

MODELAGEM DA OPERAÇÃO FERROVIÁRIA ATRAVÉS DE REDES COMPLEXAS

Rômulo Murucci Oliveira, Eng.º, Universidade Federal de Juiz de Fora, Brasil.

Pedro Duarte Gomes Tostes, MEng, MRS Logística, Brasil.

RESUMO

O presente trabalho apresenta a modelagem da operação ferroviária de uma concessionária brasileira através de redes complexas, aproveitando-se do caráter cíclico da grade de formação dos trens de carga geral. O principal objetivo é aplicar estudos de ciência de redes para melhorar o entendimento do comportamento das operações ferroviárias. A metodologia foi implementada em linguagem *Python* e aplicada em um cenário de uma grade de trens conhecida, sendo o modelo operacional baseado em uma rede espaço-tempo com a utilização de restrições adicionais para encontrar o fluxo de custo mínimo de multiprodutos, estes representados no problema ferroviário pelas locomotivas dos trens. As principais propriedades oriundas da ciência de redes são estudadas e analisadas, como a distribuição de graus, a densidade do grafo, o coeficiente de agrupamento, a centralidade e a detecção de comunidades. Os resultados esperados foram alcançados, pois foi possível abordar e aplicar a ciência de redes para melhorar a compreensão sobre o comportamento da operação ferroviária.

Palavras-chaves: Operação ferroviária; Ciência de redes; Rede de multiprodutos; Carga Geral; Alocação de locomotivas.

1. INTRODUÇÃO

1.1. Grade de Trens de Carga Geral

A maior parte do volume de trens em circulação pode ser enquadrada em uma de duas categorias principais: trens unitários (aqueles formados por grandes lotes de vagões, que podem ou não ser de um mesmo tipo e podem ou não atender a um único cliente), que circulam normalmente sem grandes alterações em sua composição ao longo da viagem, parando estritamente entre as estações de origem e destino; e trens de carga geral (aqueles formados por lotes menores de vagões, comumente de tipos diferentes e atendendo a diversos clientes ao mesmo tempo), que circulam com um planejamento de paradas mais numerosas e contemplando diversas alterações em sua composição [Reis 2018].

Além das diferenças no que tange o modelo de formação dos trens, é importante diferenciar dois modos de operação encontrados nesta empresa: os chamados “trens de ciclo” operam obedecendo às programações baseadas na duração planejada de cada atividade e etapa do processo, enquanto os trens de grade operam obedecendo a uma tabela de programação (grade de trens) que delimita os horários de início e fim de cada atividade e etapa do processo [Tostes 2022]. A Figura 1 ilustra um exemplo hipotético de grade contendo dois trens, e as informações necessárias para programação, execução e controle das atividades previstas.

Grupo de Trens Carga Geral																
ID	Trem	Origem		Horário	Destino	Horário	Circulação	Seg	Ter	Qua	Qui	Sex	Sab	Dom	Fer	Observação
1	TREM_1	Pátio_01		13:10	Pátio_18	05:00	15:50	X	X	X	X	X				Especificidades de Carga, Itinerário e Excludência
	Pátio	Atividade		Chegada	Partida											
	Pátio_03	Pesar Composição		13:30	14:00											
	Pátio_06	Trocar Tripulação		16:15	16:30											
	Pátio_10	Aguardar faixa de circulação		23:15	01:00											
2	TREM_2	Pátio_17		22:00	Pátio_22	13:00	15:00	X		X		X			X	Especificidades de Carga, Itinerário e Excludência
	Pátio	Atividade		Chegada	Partida											
	Pátio_18	Trocar Tripulação		22:20	22:50											
	Pátio_19	Anexar e Retirar Vagões		03:30	09:00											
	Pátio_21	Pesar Composição		10:00	10:30											

Figura 1. Grade de trens hipotética contendo dois trens. Fonte: Tostes, 2022.

1.2. Redes Complexas

O termo redes complexas refere-se a um grafo que apresenta uma estrutura topográfica não trivial, composto por um conjunto de vértices (nós) que são interligados por meio de enlaces (arcos). Os arcos estabelecem algum tipo de relação entre dois vértices de acordo com o problema modelado. Além disso, o grafo pode ser direcionado (dígrafo) ou não [Barabási 2016].

O estudo de redes na forma de grafos é um dos pilares da matemática discreta e teve início em 1735, quando Euler propôs uma solução para o problema das pontes de Königsberg, originando a teoria dos grafos.

Os termos "grafo" e "rede" são considerados sinônimos na literatura científica. Entretanto, podemos estabelecer uma pequena distinção, em que as redes representam sistemas reais (como a sociedade sendo uma rede de indivíduos ligados pelas famílias) e o grafo é a representação matemática de tais sistemas (como um grafo social) [Barabási 2016].

É importante lembrar que nem todo grafo pode ser considerado uma rede complexa, pois essa classificação só é possível se o grafo apresentar algumas propriedades topográficas que não estão presentes em grafos simples [Metz et al. 2007]. Algumas dessas propriedades são descritas brevemente a seguir.

O grau de um nó i representa quantas ligações o mesmo possui com outros nós da rede. O grau médio para os grafos direcionados é calculado como: a soma dos arcos que apontam para o nó i e das arestas que apontam do nó i para outros nós da rede [Barabási 2016].

A distribuição de graus irá nos mostrar qual a probabilidade de que cada nó possua determinado grau na rede. Esta propriedade é essencial para avaliar os tipos de redes complexas [Metz et al. 2007].

As redes aleatórias, ou modelo Erdős-Rény (ER) são definidas a partir da ligação aleatória entre seus nós. A distribuição de graus se comporta como uma distribuição de Poisson [Barabási 2016].

Nas redes pequeno-mundo (*small-world*), ou modelo Watts-Strogatz (WS), grande parte das conexões são estabelecidas entre vértices mais próximos, mas algumas conexões aleatórias entre grupos são estabelecidas. Dessa forma, a distância entre dois nós aleatórios é pequena, ou seja, as redes WS possuem coeficiente de agrupamento (ou de clusterização) médio alto quando comparadas com as redes aleatórias [Watts and Strogatz 1998].

Nas redes livre escala (*scale-free*), ou modelo Barabási-Albert (BA), uma das principais características é denominada conexão de preferência (*preferential attachment*) que é a tendência de um novo vértice se conectar a um vértice da rede que tem um grau elevado de conexões. A distribuição de graus possui comportamento exponencial, portanto, ao plotar em um gráfico "log x log" a relação se aproxima de uma reta [Barabási 2016].

A densidade de um grafo quantifica se o mesmo é esparso (valores próximos de “zero”) ou completo (valores próximos de “um”), ou seja, representa se ele possui poucas ou muitas ligações entre os nós [Kaiser 2008]. O coeficiente de agrupamento representa se os vizinhos de um determinado nó estão fortemente conectados [Saramäki et al. 2007].

Um algoritmo de busca é qualquer algoritmo que visita todos os vértices de um grafo andando pelos arcos de um vértice a outro. Há muitas maneiras de fazer essa busca. Cada algoritmo de busca é caracterizado pela ordem em que visita os vértices. A busca em largura está intimamente relacionada com os conceitos de distância e caminho mínimo, e a busca em profundidade nos ajuda a encontrar ciclos e componentes fortemente conectados [Sedgewick 2002].

A centralidade de um grafo basicamente representa o grau de cada nó, sendo que para dígrafos, possuímos dois tipos: entrada e saída. Centralidade *betweenness* é alta quando o nó pertence a vários caminhos curtos do grafo, ou seja, funciona como um elemento de intermediação (ou “ponte”) da rede. Centralidade *closeness* é alta quando o nó está muito próximo dos demais nós da rede. *PageRank* representa a centralidade de um nó que é proporcional a soma das centralidades de seus vizinhos, para grafos direcionados [Barabási 2016].

Na ciência de redes, chamamos de comunidade um grupo de nós que têm maior probabilidade de se conectarem uns aos outros do que a nós de outras comunidades. Como exemplos, temos as redes sociais e as redes celulares (biológicas). Temos diversos algoritmos desenvolvidos com o objetivo de detectar comunidades, e o algoritmo de Louvain é um dos mais utilizados por possuir baixa complexidade computacional: ele emprega uma heurística iterativa baseada na maximização da modularidade [Blondel et al. 2008]. Valores de modularidade próximos de “zero” representam redes aleatórias, e valores maiores que 0,30 representam redes modulares, ou seja, com comunidades bem definidas [Newman and Girvan 2004].

A rapidez com que um vírus se propaga através de uma rede é conhecida como vulnerabilidade e é definida como uma medida de suscetibilidade à disseminação de entidades [Tong et al. 2010]. Em contrapartida, a robustez é definida como uma medida da capacidade de uma rede continuar a funcionar mesmo quando parte dela é naturalmente danificada ou alvo de um ataque. O fenômeno de maior interesse em termos de robustez é encontrar o limiar crítico do impacto das falhas dos nós na integridade de uma rede (f_c). Ao removermos uma fração “ f ” de nós, para qualquer $f < f_c$, a componente gigante se mantém, mas quando “ f ” ultrapassa “ f_c ” a componente gigante desaparece. [Barabási 2016].

2. DESENVOLVIMENTO

A representação do modelo operacional se baseia na criação de uma rede espaço-tempo com a utilização de restrições adicionais para encontrar o fluxo de custo mínimo de multiprodutos (*multicommodity network flow*) [Tostes 2022].

A primeira etapa consiste em adaptar os pátios ferroviários em uma representação de grafo e definir o passo temporal. A Figura 2 mostra um exemplo hipotético para preservar as informações da concessionária.

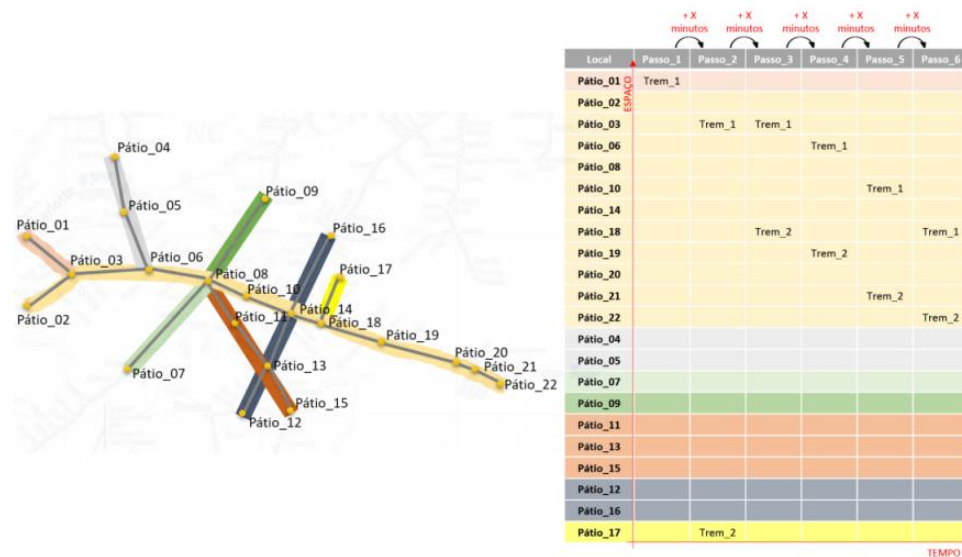


Figura 2. Grafo ferroviário e passo temporal. Fonte: Tostes, 2022.

Para criar a rede de espaço-tempo é necessário definir os movimentos viáveis entre os nós, que são determinados pela grade de trens vigente [Tostes 2022]. Dessa forma criou-se uma matriz de vizinhança que define quais nós possuem ligação entre si (Figura 3), e uma matriz dos tempos de viagem que define quantos passos são necessários para que a ligação entre os pátios vizinhos ocorra (Figura 4).

Local	Pátio_01	Pátio_02	Pátio_03	Pátio_06	Pátio_08	Pátio_10	Pátio_14	Pátio_18	Pátio_19	Pátio_20	Pátio_21	Pátio_22	Pátio_04	Pátio_05	Pátio_07	Pátio_09	Pátio_11	Pátio_13	Pátio_15	Pátio_16	Pátio_17
Pátio_01	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_02	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_03	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_06	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Pátio_08	0	0	0	1	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0
Pátio_10	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_14	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	1	0	0	1	0
Pátio_18	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0	0	0	1
Pátio_19	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0
Pátio_20	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0	0
Pátio_21	0	0	0	0	0	0	0	0	1	1	1	1	0	0	0	0	0	0	0	0	0
Pátio_22	0	0	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	0	0
Pátio_04	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Pátio_05	0	0	0	1	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0
Pátio_07	0	0	0	0	1	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
Pátio_09	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Pátio_11	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0
Pátio_13	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	1	1	1	1	0	0
Pátio_15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0
Pátio_16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	0	0
Pátio_17	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	1	1

Figura 3. Matriz de vizinhança. Fonte: Tostes, 2022.

Local	Pátio_01	Pátio_02	Pátio_03	Pátio_06	Pátio_08	Pátio_10	Pátio_14	Pátio_18	Pátio_19	Pátio_20	Pátio_21	Pátio_22	Pátio_04	Pátio_05	Pátio_07	Pátio_09	Pátio_11	Pátio_13	Pátio_15	Pátio_16	Pátio_17
Pátio_01	1	0	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_02	0	1	4	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_03	2	4	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_06	0	0	2	1	1	0	0	0	0	0	0	0	0	3	0	0	0	0	0	0	0
Pátio_08	0	0	0	4	1	2	0	0	0	0	0	0	0	0	2	2	4	0	0	0	0
Pátio_10	0	0	0	0	4	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Pátio_14	0	0	0	0	0	1	1	3	0	0	0	0	0	0	0	0	0	3	0	2	0
Pátio_18	0	0	0	0	0	0	4	1	1	0	0	0	0	0	0	0	0	0	0	0	3
Pátio_19	0	0	0	0	0	0	0	2	1	3	0	0	0	0	0	0	0	0	0	0	0
Pátio_20	0	0	0	0	0	0	0	0	4	1	4	0	0	0	0	0	0	0	0	0	0
Pátio_21	0	0	0	0	0	0	0	0	2	1	3	0	0	0	0	0	0	0	0	0	0
Pátio_22	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0
Pátio_04	0	0	0	0	0	0	0	0	0	0	0	1	3	0	0	0	0	0	0	0	0
Pátio_05	0	0	0	1	0	0	0	0	0	0	0	4	1	0	0	0	0	0	0	0	0
Pátio_07	0	0	0	0	1	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Pátio_09	0	0	0	0	3	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
Pátio_11	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	1	1	0	0	0	0
Pátio_13	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	2	1	1	4	0	0
Pátio_15	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	1	0	0	0	0
Pátio_16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	1	0	0
Pátio_17	0	0	0	0	0	0	4	0	0	0	0	0	0	0	0	0	0	0	0	1	0
Pátio_12	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	0	0	0	1

Figura 4. Matriz de tempos de viagem. Fonte: Tostes, 2022.

Considerando a influência das matrizes de vizinhança e de tempos de viagem, foi possível criar a matriz de adjacências. A Figura 5 traz a representação da rede espaço-tempo criada a partir da Figura 2 e da matriz de adjacências, ou seja, um grafo direcionado ponderado. Foram incluídos nós artificiais que representam a fonte (SF) e depósito (SD) de locomotivas, necessários para realizar o balanceamento do sistema em caso de sobra ou falta de locomotivas.

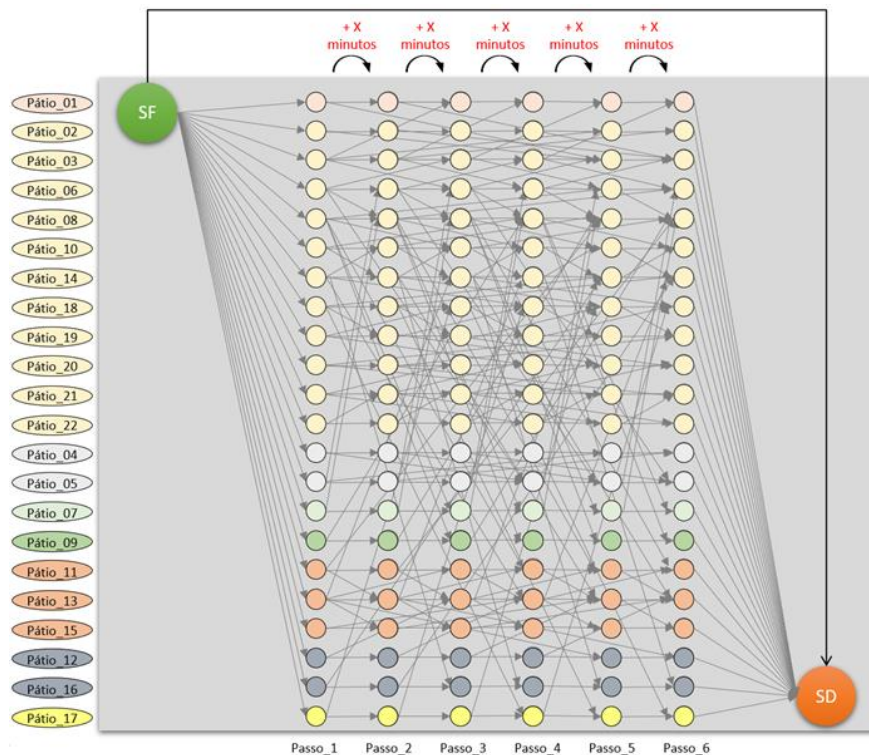


Figura 5. Rede espaço-tempo do exemplo hipotético. Fonte: Tostes, 2022.

O algoritmo com o pseudo-código da criação da rede espaço-tempo está na Figura 6. A matriz de adjacências possui todos os nós tanto nos índices das linhas quanto das colunas, dessa forma é possível diferenciar a direção da conexão. Só haverá conexão entre os nós caso o elemento correspondente da matriz esteja preenchido, em que a direção será definida como a linha sendo a origem e a coluna sendo o destino. O

valor de cada elemento da matriz representa o peso do arco. Note que a rede é representada por uma estrutura de dados de dicionário de dicionários (nós, antecessores e sucessores). A criação do grafo foi realizada através da matriz de adjacências com a linguagem de programação *Python*, e apoio das bibliotecas *NetworkX* [Hagberg et al. 2008] e *Pandas* [McKinney 2010].

Algoritmo 1: Criação da rede espaço-tempo

```

1: Entrada:  $M(m,n)$  = matriz de adjacências com 'm' linhas e 'n' colunas
2: Saída:  $G$  = grafo direcionado ponderado por 'e'
3:
4: início
5: cria dicionário nós
6:   para cada  $i \in \text{índices de } M[i][j]$  faça
7:     índice = índice de  $i$ 
8:     valor = 0
9:
10: cria dicionário antecessores
11:   para cada  $j \in M[i][j]$  faça
12:     índice  $\leftarrow$  índice de  $j$ 
13:     para cada  $i \in M[i][j]$  faça
14:       se  $e[i][j] \neq 0$  então
15:         valor  $\leftarrow$  índice = índice de  $i$ 
16:         valor =  $e[i][j]$ 
17:
18: cria dicionário sucessores
19:   para cada  $i \in M[i][j]$  faça
20:     índice  $\leftarrow$  índice de  $i$ 
21:     para cada  $j \in M[i][j]$  faça
22:       se  $e[i][j] \neq 0$  então
23:         valor  $\leftarrow$  índice = índice de  $j$ 
24:         valor =  $e[i][j]$ 
25:
26:  $G \leftarrow$  dicionário (nós, antecessores e sucessores)
27: fim

```

Figura 6. Pseudo-código para criação da rede espaço-tempo.

3. RESULTADOS E DISCUSSÕES

A rede possui 1586 nós, 4313 arcos e o grau médio é 5,44. A Figura 7 mostra a distribuição e o histograma de graus, na qual não se observa uma aproximação da distribuição de Poisson e nem de uma curva exponencial. A densidade do grafo vale 0,0017, um valor muito próximo de “zero”, ou seja, o grafo é muito esparsos.

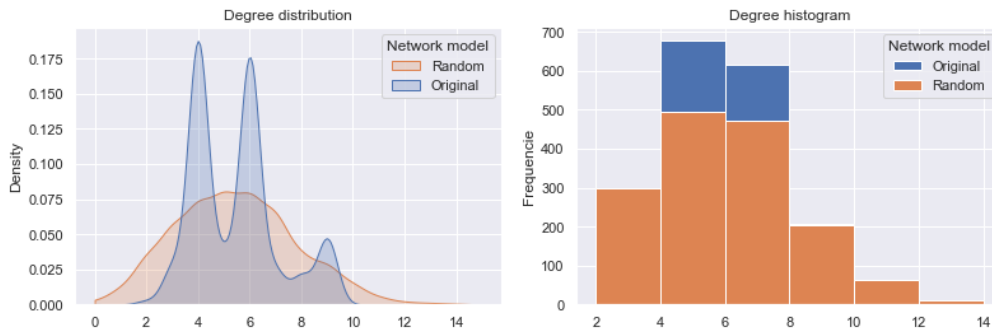


Figura 7. Distribuição e histograma de graus.

A Tabela 1 mostra os valores encontrados de coeficiente de agrupamento para a rede original, e para uma rede aleatória gerada com a mesma quantidade de nós e arcos (os pesos foram mantidos, mas embaralhados aleatoriamente pelos arcos). Foram utilizadas duas métricas, uma local [Watts and Strogatz 1998] e outra global [Newman 2003]. Observa-se que as redes original e aleatória possuem valores semelhantes e próximos de “zero”, portanto, o grafo original não se classifica como uma rede *small-world*. Caso a rede fosse classificada como WS, atenção especial se daria aos pátios que se conectam aos demais grupos, pois seriam elos importantes na cadeia de transportes ao encurtar distâncias entre os elementos da rede.

Tabela 1. Coeficientes de agrupamento por modelo de rede.

	Rede Original	Rede Aleatória
Coeficiente Local	0,000000	0,000001
Coeficiente Global	0,000000	0,001137

Foi executado o algoritmo de busca em largura para verificar se o grafo é conexo. Foram encontrados todos os nós, exceto o nó artificial do depósito, o que é esperado, pois a intenção é que não tenha sobra de locomotivas e, portanto, não tenhamos ligação com o depósito (os arcos que se ligam ao depósito possuem peso “zero”).

Ao calcular a centralidade *betweenness*, todos os nós do último passo temporal (23:40h), a fonte e o depósito possuem valor “zero”. Esse comportamento é esperado, pois esses são todos os nós que estão nas extremidades da rede, ou seja, não funcionam como elementos de ponte.

Os trinta maiores valores de centralidade *betweenness* foram ocupados por nós que representam o pátio de “FPU” com passos temporais que variaram de “06:20h” até “16:40h”. O comportamento temporal é explicado por ser próximo do valor central no dia (12:00h), o que gera mais conexões e os nós funcionam melhor como ponte; além disso, fisicamente o pátio de “FPU” é o que possui maior número de ligações com outros pátios, afinal ele é o pátio da grade de trens mais próximo do maior entroncamento ferroviário da América Latina.

Algo similar ocorreu quando se calcula a centralidade *closeness*, todos os nós do último passo temporal (23:40h) e o nó artificial de depósito possuem valor “zero”, afinal eles não se conectam com nenhum outro nó, ou seja, não possuem vizinhos sucessores (não possuem arcos saindo deles). A fonte possui maior valor de centralidade *closeness*, seguido por cinco valores com nós que representam o pátio de “FPU”.

Os nós fonte e depósito possuem os menores valores de centralidade *pagerank*, além de possuírem valores iguais, o que está de acordo com a teoria, pois são os únicos nós que não recebem nenhuma informação (arco de entrada).

O nó que representou o maior valor de centralidade *pagerank* é um pátio que recebe dois trens simultaneamente, ou seja, no mesmo espaço-tempo. Essa situação ocorre uma única vez em toda a rede espaço-tempo, e por isso explica o elevado valor para esse pátio e horário. Como consequência direta, nessa faixa de horário específica tem-se maior carga de trabalho que o normal, portanto, essa propriedade é útil para direcionar o planejamento dos recursos de manobra do pátio, o que reduz custos de pessoal e otimiza a alocação de equipes, além de garantir maior número de manobreadores e maquinistas no horário necessário para manter a fluidez da operação ferroviária.

Para a detecção de comunidades foi utilizado o algoritmo de Louvain, o qual encontrou 286 comunidades diferentes e obteve valor de modularidade igual a 0,32. Observou-se que as comunidades foram geradas de acordo com a proximidade física entre os pátios, afinal esse é um pré-requisito para que tenha conexão entre eles. O valor da modularidade mostra que as comunidades foram bem definidas (maior que 0,30), ou seja, ela não é aleatória.

A Tabela 2 resume as comunidades por quantidade de nós que cada uma possui. O depósito está em uma comunidade de um único elemento por não possuir ligações com os demais nós da rede. A comunidade com maior quantidade de elementos (23) possui a fonte e todos os pátios do primeiro passo temporal ligados à ela.

Tabela 2. Quantidade de comunidades por quantidade de nós.

Quantidade de nós	Quantidade de comunidades
0	0
1	1
2	9
3	33
4	41
5	79
6	42
7	49
8	14
9	4
10	12
11	1
12	1
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	1

Das nove comunidades que possuem dois nós, sete possuem os mesmos dois pátios, só que com passos temporais distintos, pois apenas um trem realiza parada nesses pátios. Esse comportamento é um indicativo para o planejamento operacional:

como apenas um trem realiza parada nesses pátios, caso ocorra algum imprevisto na circulação do trem, clientes que são atendidos nos referidos pátios podem ser prejudicados. Essa característica pode direcionar mudanças no planejamento da grade de trens como, por exemplo, mais trens realizarem paradas nesses pátios, adicionar atendimentos alternativos para esses pátios/clientes, ou aumentar a confiabilidade de circulação do referido trem. Dessa forma, diminuem-se atrasos e impactos na cadeia logística do cliente, no faturamento do transporte e no desgaste da relação da companhia com o cliente.

Nota-se que as comunidades com maior quantidade de nós possuem os pátios que estão mais próximos do maior entrocamento da malha ferroviária, ou seja, aqueles que possuem mais conexões (e consequentemente quantidade maior de trens que param nos respectivos pátios) que os demais pátios. Esse padrão denota a importância estratégica do entroncamento ferroviário para a operação, além de reafirmar o seu comportamento de elemento ponte na rede. Logo, garantir a operação plena do entrocamento é imperativo para reduzir impactos de funcionamento da malha ferroviária e, com isso, direciona o planejamento estratégico da manutenção a aportar maiores recursos no entroncamento para minimizar riscos de falhas e ocorrências.

Na simulação de ataque na rede utilizou-se estratégia de ataque aleatório e direcionado, com remoção de 10% dos nós. Nos ataques direcionados utilizaram-se como referência para remoção dos nós o grau, a centralidade *betweenness* e a centralidade *closeness*. Para a análise foram utilizados como parâmetros os valores de fração inicial de nós em relação a componente gigante (initial), fração de nós em relação a componente gigante após a remoção (frac), a média dos caminhos mínimos (apl) e a diferença percentual dos caminhos mínimos em relação ao valor inicial (dif. % apl), conforme mostra a Tabela 3, com os valores para os primeiros 1% dos nós removidos da rede.

Tabela 3. Simulação de ataque em 1% dos nós da rede.

Qtd. nós removidos	initial	frac	Aleatório		Grau		Betweenness		Closeness	
			apl	dif. % apl	apl	dif. % apl	apl	dif. % apl	apl	dif. % apl
1	0,999	0,999	6940	0,0%	5997	0,0%	6995	0,0%	6939	0,0%
2	0,999	0,998	6941	0,0%	5999	0,0%	6999	0,1%	6941	0,0%
3	0,999	0,997	6941	0,0%	6001	0,1%	7001	0,1%	6943	0,1%
4	0,999	0,997	6941	0,0%	6003	0,1%	7003	0,1%	6945	0,1%
5	0,999	0,996	6941	0,0%	6004	0,1%	7005	0,2%	6947	0,1%
6	0,999	0,996	6941	0,0%	6006	0,1%	7007	0,2%	6948	0,1%
7	0,999	0,995	6941	0,0%	6008	0,2%	7009	0,2%	6950	0,2%
8	0,999	0,994	6948	0,1%	6010	0,2%	7012	0,2%	6952	0,2%
9	0,999	0,994	6951	0,2%	6011	0,2%	7015	0,3%	6954	0,2%
10	0,999	0,993	6956	0,2%	6015	0,3%	7019	0,4%	6956	0,2%
11	0,999	0,992	6958	0,3%	6017	0,3%	7021	0,4%	6958	0,3%
12	0,999	0,992	6963	0,3%	6018	0,3%	7023	0,4%	6960	0,3%
13	0,999	0,991	6963	0,3%	6020	0,4%	7025	0,4%	6962	0,3%
14	0,999	0,991	6966	0,4%	6023	0,4%	7028	0,5%	6963	0,3%
15	0,999	0,990	6984	0,6%	6026	0,5%	7030	0,5%	6966	0,4%

Entre as diferentes estratégias utilizadas para realizar o ataque na rede, utilizar a centralidade *betweenness* foi a mais eficaz, pois aumentou a média dos caminhos mínimos da rede em quase 25%, afinal são nós que funcionam como pontes na rede. Essa característica é importante para avaliação da resiliência da rede para casos de interrupções da malha ferroviária (por exemplo, decorrentes por condições climáticas adversas) que pode quantificar o montante de investimento em manutenção de acordo com o nível de integridade desejado.

A Tabela 4 mostra uma comparação da média dos caminhos mínimos antes e depois da remoção de todos os nós (10% dos nós). O limiar crítico da rede (f_c) possui valor de 0,80, o que denota a necessidade de remover 80% dos nós para a componente gigante desaparecer.

Tabela 4. Simulação de ataque em 10% dos nós da rede.

			Aleatório		Grau		Betweenness		Closeness	
Qtd. nós removidos	initial	frac	apl	dif. % apl	apl	dif. % apl	apl	dif. % apl	apl	dif. % apl
0	1,0	1,0	6938	0,0%	6938	0,0%	6938	0,0%	6938	0,0%
158	1,0	0,9	8106	16,8%	6805	-1,9%	8640	24,5%	7624	9,9%

4. CONCLUSÃO

A rede modelada não se aproxima de nenhum dos modelos clássicos estudados, pois a distribuição de graus não se assemelha à distribuição de Poisson e a modularidade é maior que 0,30 (não é um modelo ER); a distribuição de graus não se assemelha à curva exponencial (não é um modelo BA); e os valores de coeficiente de agrupamento da rede original e da rede aleatória são semelhantes (não é um modelo WS). Entretanto, ao realizarmos a discretização da distribuição de graus através de um histograma, podemos observar que os valores se aproximam da distribuição de Poisson (Figura 7). Podemos concluir que, apesar da rede não ser aleatória, ela se aproxima mais do modelo ER.

Verificou-se que o grafo é muito esparso, mas é conexo. A grade de trens foi criada com o intuito de existir pelo menos uma ligação entre os pátios, independentes de quantos passos temporais sejam necessários, pois eles estão fisicamente conectados.

Ao se calcular os valores de centralidade, notam-se valores *betweenness* e *closeness* similares para a maioria dos nós, exceto para a fonte que possui centralidade *betweenness* igual a “zero” e centralidade *closeness* alta, o que denota um nó que não é uma ponte mas que possui muitos vizinhos.

Apesar da rede ser modular (modularidade maior que 0,30), as comunidades não são fortemente conectadas, pois encontrou-se valores próximos de “zero” de coeficientes de agrupamento. Vale ressaltar que a modularidade da rede (0,32) está próximo do limite da definição de rede modular.

Nota-se que em todas as estratégias de ataque da rede a tendência inicial é de aumento da média dos caminhos mínimos, entretanto, após a remoção de 10% dos nós, apenas a estratégia de remoção por graus apresentou redução no valor da média dos caminhos mínimos.

Conclui-se que, as propriedades oriundas da ciência de redes com maior aplicação prática para o modelo em estudo são a centralidade *pagerank*, a detecção de comunidades, e a simulação de ataques na rede por centralidade *betweenness*. Com a centralidade *pagerank* encontram-se pátios estratégicos para a realização de

planejamento de recursos de manobras em horários específicos. Com as comunidades pode-se identificar o entroncamento ferroviário e os pátios mais próximos dele, que são também os elementos com maior característica de ponte; além disso, é possível avaliar trens que são vulneráveis para o sistema por serem os únicos que atendem determinados pátios. Com a simulação de ataques mediu-se a resiliência da rede. Dessa forma, com o estudo dessas propriedades temos um bom diagnóstico para validar o planejamento operacional e direcionar a estratégia de manutenção da malha ferroviária.

Em linhas gerais, os resultados evidenciam a utilidade da aplicação de estudos em ciência de redes para melhorar o entendimento e a compreensão do comportamento das operações ferroviárias.

Como recomendações para trabalhos futuros, sugere-se parametrizar a grade de trens para possibilitar a inserção de informações de horizonte operacional (ou seja, verificar o que realmente é programado diariamente frente ao planejado). Espera-se um ganho na otimização da frota de locomotivas, mas com um grande custo computacional.

Testar a utilização de lista de adjacências em detrimento da matriz de adjacências para comparar e verificar ganhos de desempenho computacional. Espera-se ganhos para matrizes esparsas [Hagberg et al. 2008].

Aumentar o passo utilizado (por exemplo, de 20 minutos para 1 hora), pois se espera ganho de desempenho computacional e sem perda na otimização da frota de locomotivas.

Acrescentar na simulação de ataque da rede pesos para os diferentes pátios a serem removidos, de acordo com parâmetros do planejamento operacional como, por exemplo, clientes que são mais estratégicos para atendimento. Além disso, podem-se utilizar ataques direcionados a trechos que já são conhecidos por apresentarem maior susceptibilidade à interrupções da malha ferroviária.

5. REFERÊNCIAS

- Barabási, A.-L. (2016). *Network Science*. Cambridge University Press.
- Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E. (2008). *Fast unfolding of communities in large networks*. Journal of Statistical Mechanics: Theory and Experiment.
- Hagberg, A. A., Schult, D. A. and Swart, P. J. (2008). *Exploring network structure, dynamics, and function using NetworkX*. In *Proceedings of the 7th Python in Science Conference (SciPy2008)*. Gael Varoquaux, Travis Vaught, and Jarrod Millman (Eds), 11-15, Pasadena, CA USA.
- Kaiser, M. (2008). *Mean clustering coefficients – the role of isolated nodes and leafs on clustering measures for small-world networks*. New Journal of Physics, 10.
- McKinney, W. (2010). *Data Structures for Statistical Computing in Python*. In *Proceedings of the 9th Python in Science Conference (SciPy2010)*. Stéfan van der Walt and Jarrod Millman (Eds), 445:56-61.
- Metz, J., Calvo, R., Seno, E. R. M., Romero, R. A. F., and Liang, Z. (2007). *Redes complexas: conceitos e aplicações*. Technical report, Instituto de Ciências Matemáticas e de Computação da Universidade de São Paulo.
- Newman, M. E. J. (2003). *The structure and function of complex networks*. SIAM Review, 45(2).

- Newman, M. E. J. and Girvan, M. (2004). *Finding and evaluating community structure in networks*. Physical Review E, 69(2).
- Reis, C. A. A. (2018). Alocação eficiente de locomotivas no transporte ferroviário de cargas. *Master's thesis*, Instituto Militar de Engenharia.
- Saramäki, J., Kivelä, M., Onnela, J.-P., Kaski, K., and Kertész, J. (2007). *Generalizations of the clustering coefficient to weighted complex networks*. Physical Review E, 75(027105).
- Sedgewick, R. (2002). *Algorithms in C, Part 5: Graph Algorithms*. Addison-Wesley/Longman.
- Tong, H., Prakash, B. A., Tsourakakis, C., Eliassi-Rad, T., Faloutsos, C., and Chau, D. H. (2010). *Studying mobile context-aware social services in the wild*. In *Proceedings of 2010 IEEE International Conference on Data Mining*, 1091-1096, Institute of Electrical and Electronics Engineers.
- Tostes, P. D. G. (2022). Metodologia para dimensionamento de frota da grade de carga geral para aplicação em uma concessionária do setor ferroviário. *Master's thesis*, Pontífica Universidade Católica do Rio de Janeiro.
- Watts, D. J. and Strogatz, S. H. (1998). *Collective dynamics of 'small-world' networks*. Nature, 393:440-442.

MODELING RAILWAY OPERATIONS USING COMPLEX NETWORKS

Abstract

This paper presents the modeling of railway operations by a Brazilian concessionaire using complex networks, leveraging the cyclical nature of the formation grid of general freight trains. The primary objective is to apply network science studies to enhance the understanding of railway operations behavior. The methodology was implemented in Python and applied to a known train grid scenario. The operational model is based on a space-time network with additional constraints to determine the minimum cost flow of multi-products, represented in the railroad problem by the locomotives of the trains. The paper examines and analyzes key properties of network science, including degree distribution, graph density, clustering coefficient, centrality, and community detection. The expected results have been achieved, demonstrating the application of network science to enhance comprehension of railway operations behavior.

Keywords: Rail operations; Network science; Multicommodity network flow; General cargo; Locomotive assignment.