde o agente não precisa considerar o que está ocorrendo com os outros agentes. Decisões que consideram o estado dos agentes vizinhos podem ser mais eficientes, mas são mais complexas: um agente com muitos vizinhos precisaria gastar tempo se comunicando com todos eles para decidir pela distribuição ou não de uma tarefa. Novamente, o estudo de outros algoritmos para decisão de distribuição de tarefas não faz parte deste trabalho. O número 4 é de certa forma arbitrário, mas foi baseado em alguns estudos preliminares, que mostram que valores menores implicam uma grande quantidade de movimentação de tarefas de um agente para o outro, enquanto valores maiores não apresentam muita diferença na dinâmica. Um estudo mais aprofundado do melhor valor, e a verificação se ele depende dos outros parâmetros da simulação, deve ser feito no futuro, mas não faz parte deste trabalho.
- Ao entregar uma tarefa para um vizinho, qual tarefa da fila de espera deve ser escolhida?
A tarefa no final da fila (a que foi inserida mais recentemente), isso se deve a que tarefas que estão há mais tempo na fila local estão mais próximas de serem

executadas pelo agente, enquanto que foi inserida recentemente é a que teria que esperar mais tempo para ter acesso ao agente local, então vale a pena ela tentar se executada em outro agente. Outros esquemas para a escolha da tarefa a enviar podem ser considerados, mas não fazem parte deste trabalho.

- Ao entregar uma tarefa para um dos vizinhos, qual deles escolher?

O vizinho escolhido é aleatório porque é um critério puramente local, que não precisa de informações sobre a situação dos vizinhos. Novamente, usar informações sobre o vizinho poderia possibilitar uma melhor distribuição de tarefas, mas constituiu um método mais complexo, que não faz parte deste trabalho.

A cada 100 passos de execução eram registrados o tamanho da fila em cada agente da rede de distribuição em arquivos que foram usados para os estudos e construções dos gráficos que são apresentados na seção de resultados.

4 Resultados

Para a aquisição dos dados para estudos e ter números suficientes para garantir uma veracidade dos dados apresentados a seguir foram feitos para cada modelo entre ER, BA, HK e LFR-benchmark, foram criada dez redes distintas de cada e para cada uma dessa rede foi executado 10 vezes a simulação gerando um total de 100 simulações.

Foi usado, e calculado, o valor médio desses 100 valores para a construção dos gráficos. Também foi calculado o desvio padrão para dar informação sobre a variabilidade dos valores.

Um valor que queremos observar nas simulações seria o ponto de sobrecarga de tarefas, denominado como ρ_c que pode ser calculado pela equação (6)

$$\begin{aligned} \rho_c &= \frac{\sum_{i=1}^L t_i}{L} \\ \rho_c &= 1 \\ \rho_c &= \frac{1}{L} \quad (6) \end{aligned}$$

sendo ρ_c a representação do tempo todo ocupado. E para este trabalho como temos o mesmo tamanho das tarefas, o valor encontrado para ρ_c foi de 0,05. Então a partir deste ponto os agentes começam a apresentar uma sobrecarga de tarefas.

4.1 Efeito de distribuição da rede

Antes de avaliar a influência de parâmetros topológicos nas redes, será verificado se apenas o uso ou não de distribuição de tarefas para redes é vantajoso e para isso foi realizada a simulação para a comparação do caso com e sem distribuição para redes com as mesmas características.

A rede escolhida para essa comparação foi a rede com a topografia ER, na realização desta simulação com e sem distribuição foi escolhido o valor de 4 (quatro) para a distribuição, que significa que quando aquele nó tiver mais do que quatro tarefas esperando para serem executadas, uma delas será redistribuídas para seus vizinhos para ser executada.

Outros parâmetros, como quantidade de nós, valor do grau médio, tamanho das tarefas, passos de simulação e guardar o tamanho da fila depois de determinado instante de unidade de tempo, foram mantidos iguais nas duas simulações, lembrando que os valores usados são: para quantidade de nós foi 1000 (mil); para o grau médio foi 4 (quatro), tamanho

da tarefa de 20 (vinte); 10.000 (dez mil) passos de simulação; e a cada 100 (cem) instantes guardar o tamanho da fila. Também foi usado o intervalo de $[0,02;0,06]$ para a taxa de geração de tarefas.

O resultado da simulação entre as duas condições é mostrado na Figura 6.

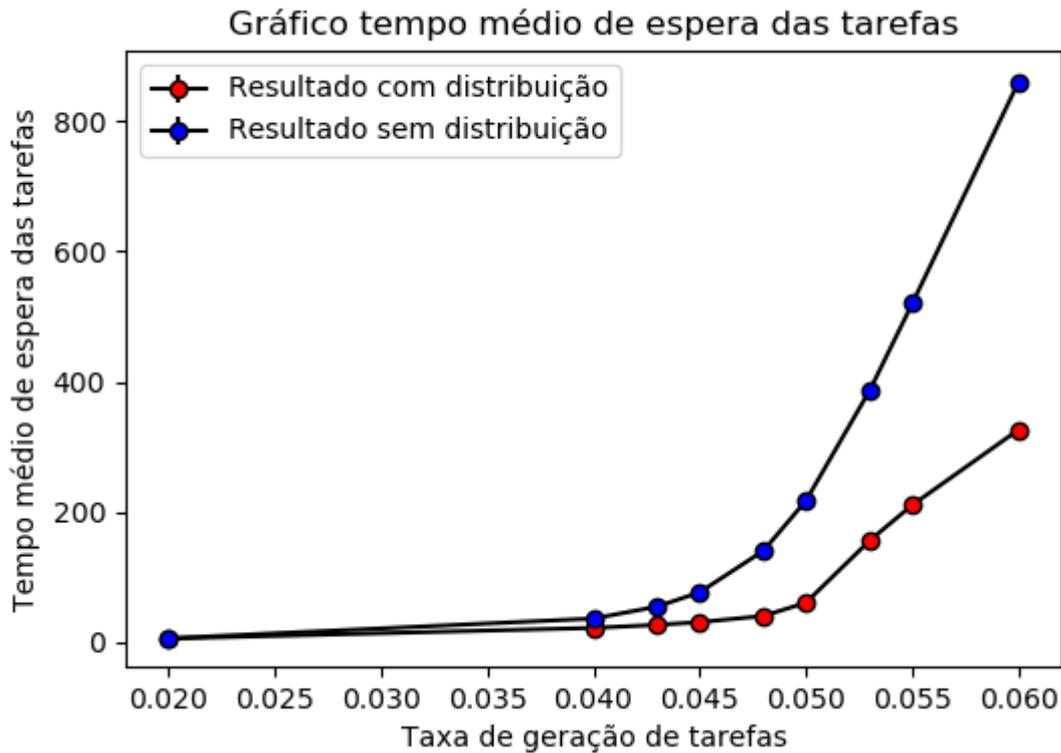


Figura 6 - Gráfico de tempo médio de espera das tarefas por ρ para uma rede sem distribuição e um resultado com distribuição na topografia ER.

Fonte: Elaborada pelo autor

Neste gráfico apresentamos o conceito sobre a taxa de geração de tarefas, que iremos denominar com a letra grega ρ . Um valor que é interessante procurarmos seria o ponto crítico de geração de tarefas, denominada como ρ_c , na teoria esse ponto tem o valor de 0,05. Mas no gráfico percebemos que antes mesmo do valor teórico já é perceptível um aumento rápido.

O ponto ρ_c é onde ocorre uma sobrecarga de tarefas na rede, a partir deste ponto os agentes começam a ter mais tarefas do que conseguem realizar. Comparado ao valor teórico, mencionado anteriormente, nota-se que a sobrecarga começa um pouco antes pelo fato que esse valor seria o encontrado se existisse uma distribuição perfeita de tarefas entre os agentes, o que não é o caso, pois existe uma flutuação de distribuição, assim valores abaixo do crítico já apresentam um certo nível de sobrecarga.

No caso em que não existe a distribuição o tempo de espera após o ponto de sobrecarga se comporta de forma exponencial, já no caso em que ocorre a distribuição da rede o tempo aumenta de forma mais devagar.

O gráfico também tem em sua composição uma barra de erro que corresponde a um desvio padrão em cada marcador e em alguns gráficos, como no anterior e nos seguintes em que não se vê esta barra é porque esta barra é menor do que o marcador.

Outros valores a serem observados neste tipo de comparação seriam tamanho de fila médio e o tempo ocioso médio e estas comparações são apresentadas nas Figura 2 e Figura 3.

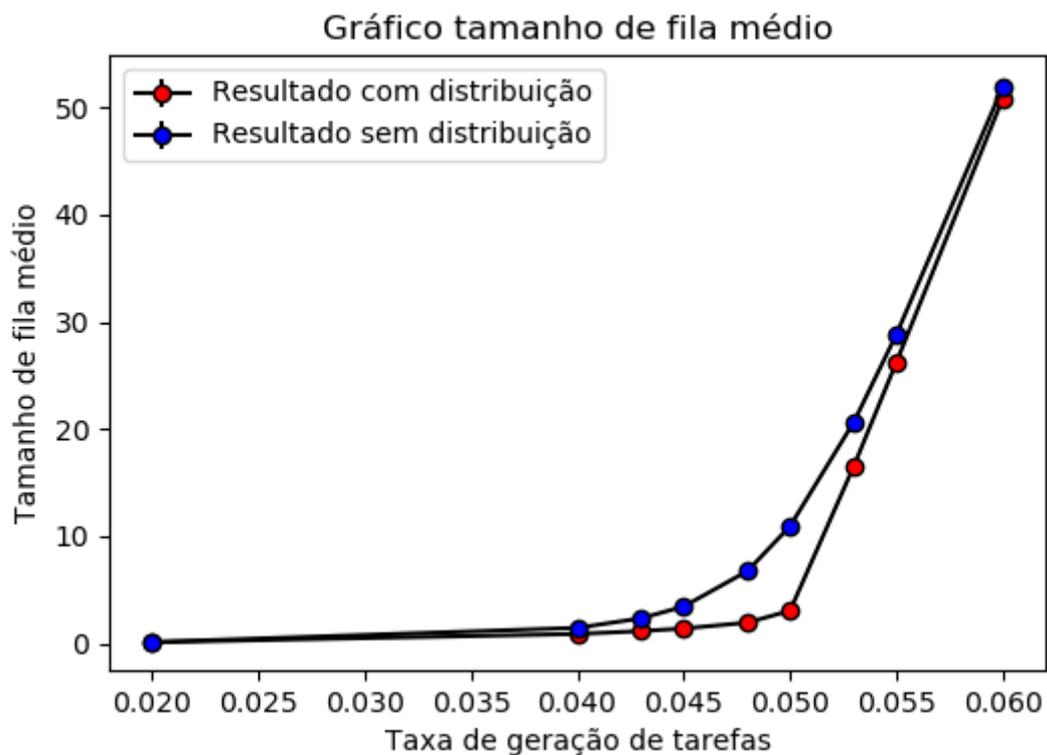


Figura 7 - Gráfico do tamanho de fila médio por ρ para uma rede sem distribuição e um resultado com distribuição na topografia ER.

Fonte: Elaborada pelo autor

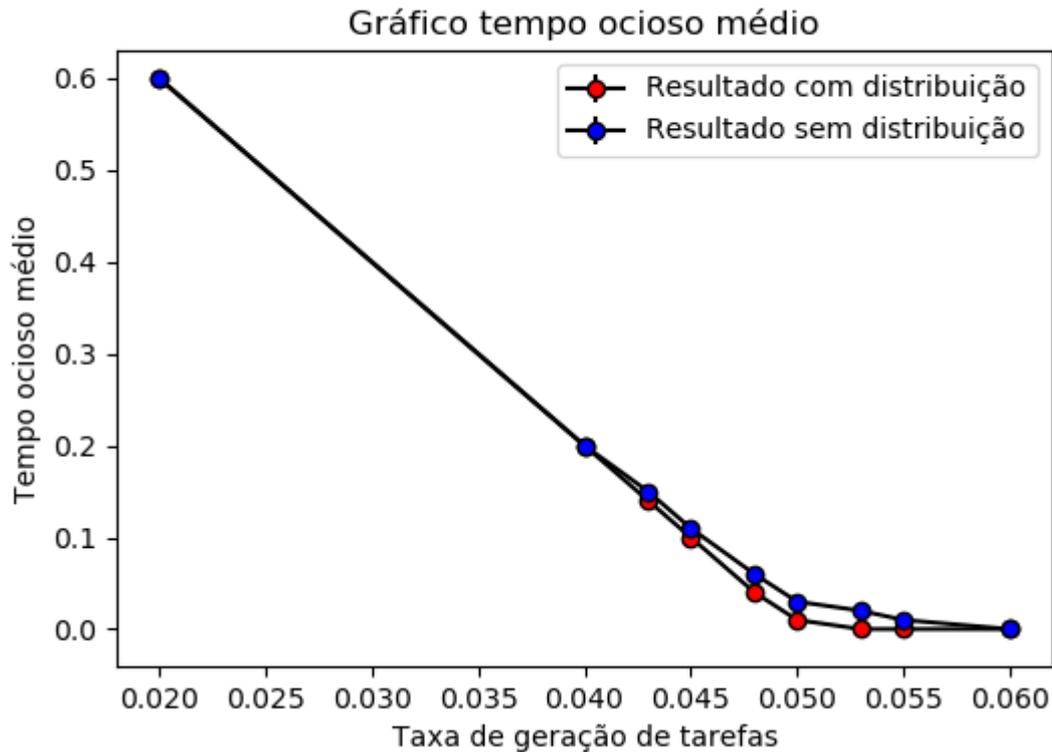


Figura 8 - Gráfico do tempo ocioso médio por ρ para uma rede sem distribuição e um resultado com distribuição na topografia *ER*.
 Fonte: Elaborada pelo autor

Ao analisar os Figura 7 e Figura 8, é perceptível que por volta do ponto $\rho_c = 0.05$ o tempo médio, tamanho de fila médio e tempo ocioso médio, existe uma leve diferença onde a rede com distribuição apresenta um melhor rendimento.

No tamanho de fila médio, as filas são menores do que em uma rede sem distribuição. Já para o tempo ocioso médio para um nó pertencente a uma rede com distribuição ele tem menos tempo de ociosidade. Uma observação a ser apresentada aqui ao analisar os gráficos é que apenas ao redor da faixa de 0,04 e 0,06, ou seja, próximo à taxa crítica que temos uma diferença entre o resultado sem distribuição e a rede *ER*, fora desta faixa os resultados são iguais, isso pode ser explicado porque para baixa taxa de geração de tarefas a fila está sempre vazia e então não é necessário uma distribuição das tarefas, já no caso de uma alta taxa de geração de tarefas as filas estão sempre lotadas. logo não adianta distribuir novamente, pois todos os outros nós estão lotados

4.2 Efeito do grau médio

Após feita a análise de comparação em relação o efeito da distribuição com uma rede sem uma distribuição, iremos analisar o efeito do grau médio dentro da rede com distribuição. para melhor visualização deste efeito foi usado a topografia de rede *ER*.

Para a realização desta simulação os parâmetros de número de nós foi mantido em 1000 (mil), foi usado o intervalo de $[0,02; 0,06]$ para a taxa de geração de tarefas e foram usados os valores de grau médio são 2, 4, 10, 14, 20, 30 e 100 . O resultado para esta simulação é mostrado na Figura 9.

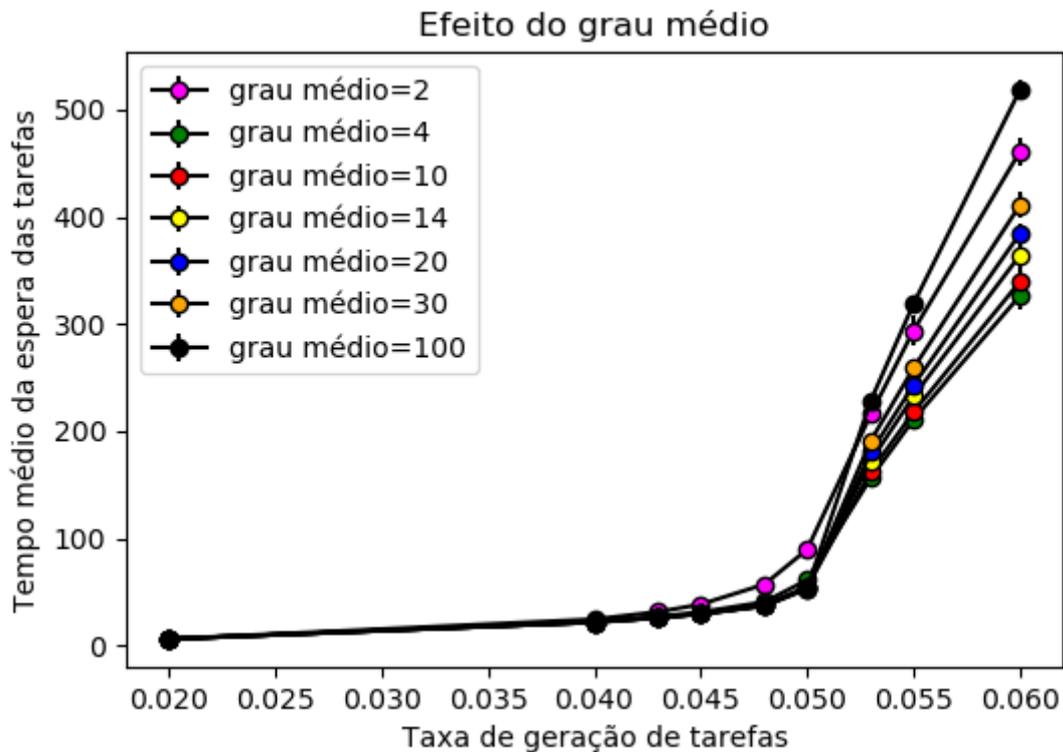


Figura 9 - Gráfico de tempo médio de espera das tarefas por ρ para uma rede com distribuição topografia *ER*.

Fonte: Elaborada pelo autor

Por este gráfico comparativo com a finalidade de saber a interferência do grau médio percebe-se que o grau médio tem efeito sobre o tempo médio da realização da tarefa, mas pode se observar que o grau médio não afeta de modo monotônico, ou seja, não está sempre diminuindo ou sempre aumentando, esse comportamento é observado para os valores elevado para a taxa de geração de tarefas.

Sendo que, analisando o gráfico, o melhor grau médio que apresenta os melhores resultados é o grau médio igual a 4.

Um ponto a ser observado é que o menor valor do grau médio e seu maior valor, sendo o valor 1 e 50, respectivamente, os que apresentam o maior tempo médio dentro dos valores analisados para observar o comportamento em relação ao grau médio.

Analisando agora em torno do ponto ρ_c , para ver o comportamento do grau médio, foi plotado o gráfico com o intervalo da taxa de geração de tarefas vai de [0,046; 0,054] e obtivemos o seguinte gráfico.

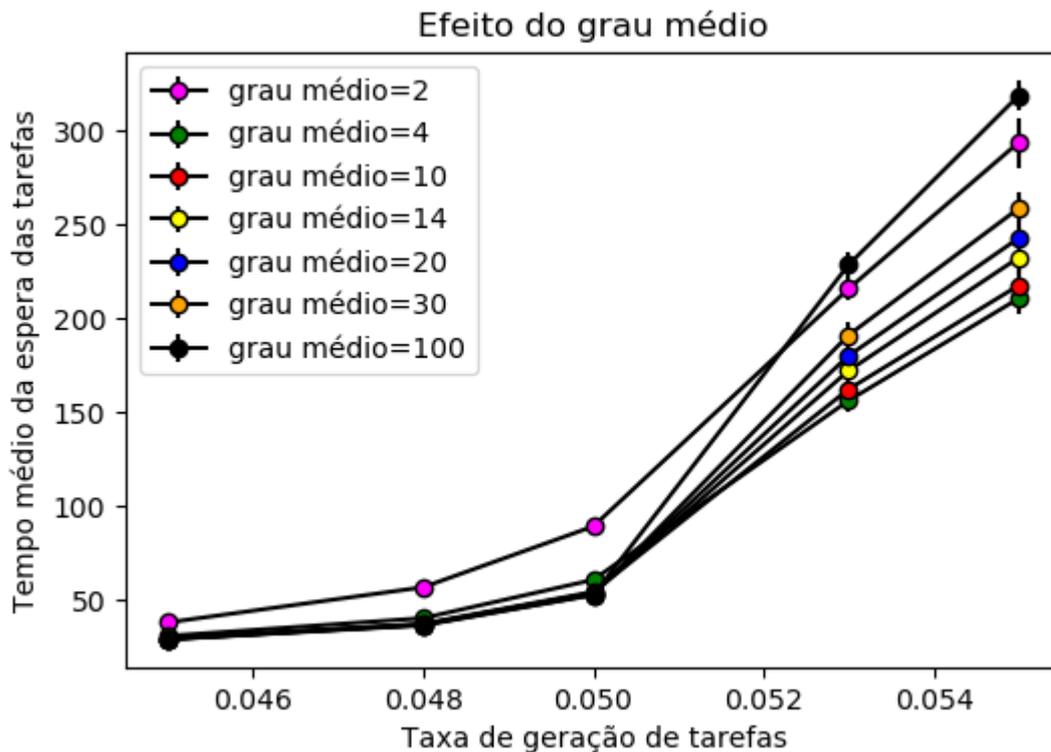


Figura 10 - Gráfico ao redor do ponto ρ_c pelo tempo médio para rede ER.
Fonte: Elaborada pelo autor

Para os valores de taxa maiores que 0,052 é observado que a desvio padrão destes pontos estão presentes e quanto maior o valor da taxa mais a barra aumenta em relação à taxa anterior daquele grau médio. dentro o maior valor entre os graus médios o grau médio de valor $m=1$ é que se encontra com maior tamanho, como pode ser observado anteriormente na Figura 4.

No ponto ρ_c , é observado que quanto maior for o valor do grau médio da rede menor é o tempo médio da rede realizar a tarefa, o que mostra que o grau médio tem interferência no tempo médio na realização da tarefa. Isto pode ser observado melhor nos gráficos a seguir

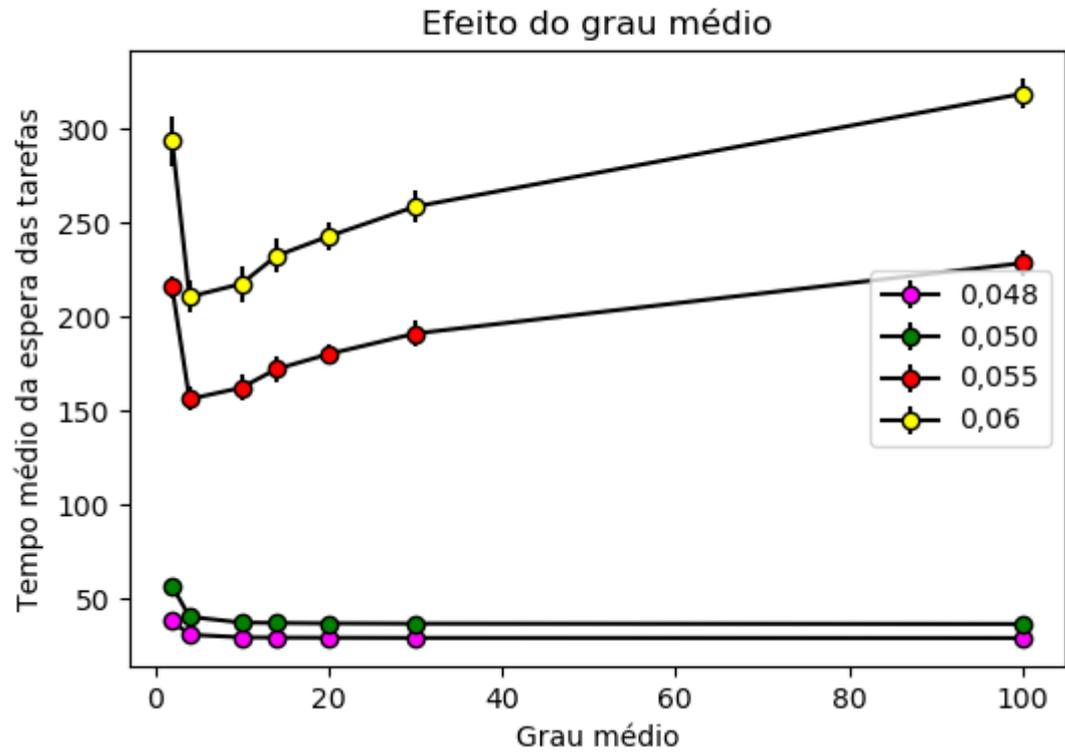


Figura 11 - Gráfico tempo médio de espera das tarefas por grau médio para uma rede ER;
Fonte: Elaborada pelo autor

Na Figura 11 pode ser observado que o comportamento antes e depois do ρ_c é diferente. Antes do ponto ρ_c o aumento de grau faz com que o tempo diminua e para depois do ρ_c o aumento do grau (acima do ponto 2) acarreta com que o tempo aumente.

Queremos mostrar o comportamento do tempo médio de espera das tarefas após a taxa crítica, e como é observado temos valores maiores de unidade de tempo médio para cada tarefa .

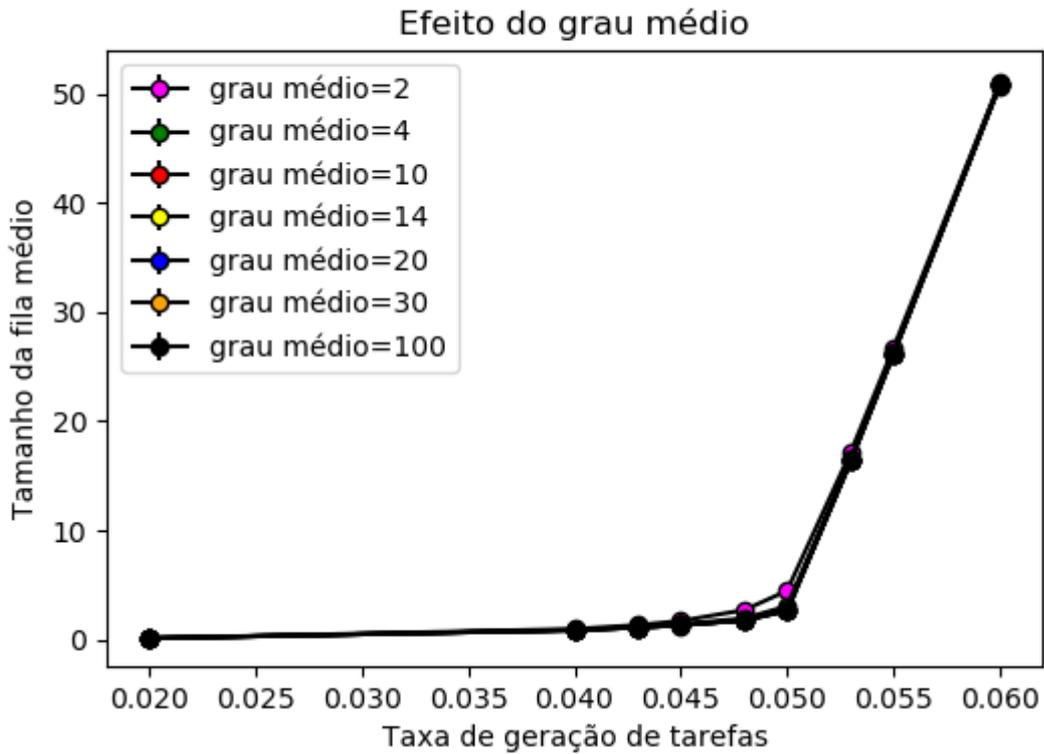


Figura 12 - Gráfico do tamanho de fila médio por ρ para uma rede de topografia de rede ER para a observação do efeito do grau médio na rede.

Fonte: Elaborada pelo autor

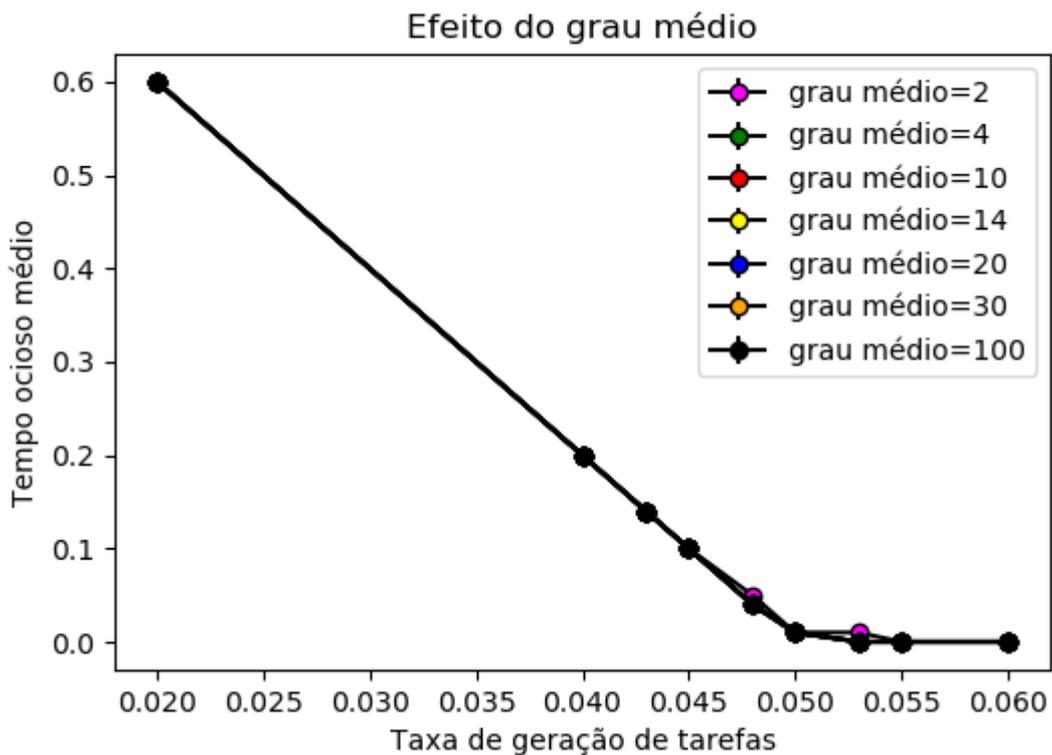


Figura 13 - Gráfico do tempo ocioso médio por ρ para uma rede de topografia de rede ER para a observação do efeito do grau médio na rede.

Fonte: Elaborada pelo autor

Ao observarmos aos dois últimos gráficos, se observa que o grau médio tem uma leve interferência na região do ponto ρ_c .

4.3 Efeito Heterogeneidade de graus

Após estudarmos qual o efeito do grau médio dentro de uma rede, agora um outro fato relevante para se existe uma interferência seria o efeito de heterogeneidade de graus em uma rede e para percebermos como atua vamos comparar as rede de topologia de *ER* com a rede de topografia de *BA*, para 3 valores de grau médio diferentes e esses valores de grau médio seriam 2, 5 e 10.

Para esta simulação foi mantido os parâmetros de número de nós igual a mil e os gráficos obtidos nestas comparações foram a Figura 14, Figura 15 e Figura 16.

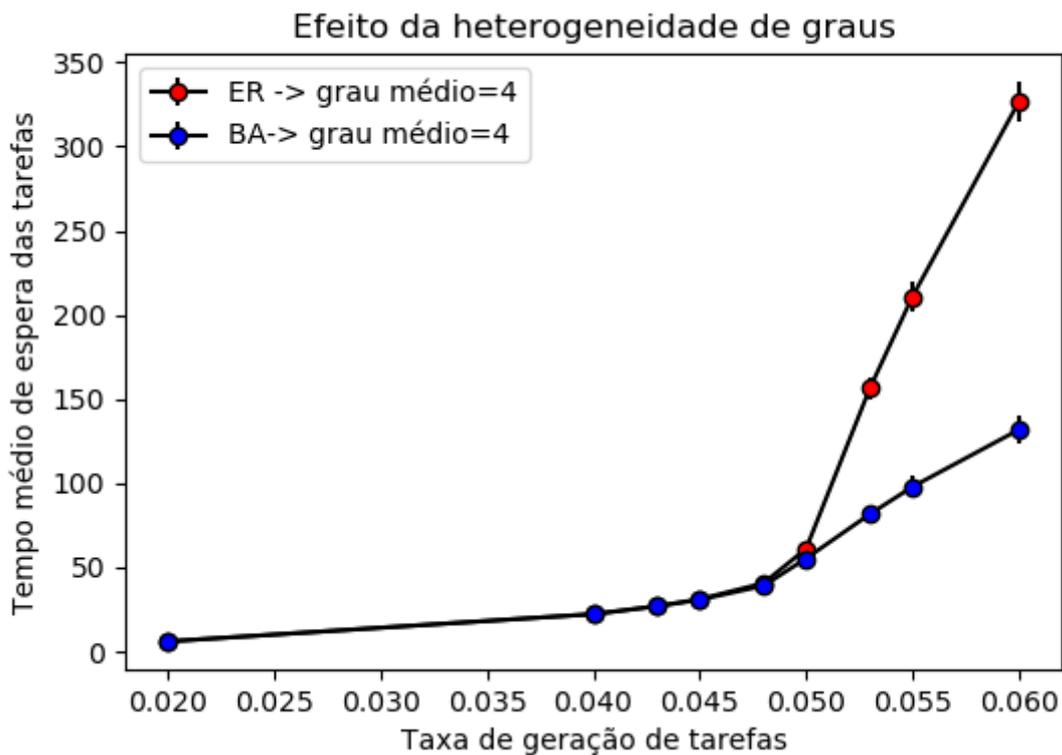


Figura 14 - Gráfico do tempo médio de espera das tarefas por ρ para comparação de efeito heterogeneidade de graus em rede *ER* e rede *BA* com grau médio = 4

Fonte: Elaborada pelo autor

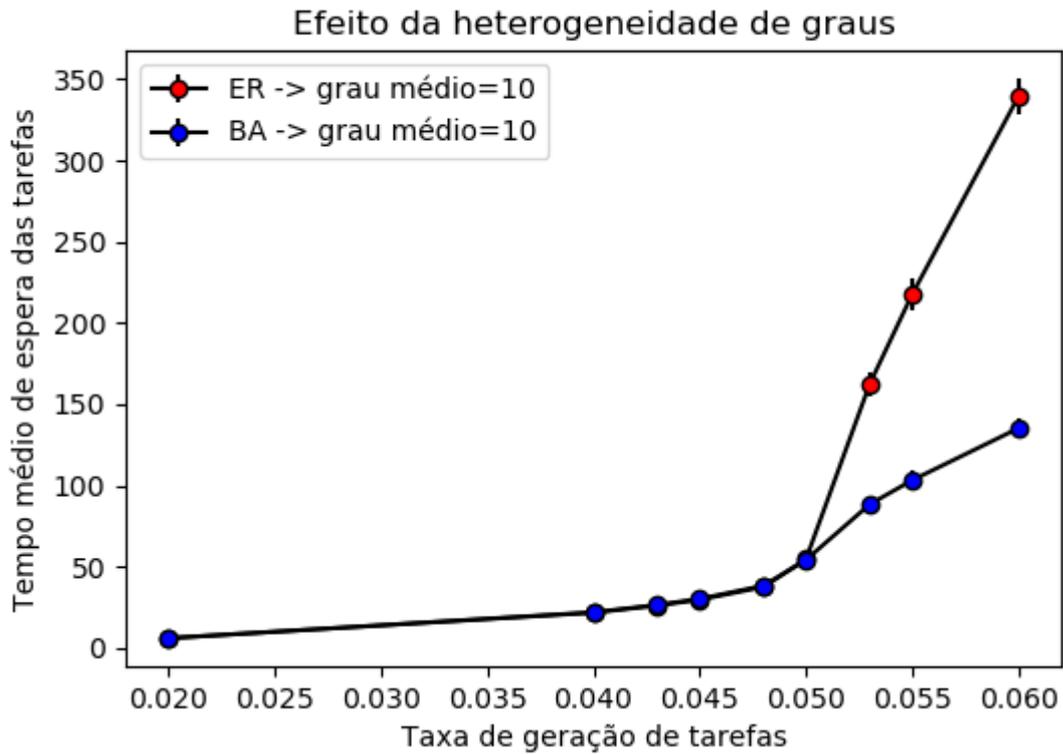


Figura 15 - Gráfico do tempo médio de espera das tarefas por ρ para comparação de efeito heterogeneidade de graus em rede *ER* e rede *BA* com grau médio=5.

Fonte: Elaborada pelo autor

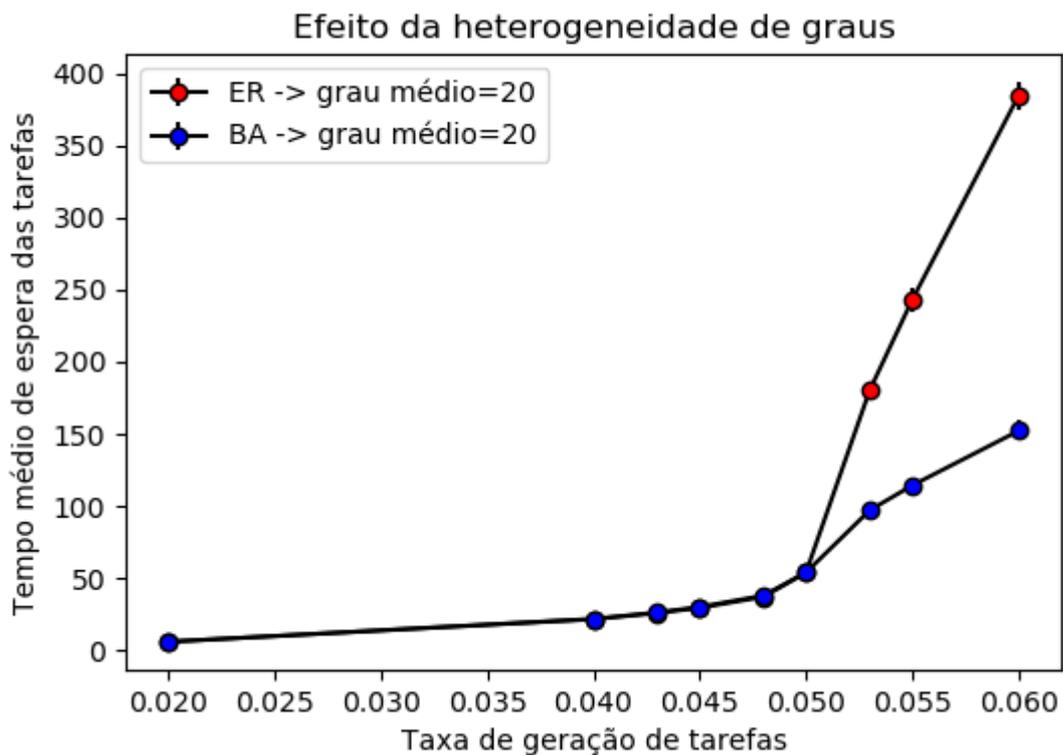


Figura 16 - Gráfico do tempo médio de espera das tarefas por ρ para comparação de efeito heterogeneidade de graus em rede *ER* e rede *BA* com grau médio=10.

Fonte: Elaborada pelo autor

Nos três gráficos anteriores, é observado que o efeito heterogeneidade de graus pode ser um dos fatores que ajudam a ter uma distribuição melhor para as tarefas para valores de geração altas, pois para as redes *BA* se tem uma melhor eficiência.

4.4 Efeito de coeficiente de aglomeração

Após a observação do efeito dos parâmetros de distribuição, do grau médio e o efeito heterogeneidade em uma rede, agora será explorado o efeito do coeficiente de aglomeração, nesta simulação vamos manter os parâmetros de número de nós e o grau médio, sendo nos valores de 1000 e 10, respectivamente. Foram obtidos a Figura 17 e Figura 18.

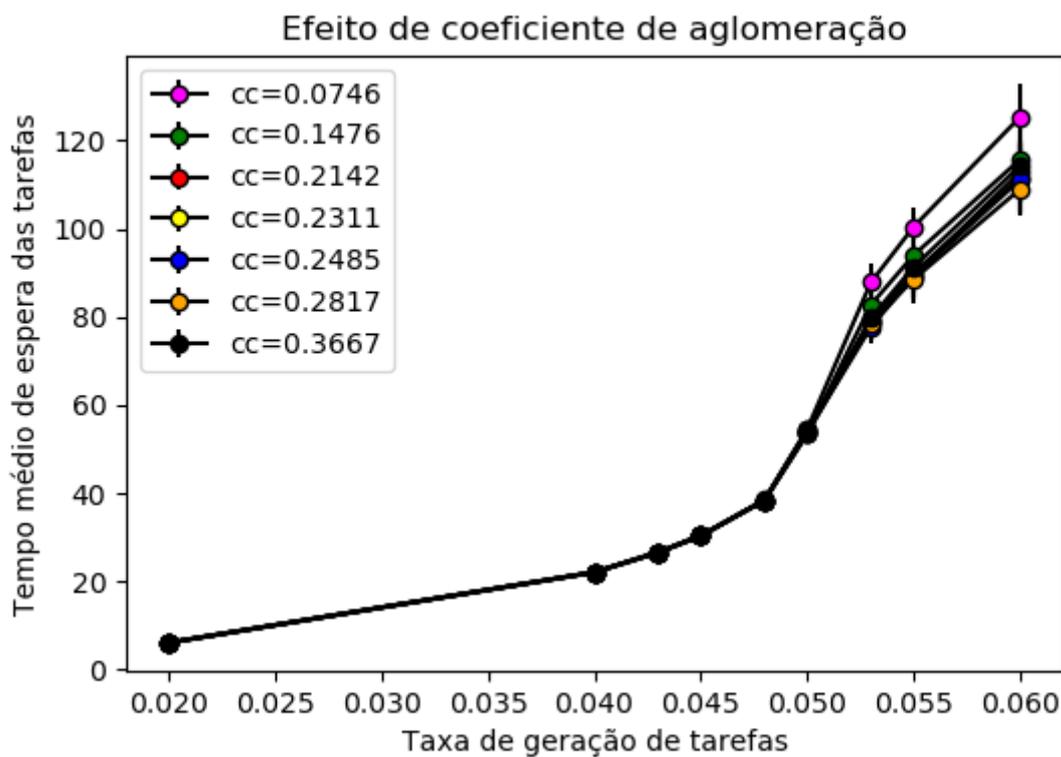


Figura 17 - Gráfico do tempo médio de espera das tarefas por ρ para comparação de efeito do coeficiente de aglomeração em uma rede *HK*

Fonte: Elaborada pelo autor

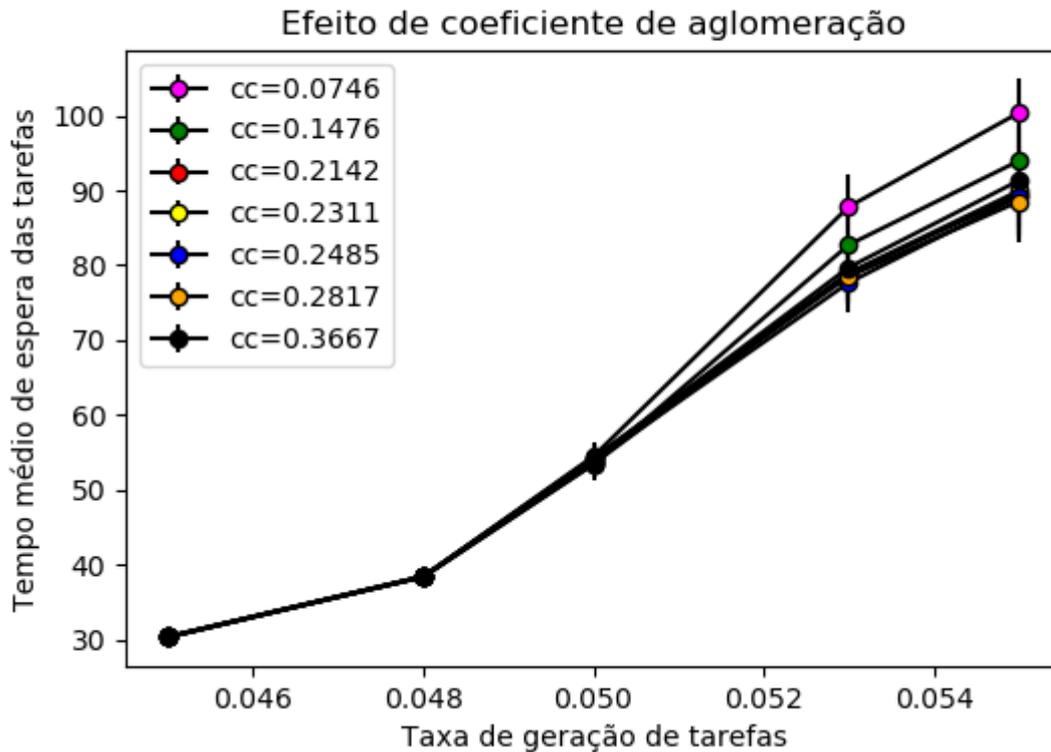


Figura 18- Gráfico do tempo médio de espera das tarefas por ρ para comparação de efeito do coeficiente de aglomeração em uma rede HK na região próxima ao ponto ρ_c .
 Fonte: Elaborada pelo autor

Pelas Figuras 17 e 18 pode ser observado que com coeficiente de aglomeração melhora a distribuição de tarefas na rede.

É possível observar também que o melhor valor de ρ para essa rede não é o maior valor, que neste caso seria $cc=0,366685$, mas sim o valor de $cc=0,248494$, uma hipótese possível para esse acontecimento seria que existe a presença de ciclos numa rede com coeficiente de aglomeração maior vai fazer com que as tarefas fiquem circulando de um nó para o outro, com mais dificuldade de encontrar um nó desocupado.

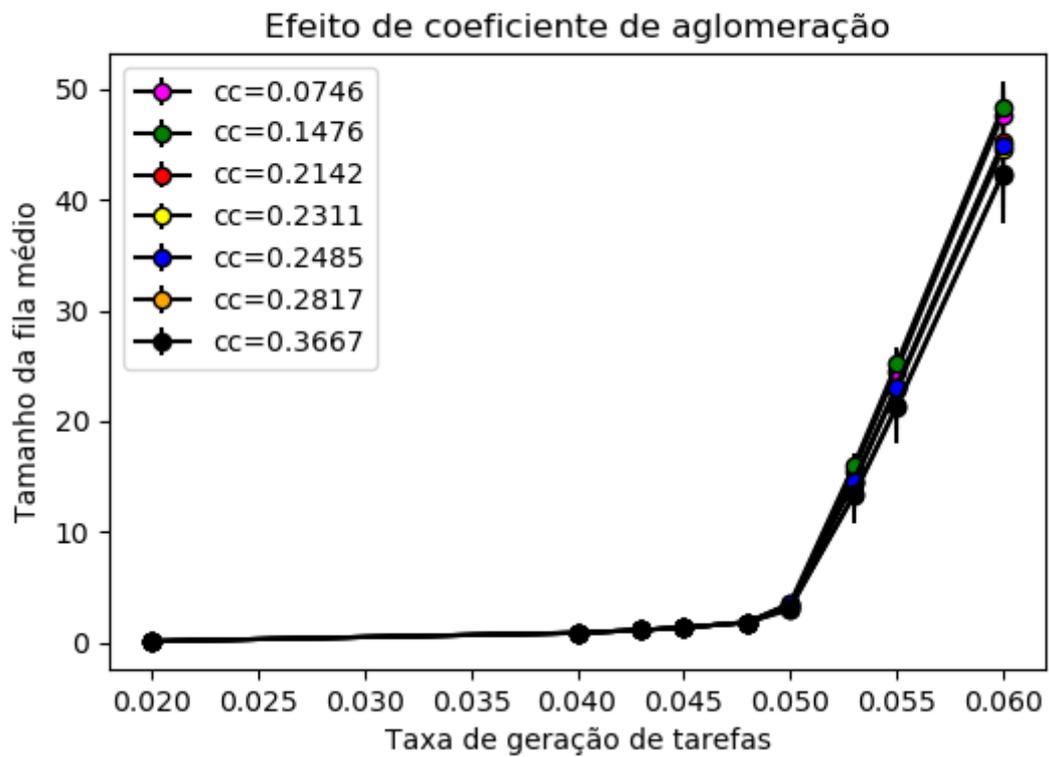


Figura 19 - Gráfico tamanho de fila médio por ρ para uma rede de topografia de rede *HK* para a observação do efeito do coeficiente de aglomeração na rede.
Fonte: Elaborada pelo autor

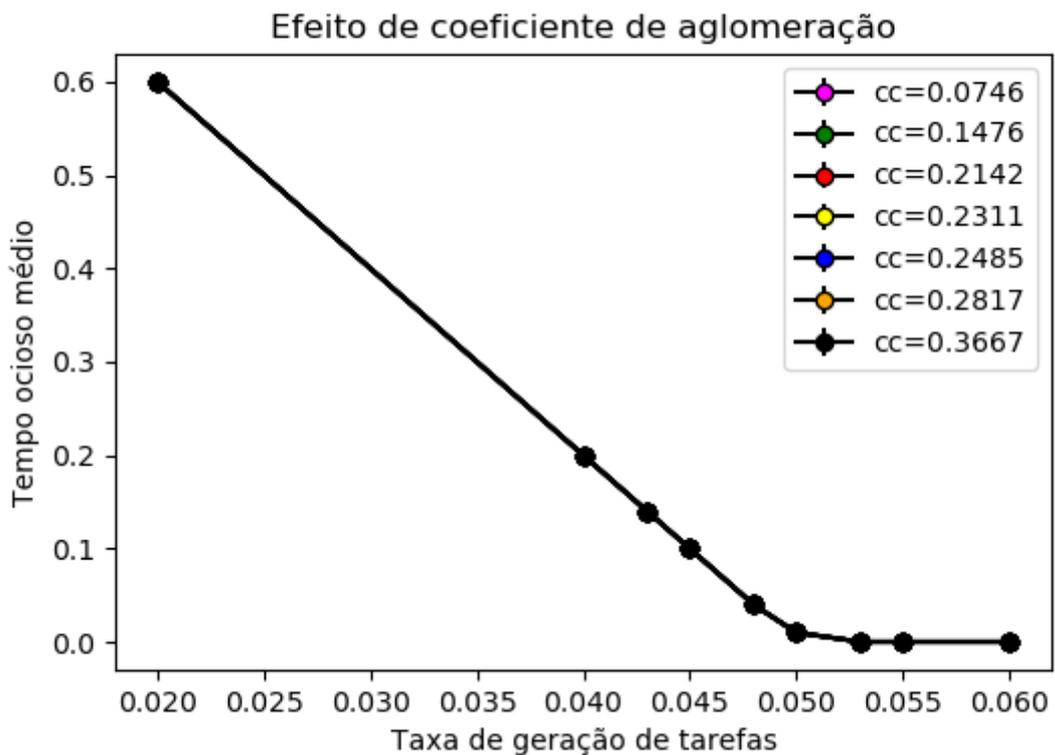


Figura 20 - Gráfico tempo ocioso médio por ρ para uma rede de topografia de rede *HK* para a observação do efeito do coeficiente de aglomeração na rede.
Fonte: Elaborada pelo autor

Na Figura 19, é possível a observação que após o valor crítico o tamanho da fila aumenta quanto maior o valor do coeficiente de aglomeração. Neste gráfico também é possível a observação da variabilidade após a taxa de geração passar o valor de ρ_c e seu tamanho aumentado conforme aumenta-se o valor da taxa de geração

Já no gráfico relacionado ao tempo ocioso médio não se observa uma diferença para valores distintos do coeficiente de aglomeração.

4.5 Efeito de estrutura de comunidade

Em seguida a observação do efeito de distribuição, do grau médio na rede, efeito de hub e o efeito de clustering, o próximo efeito que será explorado na influência no tempo médio é o Efeito de estrutura de comunidade em uma rede, e isso será observado em uma rede com topologia *LFR*.

Na realização desta simulação os parâmetros número de nós, grau médio e distribuição da rede se mantiveram iguais aos outros, sendo 1000, 10 e 4, respectivamente. Os parâmetros usados para gerar o modelo LFR foram o grau médio e μ que é a fração de arestas intercomunitárias incidentes a cada nó. Este valor deve estar no intervalo [0, 1].

E o gráfico obtido para essa simulação foi o gráfico que é apresentado na Figura 21.

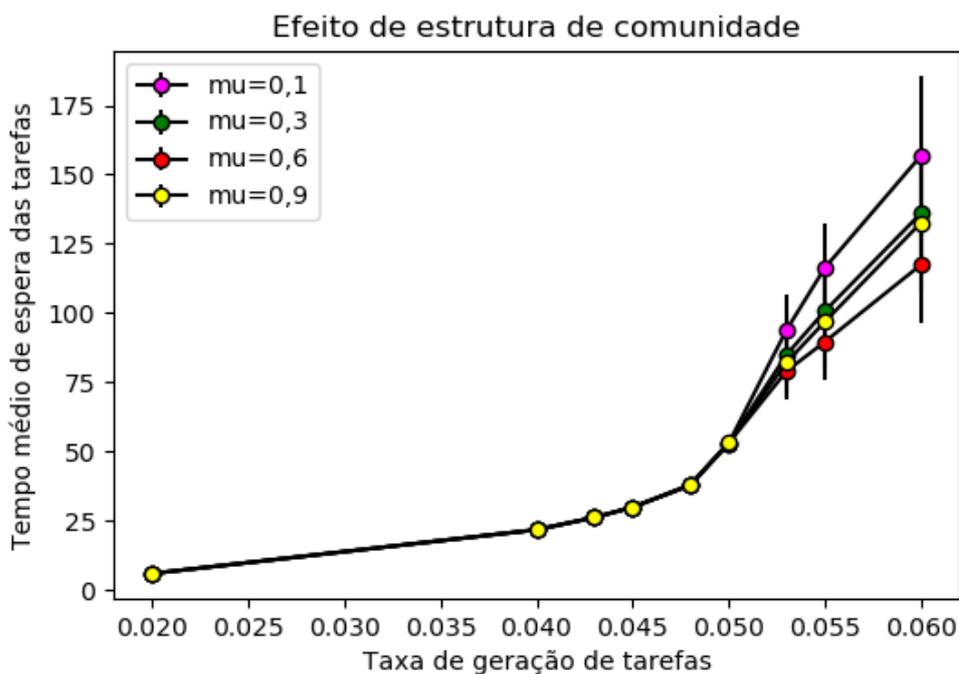


Figura 21 - Gráfico do tempo médio de espera das tarefas por ρ para comparação de Efeito de estrutura de comunidade em uma rede *LFR*

Fonte: Elaborada pelo autor

Pode ser observado que para altos valores ρ , o efeito comunidade tem uma grande interferência no tempo médio, pois quanto maior for a presença do efeito comunidade menor o tempo médio para a realização da tarefa.

No ponto ρ_c , e para pontos menores que eles a presença ou não de comunidade não altera o valor do tempo médio gasto para a realização da tarefa, como pode ser observado na Figura 22.

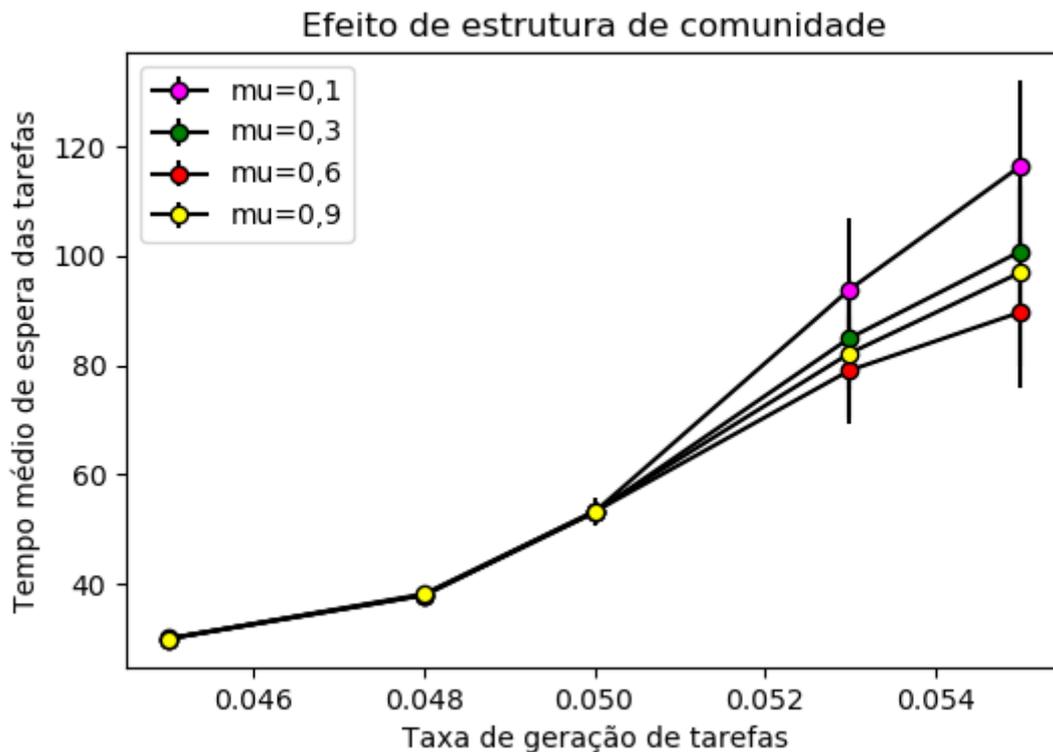


Figura 22 - Gráfico do tempo médio de espera das tarefas por ρ para comparação de Efeito de estrutura de comunidade em uma rede *LFR* na região próxima ao ponto ρ_c .

Fonte: Elaborada pelo autor

Nos últimos dois gráficos em comparação a outros gráficos referentes ao tempo médio dos outros efeitos estudados até agora se tem o tamanho da variabilidade muito maior do que os outros efeitos comparados e também já está presente mais perto do ponto ρ_c mesmo que em tamanho menor, portanto a variabilidade é maior em certos casos. Isto é provavelmente devido a grande variabilidade na estrutura de comunidade gerada pelo modelo *LFR*.

Nas Figura 23 e Figura 24. são apresentados as simulações para apresentar a interferência do Efeito de estrutura de comunidade para o tamanho de fila médio e tempo ocioso médio, respectivamente.

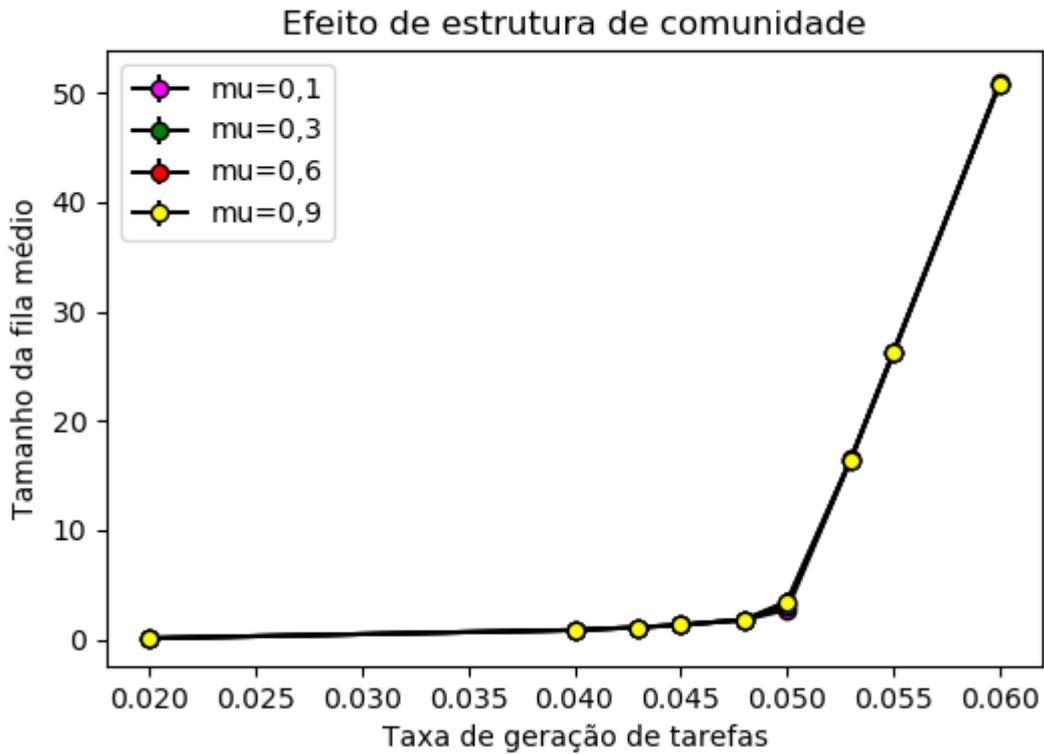


Figura 23 - Gráfico tamanho de fila médio por ρ para uma rede de topografia de rede *LFR* para a observação do efeito do grau médio na rede.

Fonte: Elaborada pelo autor

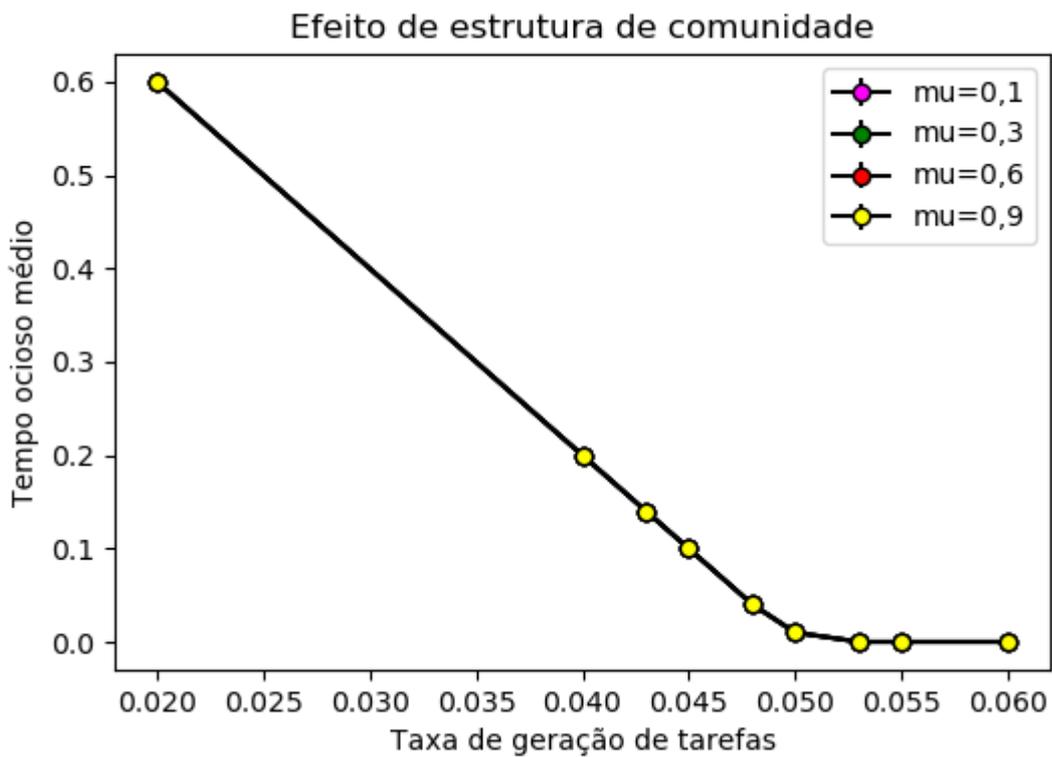


Figura 24 - Gráfico do tempo ocioso médio por ρ para uma rede de topografia de rede *LFR* para a observação do efeito do grau médio na rede.

Fonte: Elaborada pelo autor

Ao observarmos a Figura 23 se nota uma distinção entre os valores apresentados para o tamanho de fila médio e do mesmo modo na Figura 24 não se tem uma diferença para os valores de tempo ocioso médio.

5 Conclusão

Como foi apresentado no começo deste trabalho e discutido as características que seriam estudadas e levadas em consideração para análise após as simulações propostas para que possam ser analisadas o quanto é possível interferir na execução de tarefas em uma rede.

Com os dados que foram obtidos durante as simulações e análises ao plotar os gráficos percebe-se que as características, como efeito heterogeneidade de graus, Efeito de estrutura de comunidade, ou mesmo o efeito do grau médio, interferem no tempo médio de execução das tarefas, aumenta o tamanho médio na fila de espera, mas não se pode afirmar que esse fator que interfere unicamente para a diminuição destes tempos médios, pois podem existir outros fatores que podem influenciar também, mas que não foram levados em consideração neste trabalho.

Com os dados obtidos nas análises para o efeito de heterogeneidade de graus foi possível observar que existe um melhor valor de grau médio para execução, mas este valor não seria o de maior valor de grau médio e nem o de pior valor médio. Existe um valor intermediário que se encontra como o melhor valor para a execução da distribuição de tarefas colaborativas e com isso as tarefas tem um tempo médio de espera menor, ou seja, até este valor as tarefas acabam esperando menos para serem executadas, mas ao passar este valor as tarefas voltam a ter um tempo médio de espera maior.

Para a análise do efeito de coeficiente de aglomeração nota-se que o tempo de espera das tarefas aumenta conforme vai se aproximando do maior valor do coeficiente de aglomeração, mas após um certo valor o tempo médio acaba aumentando e as tarefas acabam demorando a serem realizadas.

No estudo referente ao efeito de comunidade apresenta um comportamento semelhante ao efeito de coeficiente de aglomeração, conforme a estrutura de comunidade se torna mais presente na rede o tempo médio de espera das tarefas vai diminuindo, mas após um certo valor para a taxa de geração de tarefas a tarefa volta a ter um valor elevado para o tempo médio de espera.

Para futuros trabalhos pode se pensar na possibilidade de estudar algumas outras características, como variação dos tamanhos das tarefas sendo não um valor fixo para todas. Uma outra sugestão seria existir um sistema de prioridades baseado no tamanho das tarefas ou se a tarefa foi gerada localmente ou recebida de vizinhos. Também seria possível a escolha de

outros algoritmos para a decisão de distribuição de tarefas. Pode-se também explorar a decisão de quando o agente se considera sobrecarregado.

REFERÊNCIAS

- 1 METZ, J. *et al.* **Redes complexas**: conceitos e aplicações. São Carlos: ICMC, 2007. Relatório técnico.
- 2 COSTA, L. da F. *et al.* Analyzing and modeling real-world phenomena with complex networks: a survey of applications. **Advances in Physics**, v. 60, n. 3, p. 329-412, 2011. DOI: 10.1080/00018732.2011.572452.
- 3 TRAVIESO, G. *et al.* A complex network approach to cloud computing. **Journal of Statistical Mechanics: theory and experiment**, v. 2016, n. 2, p. 023402, 2016. (modelagem de tarefas)
- 4 FORTUNATO, S. Community detection in graphs. **Physics Reports**, v. 486, n. 3-5, p. 75-174, 2010.
- 5 DE ANGELIS, A. F. *et al.* On the effects of geographical constraints on task execution in complex networks. **International Journal of Modern Physics C**, v. 19, n. 6, p. 847-853, 2008.
- 6 TRAVIESO, G. *et al.* Predicting efficiency in master–slave grid computing systems. **Journal of Complex Networks**, v. 1, n. 1, p. 63-71, 2013.
- 7 TRAVIESO, G.; COSTA, L. da F. Effective networks for real-time distributed processing. **Journal of Systems Science and Complexity**, v. 24, n. 1, p. 39-50, 2011.
- 8 COSTA, L. da F.; TRAVIESO, G.; RUGGIERO, C. A. Complex grid computing. **European Physical Journal B - condensed matter and complex systems**, v. 44, n. 1, p. 119-128, 2005. (modelagem de tarefas)
- 9 BOCCALETTI, S. *et al.* Complex networks: structure and dynamics. **Physics Reports**, v. 424, n. 4-5, p. 175-308, 2006.
- 10 NEWMAN, M. E. The structure and function of complex networks. **Siam Review**, v. 45, n. 2, p.167-256, 2003.
- 11 COSTA, L. DA F.; RODRIGUES, F. A.; G. TRAVIESO, G.; VILLAS BOAS, P. R. Characterization of complex networks: a survey of measurements. **Advances in Physics**, v. 56, n. 1, p. 167-242. DOI: 10.1080/00018730601170527.
- 12 HOLME, P.; KIM, B. J. Growing scale-free networks with tunable clustering. **Physical Review E**, v. 65, n.2, p.026107, 2002.
- 13 WATTS, D. J.; STROGATZ, S. H. Collective dynamics of ‘smallworld’ networks. **Nature**, v. 393, n. 6684, p. 440-442, 1998.
- 14 BOLLOBAS, B. Random graph. *In*: BOLLOBAS, B. (ed.). **Modern graph theory**. New York: Springer,1998. (Graduate texts in mathematics, v.184).

15 BARABASI, A. L.; ALBERT, R. Emergence of scaling in random networks. **Science**, v. 286, n. 5439, p.509, 1999.

16 GIRVAN, M.; NEWMAN, M. E. J. Community structure in social and biological networks. **Proceedings of the National Academy of sciences**, v. 99, n. 12, p. 7821-7826, 2002.

17 LANCICHINETTI, A.; FORTUNATO, S.; RADICCHI, F. Benchmark graphs for testing community detection algorithms. **Physical Review E**, v. 78, n. 4, p. 046110, 2008.