

Universidade Federal de Juiz de Fora
Instituto de Ciências Exatas
Programa de Pós-Graduação em Ciência da Computação

Ana Cláudia Martins de Souza

**Centralidade de Tempo baseada em Conectividade em Grafos Variantes no
Tempo**

Juiz de Fora

2019

Ana Cláudia Martins de Souza

Centralidade de Tempo baseada em Conectividade em Grafos Variantes no
Tempo

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Juiz de Fora, como requisito parcial para obtenção do título de Mestre em Ciência da Computação.

Orientador: Alex Borges Vieira

Coorientador: Klaus Wehmuth

Juiz de Fora

2019

Ficha catalográfica elaborada através do programa de geração automática da Biblioteca Universitária da UFJF, com os dados fornecidos pelo(a) autor(a)

Souza, Ana Cláudia Martins de.

Centralidade de Tempo baseada em Conectividade em Grafos Variantes no Tempo / Ana Cláudia Martins de Souza. -- 2019.
87 f. : il.

Orientador: Alex Borges Vieira

Coorientador: Klaus Wehmuth

Dissertação (mestrado acadêmico) - Universidade Federal de Juiz de Fora, Instituto de Ciências Exatas. Programa de Pós Graduação em Ciência da Computação, 2019.

1. Centralidade de Tempo. 2. Grafos Variantes no Tempo. 3. Conectividade. I. Vieira, Alex Borges, orient. II. Wehmuth, Klaus, coorient. III. Título.



Ana Cláudia Martins de Souza

“Centralidade de Tempo baseada em Conectividade em Grafos Variantes no Tempo”

Dissertação apresentada ao Programa de Pós-Graduação em Ciência da Computação da Universidade Federal de Juiz de Fora como requisito parcial à obtenção do grau de Mestre.

Aprovada em 29 de março de 2019.

BANCA EXAMINADORA

Prof. Dr. Alex Borges Vieira – Orientador
Universidade Federal de Juiz de Fora

Dr. Klaus Wehmuth - Coorientador

Prof. Dr. Heder Soares Bernardino
Universidade Federal de Juiz de Fora

Prof. Dr. Artur Ziviani
Laboratório Nacional de Computação Científica

Dedico este trabalho a Deus, ao meu esposo Roberto, aos meus pais José Claudio e Rosa Maria e a minha irmã Aracelle, que me forneceram sustento, fortaleza e perseverança para vencer essa etapa.

AGRADECIMENTOS

Inicialmente, agradeço ao meu professor e orientador Alex Borges pelo apoio, pelos incentivos, pelas ideias, pelo conhecimento, pela paciência, pelos momentos motivacionais, pelas correções, por acreditar que eu poderia realizar esse trabalho e me orientar. Sem seu constante apoio, eu não teria chegado até esse momento.

Agradeço também ao meu professor e coorientador Klaus Wehmuth por tantos ensinamentos, pelos exemplos, por repetir explicações de formas diferentes, por acreditar que daria certo, pelos direcionamentos, pelas diversas correções no meu texto, por todo apoio. Muito obrigada!

Meus sinceros agradecimentos ao professor Artur Ziviani por participar das reuniões sempre que possível, por me acolher no grupo de pesquisa no LNCC, pelos diversos insights, pela disponibilidade, pela motivação e alegria em visualizar os resultados.

Agradeço ao meu esposo Roberto, que me incentivou todos os dias para eu concluir este mestrado, por ter me apoiado em todos os momentos de angústia, por me ouvir explicar todos os trechos que trabalho, por ter me auxiliado sempre que eu pedia ajuda (e foram várias vezes), por caminhar comigo. Sou muito grata por ter você ao meu lado!

Agradeço a minha família por toda paciência e ausência durante esse período. E também por tantas palavras de incentivo. Vocês fizeram parte dessa trajetória.

Agradeço a todos amigos que fiz nesses dois anos dentro do PPGCC. Em especial, agradeço ao Alexandre Tavares e ao Felipe Rooke pelas conversas, pelas ideias, pelas ajudas, pela companhia.

Agradeço a todos professores do PPGCC e aos do LNCC que contribuíram para minha formação.

Agradeço ao IF Sudeste MG - *Campus* Juiz de Fora pela autorização de afastamento para capacitação e apoio financeiro, sem os quais essa trajetória seria mais difícil. Agradeço a UFJF por ser minha morada mais uma vez, aonde tive várias experiências, aprendizados e boas lembranças. Ao PPGCC por todo suporte, pela oportunidade, pela estrutura e pela confiança depositada em meu potencial.

"Para tudo há um tempo,
para cada coisa há um momento.
As coisas que Deus fez são boas a seu tempo."
Eclesiastes 3, 1.11

RESUMO

Os Grafos Variantes no Tempo (GVTs) nos permitem estudar e entender melhor a dinâmica dos sistemas do mundo real, pois conseguem expressar as variações de suas conexões ao longo do tempo. A noção de centralidade em cenários de GVT, geralmente, se referem a métricas que avaliam a importância relativa dos nós ao longo do tempo na evolução da rede dinâmica complexa. A noção de Centralidade do Tempo, que avalia a importância relativa de instantes de tempo em redes complexas dinâmicas, ainda é pouco explorada. De fato, os poucos trabalhos sobre a centralidade do tempo baseiam suas descobertas no processo de difusão, por exemplo, identificando o melhor momento para receber uma mensagem ou o instante crucial para aplicar uma vacina em determinada população. Em contraste, neste trabalho, estudamos a Centralidade do Tempo sob a ótica da conectividade de rede. Nós propomos duas métricas para identificar o instante de tempo mais importante na rede. A primeira é a Centralidade de Grau do Tempo, que possui a ideia tradicional, mas calcula o número de ligações que cada instante de tempo possui. Já a segunda, Centralidade de PageRank do Tempo, classifica os instantes de tempo de acordo com o número de arestas entrantes e do peso carregado por cada aresta. Esta segue o princípio da métrica tradicional de PageRank. Modelamos uma rede de transporte público por meio de um GVT, no qual avaliamos as métricas propostas. Nos resultados obtidos validamos que as centralidades calculadas são eficazes para identificar e classificar os instantes de tempo mais importantes do sistema. Vimos também que é fundamental aplicar as métricas no conjunto de dados que representa a rede como um todo, pois assim será identificado o instante de tempo mais importante da rede completa. No cenário avaliado, visualizamos horários que se destacam em cada dia, como por exemplo, pelo número de partidas e chegadas.

Palavras-chave: Centralidade de Tempo. Grafos Variantes no Tempo. Conectividade.

ABSTRACT

Time-varying Graphs (GVTs) allow us to study and better understand the dynamics of real-world systems because they can express the variations of their connections over time. The notion of centrality in GVT scenarios generally refers to metrics that assess the relative importance of nodes over time in the evolution of the complex dynamic network. The notion of the time centrality, which evaluates the relative importance of time instants in dynamic complex networks, is still little explored. Indeed, the few works on the time centrality base their findings on the diffusion process, for example by identifying the best time to receive a message or the crucial time to apply a vaccine in a particular population. In contrast, in this work, we study the Centrality of Time from the point of view of network connectivity. We propose two metrics to identify the most important time in the network. The first is the Time Degree Centrality, which has the traditional idea but calculates the number of connections that each instant of time has. The second, Time PageRank Centrality, classifies the time instants according to the number of incoming edges and the weight loaded by each edge. This follows the principle of traditional PageRank metrics. We model a Public Transport Network through a GVT, in which we evaluate the proposed metrics. In the obtained results we validate that the calculated centralities are effective to identify and classify the most important instants of time of the system. We have also seen that it is fundamental to apply the metrics in the data set that represents the network as a whole because this will identify the most important time instant of the complete network. In the scenario evaluated, we visualize times that stand out each day, such as the number of departures and arrivals.

Key-words: Time Centrality. Time-varyng graphs. Connectivity.

LISTA DE ILUSTRAÇÕES

Figura 1 – Exemplo de um Grafo A.	19
Figura 2 – Exemplo de um Dígrafo B.	20
Figura 3 – Exemplo ilustrativo de três arestas MAG (WEHMUTH et al., 2017).	22
Figura 4 – MAG (WEHMUTH et al., 2017).	23
Figura 5 – Sub-MAG da Figura 4.	24
Figura 6 – MAG subdeterminado.	24
Figura 7 – Exemplo de <i>snapshot</i>	26
Figura 8 – Exemplo de intervalos de tempo contínuo.	27
Figura 9 – Exemplo de arestas espaciais e temporais.	28
Figura 10 – Exemplo de arestas temporais e mistas.	28
Figura 11 – Exemplo de um GVT W.	30
Figura 12 – A matriz de adjacência do GVT W (WEHMUTH et al., 2015).	32
Figura 13 – Tempo de Cobertura (COSTA et al., 2015).	34
Figura 14 – Cobertura com Restrição de Tempo (COSTA et al., 2015).	35
Figura 15 – Exemplo MAG S	38
Figura 16 – Exemplo numérico do Pagerank em um sistema pequeno.	39
Figura 17 – ECDF das durações das viagens.	44
Figura 18 – Estrutura de uma aresta ponderada.	46
Figura 19 – Tipos de Arestas.	47
Figura 20 – Exemplos de Arestas.	49
Figura 21 – Arestas Intermodais.	50
Figura 22 – Grau do Tempo – <i>Air, Ferry e Rail</i>	55
Figura 23 – Grau do Tempo – <i>Coach</i> e Todos juntos.	56
Figura 24 – Grau Total do Tempo combinando meios de transporte três a três.	57
Figura 25 – Classificação dos Graus do Tempo em todos tempos.	60
Figura 26 – PageRank do Tempo – <i>Air, Ferry e Rail</i>	62
Figura 27 – PageRank do Tempo – <i>Coach</i> e Todos juntos.	63
Figura 28 – Grafos.	65
Figura 29 – PageRanks dos instantes de tempo antecessores.	66
Figura 30 – Classificação dos PageRank do Tempo.	67
Figura 31 – PageRank do Tempo combinando meios de transporte três a três.	68
Figura 32 – Grau de Entrada do Tempo – <i>Air, Ferry e Rail</i>	74
Figura 33 – Grau de Entrada do Tempo – <i>Coach</i> e Todos juntos.	75
Figura 34 – Grau de Entrada do Tempo combinando meios de transporte três a três.	76
Figura 35 – Grau de Saída do Tempo – <i>Air, Ferry e Rail</i>	77
Figura 36 – Grau de Saída do Tempo – <i>Coach</i> e Todos juntos.	78
Figura 37 – Grau de Saída do Tempo combinando meios de transporte três a três.	79
Figura 38 – Maiores PageRanks que chegam no <i>Domingo, 8 h 35 min</i>	83

Figura 39 – Maiores PageRanks que chegam na <i>Terça – feira</i> , 6 h 10 min.	84
Figura 40 – Maiores PageRanks que chegam na <i>Quarta – feira</i> , 6 h 10 min.	85
Figura 41 – Maiores PageRanks que chegam na <i>Quinta – feira</i> , 6 h 10 min.	86
Figura 42 – Maiores PageRanks que chegam na <i>Sexta – feira</i> , 6 h 10 min.	87
Figura 43 – Maiores PageRanks que chegam no <i>Sabado</i> , 6 h 10 min.	88

LISTA DE TABELAS

Tabela 1 – Dimensões do Conjunto de Dados.	42
Tabela 2 – Informações das Durações das Viagens.	44
Tabela 3 – Valores do Aspecto Camada.	45
Tabela 4 – Arestas Construídas.	51
Tabela 5 – Locais com somente uma Aresta de Embarque ou de Desembarque. . .	52
Tabela 6 – Arestas após a subdeterminação.	53
Tabela 7 – Maiores valores de PageRank diários por Meio de Transporte.	61
Tabela 8 – Tipos de Arestas que conectam ao instante.	64
Tabela 9 – Graus de Entrada do Tempo.	80
Tabela 10 – Graus de Saída do Tempo.	81
Tabela 11 – Graus Totais do Tempo.	82

LISTA DE ABREVIATURAS E SIGLAS

GVT Grafos Variantes no Tempo

MAG *Multi-Aspect Graph*

LISTA DE SÍMBOLOS

α	Alpha
β	Beta
\subseteq	Está contido em ou igual a
$=$	Igual
\neq	Diferente
\in	Pertence
\times	Produto cartesiano
$\%$	Resto da divisão
Σ	Somatório
σ	Sigma
$ $	Tal que
ζ	Zeta

SUMÁRIO

1	INTRODUÇÃO	16
1.1	IMPACTO NA ÁREA	17
1.2	ORGANIZAÇÃO DA DISSERTAÇÃO	18
2	FUNDAMENTAÇÃO TEÓRICA	19
2.1	GRAFOS TRADICIONAIS	19
2.2	MAG	20
2.2.1	Definição	20
2.2.2	Propriedades	21
2.2.2.1	<i>Isomorfismo entre MAGs e Grafos Direcionados</i>	21
2.2.2.2	<i>Predecessor e Sucessor</i>	22
2.2.3	Sub-MAG	23
2.2.4	Subdeterminação	24
2.2.5	Centralidades em MAG	25
2.3	GVT	25
2.3.1	Modelos para representação de GVTs	26
2.3.1.1	<i>Baseados em snapshots</i>	26
2.3.1.2	<i>Baseados em intervalos de tempo contínuos</i>	27
2.3.1.3	<i>Baseados em arestas espaciais e temporais</i>	27
2.3.1.4	<i>Baseados em arestas temporais e mistas</i>	28
2.3.1.5	<i>Modelo de representação unificado</i>	28
2.3.1.6	<i>Utilização do MAG para Representação de GVTs</i>	30
2.3.2	Matriz de Adjacência para Representação Algébrica de GVTs	30
3	TRABALHOS RELACIONADOS	33
4	CENTRALIDADE DE TEMPO	36
4.1	CENTRALIDADE DE GRAU DO TEMPO	37
4.2	CENTRALIDADE DE PAGERANK DO TEMPO	39
5	APLICAÇÃO E ANÁLISE DE CENTRALIDADE DE TEMPO	42
5.1	CONJUNTO DE DADOS	42
5.2	MODELO PROPOSTO	44
5.2.1	Descrição do Modelo	44
5.2.2	Validação do Modelo	51
5.2.3	Subdeterminação	53
5.3	AVALIAÇÃO DAS CENTRALIDADES DE TEMPO	53

5.3.1	Metodologia	53
5.3.2	Centralidade de Grau do Tempo	54
5.3.3	Centralidade de PageRank do Tempo	60
6	CONCLUSÕES	69
	REFERÊNCIAS	71
A	GRÁFICOS DOS GRAUS DO TEMPO	74
B	TABELAS DOS PICOS DOS GRAUS DO TEMPO	80
C	OUTROS GRAFOS SIMULADOS	83

1 INTRODUÇÃO

Compreender o comportamento de sistemas do mundo real vem impulsionando estudos que investigam sua dinâmica através do uso de redes complexas. São chamadas redes complexas todas as redes que possuem grande volume de informações, tais como redes neurológicas, redes de transporte, redes de telefonia, internet, entre outras. Existem vários desafios que envolvem rede complexas. Um deles é representar a rede e extrair informações dela.

Os Grafos Variantes no Tempo (GVTs), também conhecidos na literatura como redes temporais ou dependentes do tempo, foram apresentados em várias pesquisas para compreender e representar melhor a dinâmica das redes complexas (CASTEIGTS et al., 2012; HOLME; SARAMÄKI, 2012; KOSTAKOS, 2009; WEHMUTH et al., 2015). Com o uso de GVTs conseguimos visualizar e estudar o comportamento dos sistemas do mundo real. Um GVT representa a informação temporal além dos vértices e arestas. Assim, é possível observar o surgimento e comportamento temporal desses vértices e arestas.

Para exemplificar o uso de um GVT consideremos uma rede telefônica. Cada aparelho seria um vértice e cada telefonema entre dois aparelhos seria uma aresta. Cada aresta só existe enquanto os dois aparelhos estiverem se comunicando. Assim, cada aresta carrega as informações de vértice de origem e de destino da ligação, instante inicial da comunicação e a duração da chamada. Além disso, podemos representar a disponibilidade de cada aparelho, indicando instantes nos quais ele pode realizar ligações. Nesse cenário é possível observar diversas características, tais como, qual momento do dia com mais ligações, qual dia e horário da semana possui menos aparelhos ativos, entre outros.

Como nos grafos tradicionais, o uso das métricas de centralidade possibilita identificar a importância relativa de vértices ou arestas. De fato, a noção de centralidade nos cenários de GVT geralmente ainda se refere a métricas que avaliam a importância relativa dos vértices ao longo da evolução temporal da rede complexa dinâmica.

Por exemplo, diferentes métricas de centralidade temporal foram apresentadas para analisar a importância dos vértices na evolução do tempo em GVTs (KIM; ANDERSON, 2012; NICOSIA et al., 2013). Os poucos trabalhos sobre a centralidade do tempo baseiam suas descobertas no processo de difusão, como encontrar o melhor momento para levar uma mensagem ou o melhor momento para aplicar a vacina em uma população. Por exemplo, a Centralidade de Tempo, proposta por (COSTA et al., 2015), procura o instante de tempo mais importante para realizar uma determinada ação. Ele apresentou duas métricas que calculam a centralidade de tempo baseada em difusão, nas quais são detectados esses momentos-chave para agir. Porém, a noção da centralidade de tempo ainda é pouco explorada.

Entretanto, no GVT, os instantes de tempo podem ser os vértices, ligados a outros

instantes de tempo através da estrutura de uma rede (WEHMUTH et al., 2015). Nesse caso, a noção da importância de um instante de tempo baseada em um processo de difusão pode não fazer sentido. Assim, deve-se analisar a importância do tempo com base na topologia do GVT, analisando a conectividade do tempo.

Neste trabalho, estudamos a Centralidade de Tempo sob a visão de conectividade da rede. Propomos duas métricas de Centralidade de Tempo baseadas em conectividade, ou seja, as Centralidades de Grau do Tempo e de PageRank do Tempo. Ambas as métricas são análogas às tradicionais Degree e PageRank, mas neste caso, nos concentramos no instante de tempo como um vértice do grafo e modelamos os eventos que correlacionam os diferentes instantes. Por exemplo, o Grau do Tempo refere-se a quantas ações (número de conexões para outros momentos no GVT) estão associadas a cada instante de tempo. Quanto maior o número de conexões, mais importante será o instante de tempo. Já o PageRank do Tempo leva em consideração o número de arestas que chegam até ele e o peso que cada uma dessas arestas carrega para classificar o instante de tempo mais importante. Assim como na métrica tradicional de PageRank, cada vértice terá um peso calculado que é a probabilidade de ser alcançado.

Avaliamos as métricas que definimos usando o conjunto de dados de uma rede de transporte público do mundo real. Mais precisamente, modelamos os serviços de transporte público da Grã-Bretanha (GALLOTTI; BARTHELEMY, 2015) como MultiAspect Graph (WEHMUTH et al., 2016), onde o tempo é uma das principais características (aspecto). Além do tempo, consideramos outros dois aspectos: o tipo de transporte e a localização da estação. Nossos resultados expressam instantes de tempo centrais tais como horários já conhecidos como importantes em uma rede de transporte, que indicam o momento do dia em que se deve prestar mais atenção ao sistema de transporte. Tradicionalmente, ao analisar a rede de transporte, faz-se uma análise individual de cada meio de transporte. Pelo fato de considerarmos o sistema de transporte como um todo e analisando diferentes características, obtivemos instantes de tempo importantes distintos da análise de cada meio de transporte separadamente.

1.1 IMPACTO NA ÁREA

Este trabalho busca contribuir com a ampliação de conceitos sobre Centralidade do Tempo em Redes Complexas Dinâmicas. Até o momento foram identificadas propostas de Centralidade do Tempo baseadas em processos de difusão. Nós propusemos duas métricas de Centralidade do Tempo baseadas em conectividade, sendo:

- **Centralidade de Grau do Tempo:** visa classificar os instantes de tempo do maior para o menor grau dentro de uma rede complexa. O cálculo do grau parte do princípio que a rede será modelada com o maior número de características possíveis,

incluindo o fator tempo, para representar melhor o cenário do mundo real. A partir do modelo da rede, é possível ranquear todos os instantes de tempo, possibilitando identificar o mais central, conforme o número de ligações com os outros instantes de tempo.

- **Centralidade de PageRank do Tempo:** visa ranquear todos instantes de tempo da rede e identificar o instante de tempo com maior importância de PageRank em uma rede complexa. Assim como no PageRank tradicional, cada instante de tempo adquire uma reputação que é transmitida percentualmente aos próximos instantes de tempo, possibilitando classificá-los do maior para o menor. Instantes de tempo importantes nessa métrica possuem duas características: são apontados por muitos instantes e/ou são apontados por outros instantes também importantes considerando suas pontuações.

1.2 ORGANIZAÇÃO DA DISSERTAÇÃO

Esse trabalho está organizado trazendo um referencial teórico contemplando grafos, dígrafos, Grafos MultiAspectos (MAGs), grafos variantes no tempo (GVT) e representações no Capítulo 2; trabalhos relacionados no Capítulo 3; o conceito de centralidade do tempo bem como das métricas baseadas em conectividade estão dispostas no Capítulo 4; o estudo de caso com aplicação e avaliação das centralidades calculadas no Capítulo 5; e conclusões e trabalhos futuros no Capítulo 6.

2 FUNDAMENTAÇÃO TEÓRICA

Esse capítulo tem como objetivo inserir o leitor no universo da pesquisa realizada, trazendo conceitos fundamentais seguidos de exemplos. Primeiro, serão conceituados grafos tradicionais na Seção 2.1. Em seguida, apresentamos a definição, características e propriedades do Grafo MultiAspecto (MAG) na Seção 2.2. Por fim, trazemos os Grafos Variantes no Tempo (GVTs), junto a suas representações algébricas e seus modelos de representação na Seção 2.3.

2.1 GRAFOS TRADICIONAIS

Grafos são largamente utilizados para representação de redes na literatura. Seus conceitos estão relacionados. Em (EVANS; MINIEKA, 1992) uma rede é um grafo com um ou mais números associados a cada aresta ou arco. Esses números podem representar distâncias, custos, confiabilidade ou outros parâmetros relevantes.

Um grafo G é tradicionalmente definido pelo par $(V(G), E(G))$ que satisfaz $E(G) \subseteq [V(G)]^2$, sendo $V(G)$ o conjunto dos vértices e $E(G)$ o conjunto de arestas. Cada aresta será representada por $\{a, b\} \in E(G)$ sendo $a, b \in V(G)$ (BONDY; MURTY, 2008; DIESTEL, 2010). A Figura 1 ilustra um grafo A composto pelo conjunto de vértices $V(A) = \{a, b, c, d, e, f, g\}$ e pelo conjunto de arestas $E(A) = \{\{a, b\}, \{a, e\}, \{b, e\}, \{c, d\}, \{e, g\}\}$. Esse é um grafo não-orientado, pois suas arestas não possuem direção. Assim, podemos afirmar que a aresta que liga o vértice a no vértice b é a mesma que liga o vértice b no vértice a . Além disso, sempre que um grafo admitir mais de uma aresta entre os vértices $\{a, b\}$ será chamado de multigrafo (BANG-JENSEN; GUTIN, 2007).

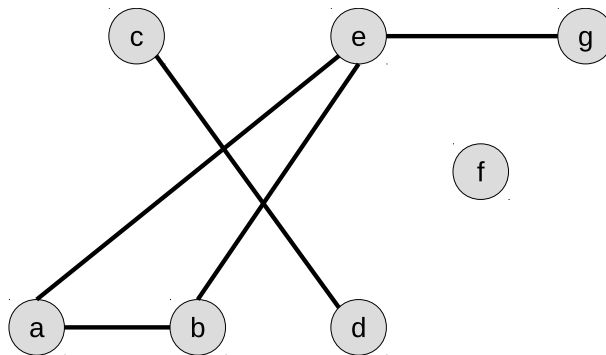


Figura 1 – Exemplo de um Grafo A.

Dígrafos são grafos orientados, nos quais a aresta que liga dois vértices $a, b \in V(G)$ é chamada de arco e possui uma direção, isto é, a aresta com origem no vértice a e destino no vértice b é diferente da que possui origem no vértice b e destino no vértice a . Por esta razão, para representar arcos são utilizados pares ordenados como (a, b) , para indicar que é uma aresta direcional com origem no vértice a e destino no vértice b . Um dígrafo

possui um conjunto finito não-vazio de vértices e um conjunto finito de pares ordenados de vértices distintos chamados arcos (BANG-JENSEN; GUTIN, 2007). Na Figura 2 temos um dígrafo B composto pelo conjunto de vértices $V(B) = \{a, b, c, d, e, f, g\}$ e pelo conjunto de arcos $E(B) = \{(a, b), (b, a), (a, e), (b, e), (c, d), (d, c), (e, g)\}$.

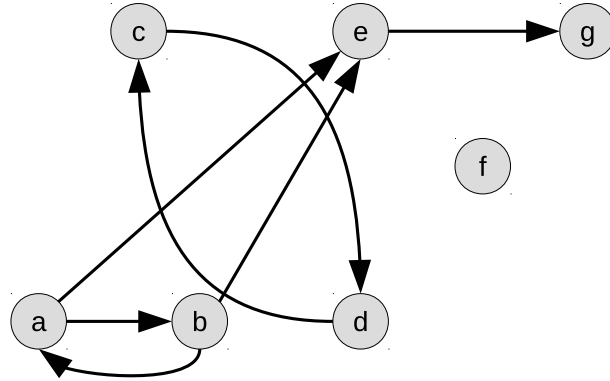


Figura 2 – Exemplo de um Dígrafo B.

2.2 MAG

Grafo MultiAspecto (MAG) é uma estrutura capaz de representar diferentes tipos de redes, tais como redes multicamadas, variáveis no tempo ou de alta ordem (WEHMUTH et al., 2016). Nas seções a seguir definiremos o MAG, apresentaremos algumas de suas características e propriedades.

2.2.1 Definição

Ele é definido como $H = (A(H), E(H))$, sendo que $E(H)$ é um conjunto de arestas e $A(H)$ é uma lista finita de conjuntos, onde cada elemento dessa lista é chamado de aspecto. Cada aspecto $\sigma \in A(H)$ é um conjunto finito, e o número de aspectos $p = |A(H)|$ é definido como da ordem de H . Cada aresta $e \in E(H)$ é uma tupla com $2 \times p$ elementos. Todas arestas são construídas seguindo o formato $(a_1, \dots, a_p, b_1, \dots, b_p)$, sendo a_1, b_1 elementos do primeiro aspecto de H , a_2, b_2 elementos do segundo aspecto de H , e assim por diante, até a_p, b_p elementos do último aspecto de H (WEHMUTH et al., 2016).

Dois conjuntos podem ser construídos a partir do produto cartesiano dos aspectos $A(H)$ na ordem p :

$$\mathbb{V}(H) = \prod_{n=1}^p A(H)[n], \quad (2.1)$$

que representa o conjunto de todos os vértices compostos possíveis do MAG H , gerado a partir do produto de todos os aspectos desse MAG H , e:

$$\mathbb{E}(H) = \prod_{n=1}^{2p} A(H)[(n-1)(\text{mod } p) + 1], \quad (2.2)$$

que representa o conjunto de todas as arestas possíveis no MAG H , sendo que $E(H) \subseteq \mathbb{E}(H)$.

Cada $u \in \mathbb{V}(H)$ é chamado de vértice composto do MAG H . Para todo $e \in \mathbb{E}(H)$ existem dois vértices compostos $u, v \in \mathbb{V}(H)$, sendo que u possui os elementos de origem, que são (a_1, a_2, \dots, a_p) , e v os elementos de destino, que são (b_1, b_2, \dots, b_p) , da aresta $e = (a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_p) \in E(H)$.

Na Figura 3 está representado um MAG composto por quatro aspectos representados pelos números, isto é, o conjunto $(A1, B1, C1, D1, \dots, K1)$ pertence ao primeiro aspecto, o conjunto $(A2, B2, C2, D2, \dots, K2)$ pertence ao segundo aspecto, e assim por diante. Esse MAG é composto pelo conjunto de onze vértices compostos

$$V = \{(A1, A2, A3, A4), (B1, B2, B3, B4), \dots, (K1, K2, K3, K4)\}$$

e pelo conjunto de três arestas descritas e ilustradas, como por exemplo, a aresta $(A1, A2, A3, A4, C1, C2, C3, C4) \in E(H)$.

Na Figura 4 temos um exemplo de uma pequena rede de transporte que pode ser representada por um MAG. Para essa rede, a lista de aspectos $A(H)$ contém três aspectos: as estações, os meios de transporte e os instantes de tempo. Cada aspecto possui um conjunto finito de informações, por exemplo, o conjunto de estações é $\{1, 2, 3\}$. O conjunto de arestas $E(H)$ nesse exemplo possui 22 arestas representadas pelas setas na figura. As setas pretas representam viagens de uma estação para outra em um meio de transporte, as setas azuis representam a troca de meio de transporte e as setas vermelhas a permanência em uma estação e meio de transporte. Nessa rede, a aresta $(2, \text{Onibus}, t1, 2, \text{Metro}, t1)$ é um exemplo de aresta que liga dois vértices em um mesmo instante de tempo; a aresta $(2, \text{Onibus}, t1, 2, \text{Onibus}, t2)$ é um exemplo de aresta que liga o mesmo vértice em dois instantes de tempo; e, a aresta $(2, \text{Onibus}, t1, 3, \text{Onibus}, t2)$ é um exemplo de aresta que liga dois vértices em dois instantes de tempo.

2.2.2 Propriedades

2.2.2.1 Isomorfismo entre MAGs e Grafos Direcionados

Embora a estrutura de uma aresta do MAG seja semelhante a uma aresta de um hipergrafo uniforme, a definição de adjacência usada em MAGs é próxima a utilizada em grafos orientados tradicionais. Pois, uma aresta do MAG pode ser vista como uma

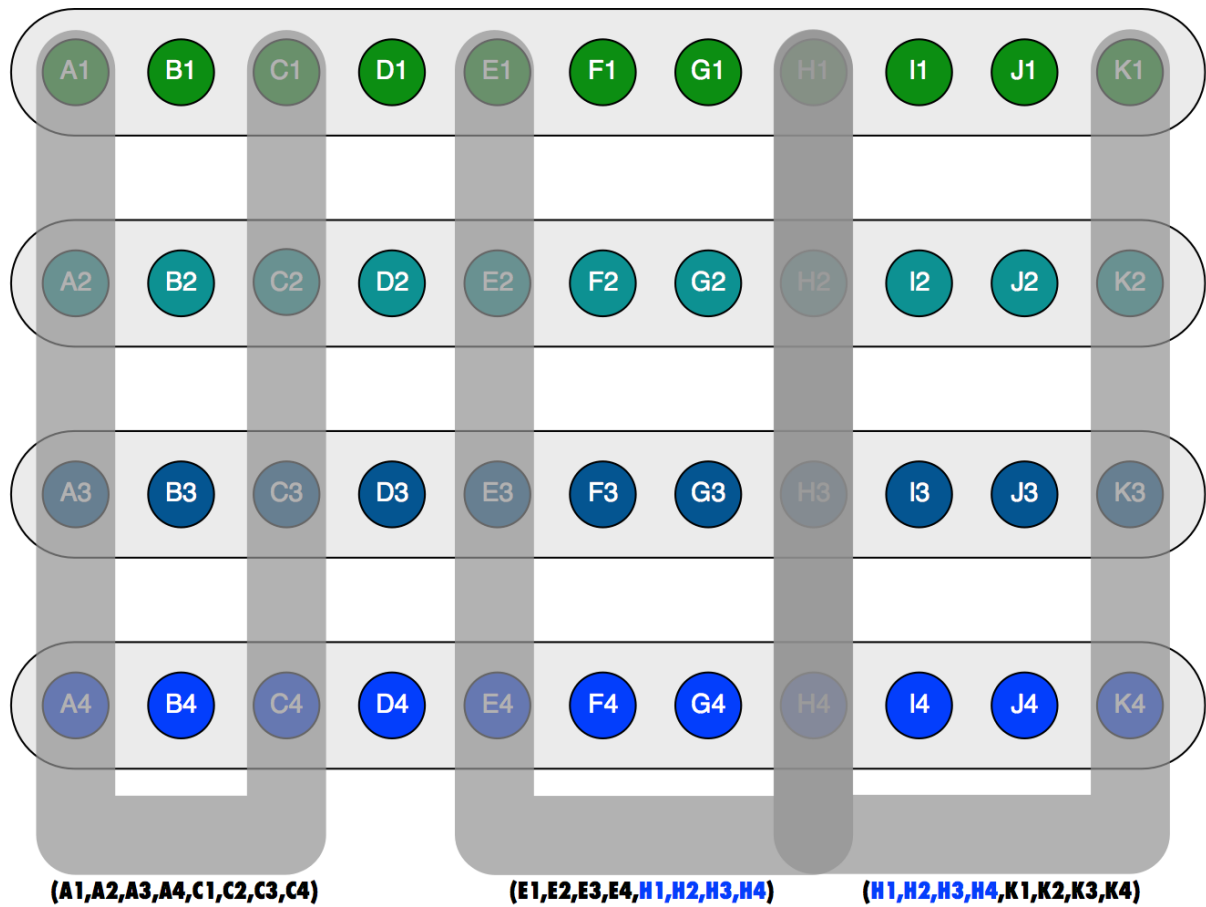


Figura 3 – Exemplo ilustrativo de três arestas MAG (WEHMUTH et al., 2017).

relação entre dois vértices compostos, que são objetos compostos construídos a partir dos elementos de aspecto. Portanto, uma aresta do MAG representa uma relação entre dois objetos (compostos) da mesma maneira que uma aresta de um grafo direcional. Esse conceito leva ao isomorfismo entre os MAGs e os grafos orientados, bem como aproxima seus conceitos sobre passeios, trilhas e caminhos WEHMUTH et al. (2016).

Dado que um MAG é isomorfo a um grafo direcional, a representação algébrica e os algoritmos para MAGs se assemelham aos dos grafos direcionados. Este isomorfismo entre MAGs e grafos direcionados abre o caminho para adaptar algoritmos de grafos bem conhecidos para aplicação em MAGs, facilitando assim o esforço para desenvolver a análise e aplicação de MAGs para modelagem de redes complexas WEHMUTH et al. (2016).

2.2.2.2 Predecessor e Sucessor

Dada a aresta $e = (a_1, a_2, \dots, a_p, b_1, b_2, \dots, b_p) \in E(H)$, na qual existem os vértices compostos $u = (a_1, a_2, \dots, a_p) \in V(H)$ e $v = (b_1, b_2, \dots, b_p) \in V(H)$, diz-se que u é predecessor de v e v é sucessor de u .

Serão adjacentes, ou vizinhas, todas as arestas que possuírem um vértice composto

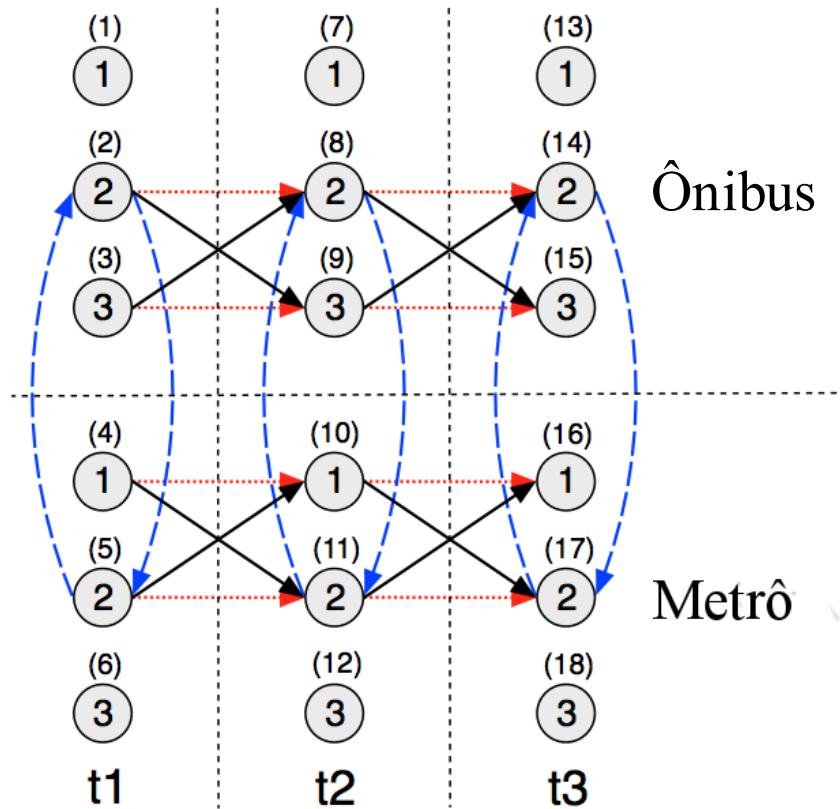


Figura 4 – MAG (WEHMUTH et al., 2017).

em comum. Isto é, se $u \in e_0$ e $u \in e_1$, e_0 é adjacente a e_1 .

2.2.3 Sub-MAG

Um dígrafo X é sub-dígrafo de Y se $V(X) \subseteq V(Y)$, $A(X) \subseteq A(Y)$, e se para todo arco pertencente a $A(X)$ existem os vértices no conjunto $V(X)$ (BANG-JENSEN; GUTIN, 2007). Uma vez que um MAG é isomorfo a um dígrafo tradicional, um sub-MAG é uma noção equivalente à noção de sub-dígrafos em dígrafos tradicionais (WEHMUTH et al., 2016).

Dado um MAG, é possível extrair uma parte dele, chamada de sub-MAG. Por definição, um sub-MAG também é um MAG. Podemos gerar um sub-MAG para realizar uma análise detalhada em um conjunto reduzido de vértices compostos de um determinado MAG (WEHMUTH et al., 2016). Por exemplo, do MAG expresso na Figura 4 poderíamos gerar um sub-MAG contendo apenas *Ônibus* como meio de transporte (Figura 2.2.3). Assim, nesse exemplo, todas as arestas que possuem o *Metrô* como parte de seus vértices e aqueles aspectos que estiverem relacionados apenas ao *Metrô* não estarão nesse sub-MAG.

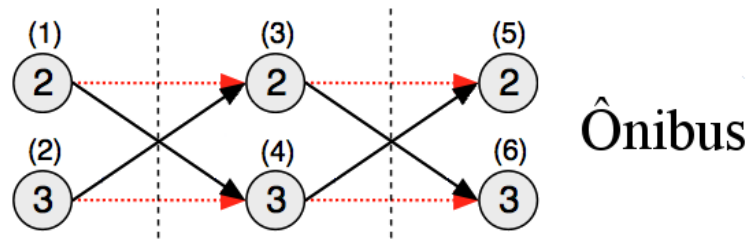


Figura 5 – Sub-MAG da Figura 4.

2.2.4 Subdeterminação

A subdeterminação é uma generalização do conceito de agregação aplicado a grafos multicamadas ou variáveis no tempo, em que todas as camadas podem ser agregadas, resultando em um grafo tradicional (WEHMUTH et al., 2017). Um MAG subdeterminado será um MAG com aspectos agregados de acordo com a subdeterminação determinada pela tupla zeta (ζ). A tupla ζ tem o mesmo número de elementos que o número de aspectos no MAG original, e para cada aspecto terá um valor 1 ou 0, de modo que somente os aspectos marcados com 1 permanecerão no MAG resultante.

Dessa forma, é possível subdeterminar um MAG H em $2^p - 2$ formas distintas, sendo $p = |A(H)|$, isto é, p é o número de aspectos do MAG H . Além disso, pode-se optar por considerar multigrafos direcionados (múltiplos arcos sem laços (*loops*)), ou pseudografos direcionados (múltiplos arcos com laços (*loops*)), ou dígrafos (sem múltiplos arcos e sem laços (*loops*)) ou grafos direcionados com laços (*loops*).

O MAG subdeterminado resultante terá as arestas do MAG original projetadas sobre a estrutura de aspecto reduzido dada pela subdeterminação definida. Sendo assim, o MAG subdeterminado terá ordem menor que a do MAG original. A operação de subdeterminação não terá inversa, pois acarreta em perda de informação.

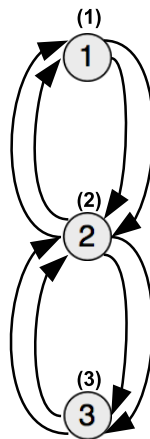


Figura 6 – MAG subdeterminado.

Como exemplo, podemos aplicar a subdeterminação no MAG representado na

Figura 4, no qual manteremos apenas o aspecto que representa as estações (1, 2 e 3) criando um multigrafo direcionado sem laços (*loops*). Durante essa subdeterminação no MAG da Figura 4, na qual consideramos apenas as estações, observe que as arestas em linhas tracejadas (azuis) e pontilhadas (vermelhas) serão laços (*loops*) no MAG resultante, e que as arestas em linhas contínuas (pretas) continuarão ligando diferentes estações. Assim, após retirar os laços, teremos a Figura 6 representando o MAG resultante dessa subdeterminação descrita.

2.2.5 Centralidades em MAG

Em teoria dos grafos e análise de redes, uma centralidade pode ser entendida como um indicador da importância relativa dos vértices ou arestas no grafo em análise. A avaliação das centralidades em um MAG pode ser feita a partir de qualquer representação do MAG, por exemplo, pode ser feita de maneira direta usando a representação de vértices compostos do MAG. Esta representação é isomorfa a um grafo orientado tradicional e transporta todas as arestas do MAG, preservando assim as suas propriedades topológicas, isto é, suas propriedades de adjacência. Assim, as centralidades podem ser calculadas utilizando os mesmos métodos aplicados aos grafos orientados tradicionais (WEHMUTH; ZIVIANI, 2018).

Como o MAG é isomorfo a um grafo direcionado no qual os vértices são os vértices compostos do MAG, qualquer métrica de centralidade conhecida para grafos direcionados pode ser aplicada diretamente para calcular as centralidades do vértice composto MAG (WEHMUTH; ZIVIANI, 2018).

Um caso particular de centralidades em MAG é a centralidade de aspecto único, onde o conjunto $A(H)[j]$ pode ser obtido por uma subdeterminação ζ_j na qual a j -ésima entrada na tupla da subdeterminação é 1 e todas as outras entradas são 0. Em particular, para qualquer MAG que tenha um aspecto designado como tempo, uma centralidade de tempo é uma centralidade de aspecto para este aspecto particular (WEHMUTH; ZIVIANI, 2018).

2.3 GVT

Uma vez que grafos estáticos não possibilitam visualizar o comportamento da rede relacionado ao tempo, os GVTs surgiram para proporcionar uma representação da dinâmica desses grafos. Na literatura eles também são conhecidos por Grafos Temporais ou Dependentes do Tempo (CASTEIGTS et al., 2012; HOLME; SARAMÄKI, 2012; KOSTAKOS, 2009). Em um GVT, podem ocorrer diferentes conexões entre os vértices com o passar do tempo, assim arestas podem existir em um determinado intervalo de tempo e não existir em outros (WEHMUTH et al., 2015).

Nas seções a seguir, serão expostos os modelos de representação para GVTs e a matriz de adjacência para representação algébrica de GVTs.

2.3.1 Modelos para representação de GVTs

Existem diferentes modelos para representar um GVT. Alguns modelos de GVTs são capazes de representar outros modelos de GVTs. Apresentaremos modelos baseados em *snapshots*, baseados em intervalos de tempo contínuo, baseados em arestas espaciais e temporais, baseados em temporais e mistas, o modelo de representação unificado e finalizamos com o MAG que também pode ser utilizado para representar um GVT e que será utilizado nessa dissertação.

2.3.1.1 Baseados em *snapshots*

Consideramos modelos baseados em *snapshots* aqueles que apresentam uma sequência de grafos (*snapshots*), onde cada grafo da sequência representa um determinado instante de tempo de um GVT. Isto é, ele expressa os vértices e arestas existentes para cada instante de tempo de um GVT. Na literatura, os trabalhos (BUI-XUAN et al., 2003; FERREIRA, 2004; HOLME, 2005; HOLME; SARAMÄKI, 2012; TANG et al., 2009, 2010a,b) se enquadram nesse modelo.

Na Figura 7, apresentamos um exemplo composto por quatro instantes de tempo t_1 , t_5 , t_7 e t_{15} , e em cada um deles são apresentadas as arestas e os vértices existentes naquele instante de tempo. Ainda nessa Figura 7, podemos visualizar um caminho partindo do vértice 0 na instante de tempo t_1 e alcançando o vértice 2 no instante t_5 .

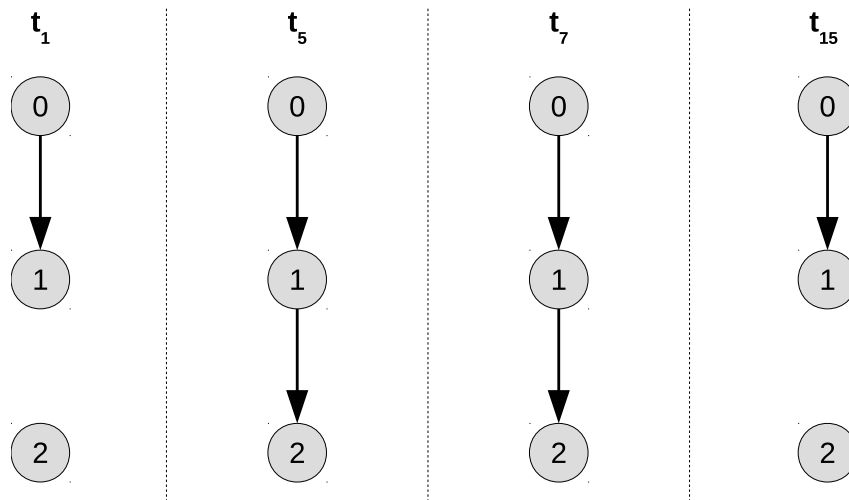


Figura 7 – Exemplo de *snapshot*.

Esse modelo é amplamente utilizado, pois ele trás uma forma fácil de visualização de um GVT, por serem, em geral, intuitivos e diretos. Porém, algumas informações ficam implícitas nesse modelo. Por exemplo, geralmente é assumido em tais modelos

de *snapshots* que os nós têm uma espécie de memória que permite que a transitividade induzida pelas arestas se propague em cada vértice ao longo do tempo. Isso significa que um caminho pode ser construído passando por um determinado vértice, mesmo se esse vértice for desconectado de todos os outros durante um período de tempo, o que significa que o vértice é capaz de reter a transitividade de aresta durante o período de desconexão. Mesmo que esse comportamento seja direto e intuitivamente esperado em muitos casos, ele não é construído dentro do modelo e deve ser assumido como uma propriedade adicional (externa) ao modelo, resultando que essa suposição também deve ser incorporada nos algoritmos usados com o modelo.

2.3.1.2 Baseados em intervalos de tempo contínuos

Uma outra maneira de representar GVTs é utilizando modelos baseados em intervalos de tempo contínuos, nos quais a existência da aresta é vinculada a um ou mais intervalos de tempo contínuo. O trabalho (CASTEIGTS et al., 2012) apresenta um modelo de intervalo contínuo.

Como exemplo, trazemos a Figura 8, na qual temos a aresta que liga os vértices 0 e 1 com duração entre os instantes 1 e 15, e a aresta que liga os vértices 1 e 2 com duração entre os instantes 5 e 7.

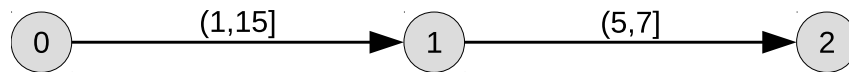


Figura 8 – Exemplo de intervalos de tempo contínuo.

2.3.1.3 Baseados em arestas espaciais e temporais

Alguns modelos como o proposto por (KOSTAKOS, 2009) são baseados na ideia de que um tipo de aresta representa conexões instantâneas entre vértices distintos enquanto outro tipo de aresta representa a “espera” de um dado vértice por uma nova conexão. Essa aresta de “espera” tem como origem um dado vértice a em um instante de tempo x e como destino o mesmo vértice a no instante de tempo y , onde o vértice a é origem em uma ou mais arestas com destino em vértices distintos a ele no instante x , não possui nenhuma conexão com outro vértice entre os instantes x e y , e é origem em uma ou mais arestas com destino em vértices distintos a ele no instante y .

Na Figura 9 temos as arestas em linhas contínuas (em preto), que indicam a conexões entre diferentes vértices no mesmo instante de tempo, as quais chamamos de arestas espaciais; e, temos as arestas em linhas tracejadas (em vermelho), que indicam o estado de “espera” do vértice nos diferentes instantes de tempo, as quais chamamos de arestas temporais. O vértice 0, por exemplo, possui arestas temporais para ele mesmo

nos próximos instantes, pois em todos instantes existe uma aresta espacial saindo dele (ligando ao vértice 1).

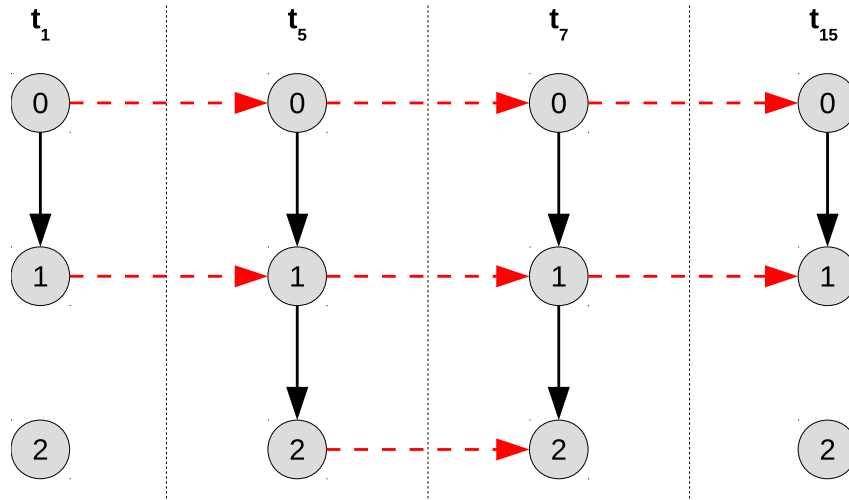


Figura 9 – Exemplo de arestas espaciais e temporais.

2.3.1.4 Baseados em arestas temporais e mistas

Alguns trabalhos, como (KIM; ANDERSON, 2012), trazem a ideia de conectar vértices distintos em instantes de tempos diferentes. Assim como mostra a Figura 10, na qual trás uma aresta que liga o vértice 0 no instante t_1 ao vértice 1 no instante t_{15} .

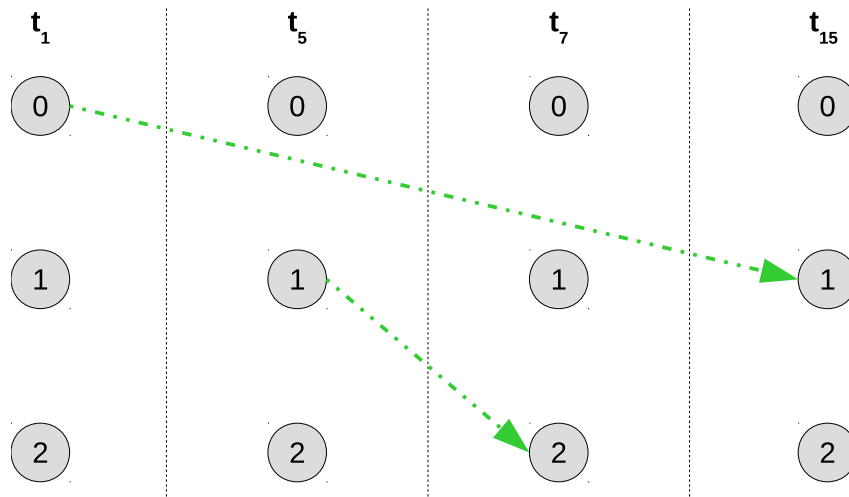


Figura 10 – Exemplo de arestas temporais e mistas.

2.3.1.5 Modelo de representação unificado

O modelo unificado foi proposto por (WEHMUTH et al., 2015), trás uma maneira de representar todos os modelos anteriormente apresentados. WEHMUTH et al. (2015) diz que um GVT é um caso particular de MAG, no qual a característica temporal

será um de seus aspectos. Ele poderá ser representado considerando o objeto $H = (V(H), E(H), T(H))$, onde $V(H)$ representa o conjunto de vértice, $T(H)$ o conjunto de instantes de tempo em que a rede está representada e $E(H) \subseteq V(H) \times T(H) \times V(H) \times T(H)$ o conjunto de arestas. Assim, $V(H)$ é o conjunto de todos os vértices de H , $E(H)$ é o conjunto de todas as arestas de H , $T(H)$ é o conjunto de todos os instantes de tempo em H . Uma aresta $e \in E(H)$ será uma quadrupla ordenada $e = (u, t_a, v, t_b)$, onde $u, v \in V(H)$ são vértices de origem e destino, respectivamente, enquanto $t_a, t_b \in T(H)$ são instantes de tempo de origem e destino, respectivamente. Entretanto, também podemos dizer que e é uma aresta direcional do vértice u no instante t_a para o vértice v no instante t_b . Sendo assim, podemos considerar (u, t_a) como um vértice composto de origem e (v, t_b) como um vértice composto de destino. O total de vértices compostos possíveis de um GVT pode ser dado pelo produto cartesiano $V(H) \times T(H)$.

Nesse modelo, cada aresta $e = (u, t_a, v, t_b)$ pode ser classificada em quatro classes de acordo com sua característica temporal:

1. *Arestas espaciais* conectam dois vértices distintos no mesmo instante de tempo, e está na forma de $e = (u, t_a, v, t_a)$, onde $u \neq v$;
2. *Arestas temporais* conectam o mesmo vértice em instantes de tempo distintos, e está na forma de $e = (u, t_a, u, t_b)$, onde $t_a \neq t_b$;
3. *Arestas mistas* conectam dois vértices distintos em instantes de tempo diferentes, e está na forma de $e = (u, t_a, v, t_b)$, onde $u \neq v$ e $t_a \neq t_b$;
4. *Arestas auto-laço (self-loop) espaço-temporais* conectam o mesmo vértice no mesmo instante de tempo, e está na forma de $e = (u, t_a, u, t_a)$.

A Figura 11 ilustra um GVT W , que possui o conjunto de vértices $V(W) = \{0, 1, 2, 3\}$, o conjunto de instantes de tempo $T(W) = \{t_0, t_1, t_2\}$. Temos a aresta $e = (0, t_0, 3, t_0) | e \in E(W)$ como exemplo. As linhas contínuas representam arestas espaciais que ligam os vértices em um mesmo instantes de tempo, por exemplo a aresta que liga o vértice 0 ao vértice 3 no instante de tempo t_0 . Já as linhas tracejadas representam arestas temporais que ligam vértices iguais em instantes de tempo distintos, por exemplo a aresta que liga o vértice 0 do instante de tempo t_0 no vértice 0 do instante de tempo t_1 .

Através do GVT W da Figura 11, podemos observar um exemplo do comportamento de um GVT, no qual vértices e arestas podem surgir e desaparecer com o passar do tempo. É notório que para cada instante de tempo há um comportamento distinto de arestas. Por exemplo, o vértice 0 alcança o vértice 3 no instante de tempo t_0 , já nos demais instantes de tempo isso não ocorre.

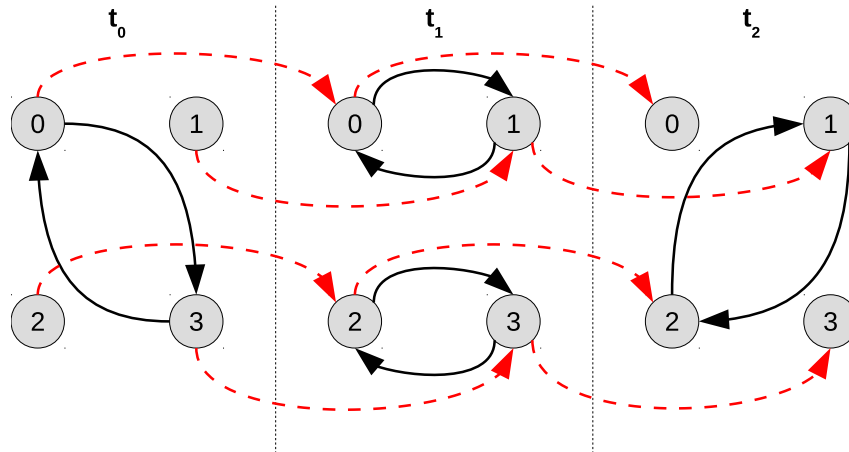


Figura 11 – Exemplo de um GVT W .

É importante observar que diferentes granularidades do tempo podem gerar grafos distintos. Por exemplo, no instante de tempo t_0 existe uma aresta ligando o vértice 0 ao 3 e no instante de tempo t_1 existe uma aresta ligando o vértice 1 ao 0. Assim, não é possível alcançar o vértice 3 partindo do vértice 1. Caso a granularidade do tempo fosse modificada, e os instantes de tempo t_0 e t_1 representassem um único instante de tempo, todas as arestas presentes nesses instantes seriam representadas no novo. Essa alteração levaria a um passeio possível do vértice 1 ao vértice 3, o que não representa a realidade. Em virtude desse fato, a ausência de granularidade (isto é, a não representação temporal) pode representar passeios que não existem.

2.3.1.6 Utilização do MAG para Representação de GVTs

O MAG, conforme descrito na Seção 2.2, é capaz de representar diferentes tipos de redes e de grafos, incluindo a representação de um GVT.

Para representar um GVT através um MAG H , inserimos o conjunto composto pelos instantes de tempo presentes no GVT como um conjunto na lista de aspectos $A(H)$. A Figura 4 (Seção 2.2.1) representa um GVT através da representação MAG.

Na Seção 5.2 descrevemos como utilizamos o MAG para representar um GVT e analisar a Centralidade de Tempo.

2.3.2 Matriz de Adjacência para Representação Algébrica de GVTs

Existem diferentes formas de representar uma rede matematicamente. Uma forma de representar grafos é utilizando uma Matriz de Adjacência (NEWMAN, 2010). A Matriz de Adjacência A de um grafo simples é uma matriz com elementos A_{ij} , tal que:

$$A_{ij} = \begin{cases} 1 & \text{se existe uma aresta entre os vértices } i \text{ e } j, \\ 0 & \text{caso contrario.} \end{cases}$$

Para um grafo não-orientado e para um grafo orientado simétrico a matriz de adjacência será simétrica, isto é, $A_{ij} = A_{ji}$. Por exemplo, para o grafo não-orientado da Figura 1, onde temos os vértices (a, b, \dots, g) , a matriz de adjacência seria:

$$A = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

Uma matriz de adjacência também poderá representar laços (*loops*), que serão inseridos na diagonal principal da matriz, e multigrafos, nos quais a existência de aresta é representado pela quantidade de arestas ligando dois vértices. Isto é, a Matriz de Adjacência M de um multigrafo é uma matriz com elementos M_{ij} , tal que:

$$M_{ij} = \begin{cases} n & \text{sendo } n \text{ a quantidade de arestas entre os vértices } i \text{ e } j, \\ 0 & \text{caso não exista aresta entre } i \text{ e } j. \end{cases}$$

Para GVTs, em que cada vértice pode possuir comportamento distinto para cada instante de tempo e as arestas são direcionadas, cada vértice composto pode ser representado por um número natural. Assim, podemos seguir o mesmo princípio da matriz de adjacência M_{ij} , i e j são números naturais que representam seus respectivos vértices compostos e também representam as linhas e colunas da matriz de adjacência. Assim, dada a aresta (u, t_a, v, t_b) , i representa a tupla de origem (u, t_a) e j a tupla de destino (v, t_b) . Como exemplo temos a Figura 12 exibe representação da matriz de adjacência do GVT W expresso na Figura 11.

Na Figura 12 cada um dos quatro nós (identificados como 0, 1, 2 e 3) do GVT W aparecem separadamente em cada um dos três instantes de tempo (t_0 , t_1 e t_2) que compõem o GVT W. Assim, o vértice composto $(0, t_0)$ está associado à linha 1 e coluna 1, o vértice composto $(1, t_0)$ está associado à linha 1 e coluna 2 e à linha 2 e coluna 1, e assim por diante. Além disso, podemos observar que ela possui cores distintas para

0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	1
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	2
1	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	3
0	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0
0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1
0	0	0	0	0	0	0	1	0	0	0	1	0	0	0	0	2
0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	1	3
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	1
0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	2
0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3
0	1	2	3	0	1	2	3	0	1	2	3					

Figura 12 – A matriz de adjacência do GVT W (WEHMUTH et al., 2015).

diferentes tipos de arestas. As arestas na diagonal principal em cada bloco são arestas espaciais que ligam dois vértices em um mesmo instante de tempo. As arestas em branco são arestas temporais, as quais ligam o mesmo vértice em instantes de tempo diferentes. Já as arestas mais escuras são arestas mistas, pois ligam vértices distintos em instantes de tempo distintos (WEHMUTH et al., 2015).

3 TRABALHOS RELACIONADOS

Muitos trabalhos na literatura buscam compreender o comportamento dos sistemas do mundo real. Em vários estudos são apresentadas diferentes formas de estruturar e analisar a dinâmica desses sistemas através do uso de redes complexas. O uso de Grafos Variantes no Tempo (GVTs) aparecem como ferramenta para modelar redes complexas com características temporais.

Dentro do tema GVT muitas propostas focam na sua representação. Por exemplo, alguns utilizam uma sequência de sub-grafos estáticos para modelar um GVT, que vão representando a evolução da rede à medida que o tempo passa (TANG et al., 2010a; HOLME; SARAMÄKI, 2012; KIVELÄ et al., 2014). Além disso, encontramos modelos mais generalistas que podem ser utilizadas para representar um GVT (WEHMUTH et al., 2017). Dentre outras representações, já citadas na Seção 2.3.1.

Existem pesquisas focadas na importância do nó observando a evolução da rede com o passar do tempo, isto é, análise da centralidade do nó. Dentro de Redes Temporais, (TANG et al., 2010a) propuseram métricas de centralidade temporal baseadas em caminhos temporais para medir a importância de um nó em uma rede dinâmica usando versões temporais de métricas de centralidade convencionais. Por exemplo, ele utilizou a métrica *clossness* que segue a ideia de encontrar o nó que possui a menor distância para os demais nós da rede. Ele também utilizou a métrica *betweenness* que segue a ideia de localizar o nó intermediário na rede, aquele que está no meio, capaz de interligar a rede.

Em outro trabalho, temos o (KIM; ANDERSON, 2012) trazendo *temporal degree*, *temporal closness* e *temporal betweenness*, sendo os três casos interessados em identificar o nó central de acordo com a evolução da rede. O *temporal degree* localiza o nó com mais conexões na evolução da rede. O *temporal closness* localiza o nó que durante a evolução da rede está mais próximo aos demais nós da rede. E, o *temporal betweenness* localiza o nó que durante a evolução da rede está intermediando melhor a ligação dos nós da rede.

A métrica de grau temporal foi estendida por (WANG et al., 2017), que propuseram o cálculo do desvio da centralidade do grau temporal. O objetivo desse trabalho é criar uma métrica que aproxime melhor ao que ocorre em uma rede real, onde os nós importantes variam com o tempo. Eles utilizam o modelo gráfico de janela de tempo para simplificar a rede para conseguir realizar os cálculos, selecionaram dados empíricos, e avaliaram a disseminação de uma doença durante uma epidemia. Diferente destes trabalhos, nosso objetivo é analisar o tempo dentro desse tipo de rede. Buscamos pelo instante de tempo mais importante tendo como foco a conectividade da rede, observando as ligações desse instante de tempo com os demais instantes de tempo. Elencamos os instantes de tempo com mais ligações e com maior probabilidade de ligação.

Nos trabalhos citados nos parágrafos anteriores, o foco principal é a importância

do nó na evolução temporal da rede. São trabalhos importantes para área, porém eles não analisam nenhum quesito referente a importância de um instante de tempo. Nós trazemos como foco deste trabalho, qualificar instantes de tempo dentro de um GVT, podendo assim dizer quais instantes de tempo são os mais importantes dentro da rede, de acordo com as métricas propostas.

A Centralidade de Tempo foi um tema abordado por (COSTA et al., 2015), com a finalidade de avaliar a importância relativa de um instante de tempo comparado a outros instantes de tempo de um mesmo Grafo Variante no Tempo. Este trabalho trás como proposta duas métricas para analisar a Centralidade do Tempo em uma rede tendo como foco processos de difusão. Assim, o objetivo dessas métricas é identificar o melhor instante de tempo para inicializar um processo de difusão que alcance um maior número de nós na rede com a menor duração possível.

A primeira métrica, denominada Tempo de Cobertura (*Cover Time*), analisa o tempo necessário para a difusão atingir um percentual ou um número dado de nós em um GVT, partindo de um determinado instante de tempo. Por exemplo, a Figura 13 mostra um processo de difusão iniciado em t_0 e terminado em t_{10} e outro processo iniciado em t_1 e terminado em t_9 . Intuitivamente, o instante de tempo t_1 é mais importante, uma vez que a difusão iniciada nele é mais rápida. Essa métrica é inversamente proporcional à importância dos instantes de tempo: quanto menor for o tempo médio de cobertura a partir de um determinado instante de tempo, maior a importância desse instante de tempo.

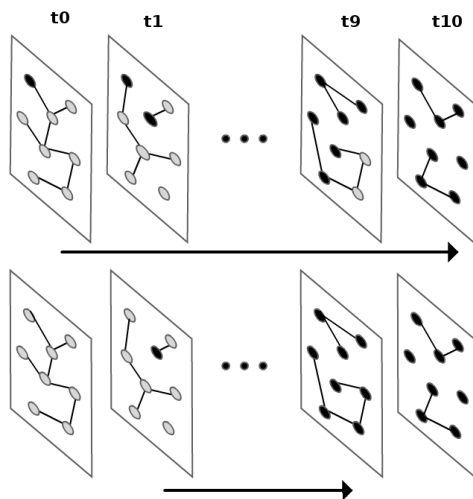


Figura 13 – Tempo de Cobertura (COSTA et al., 2015).

Já a segunda, Cobertura com Restrição de Tempo (*Time-Constrained Coverage*), estima a quantidade (ou percentual) de nós de um GVT que é atingida pela difusão da informação, dado um número de passos fixo φ do algoritmo de busca em largura. Por exemplo, a Figura 14 apresenta dois diferentes processos de difusão. Ambos os processos

ocorrem em φ passos. Percebe-se que o processo que começa em t_5 atinge um maior número de nós. Consequentemente, define-se que t_5 é mais importante que t_0 . Essa métrica é diretamente proporcional à importância dos instantes de tempo: quanto maior for a quantidade (percentual) média de nós atingidos pela difusão a partir de um determinado instante de tempo, maior a importância desse instante de tempo em comparação aos demais em um GVT. Pode-se dizer que a difusão iniciada no instante de tempo melhor avaliado tende a ser mais eficiente, pois atingiu, em média, um número maior de nós do que a difusão iniciada nos demais instantes de tempo.

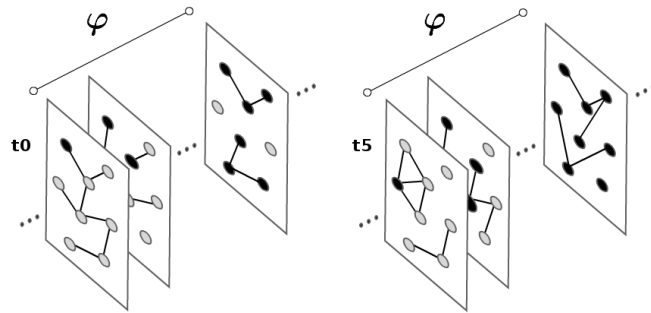


Figura 14 – Cobertura com Restrição de Tempo (COSTA et al., 2015).

Ao analisar as propostas feitas por (COSTA et al., 2015) percebemos a importância do trabalho no processo de difusão. Porém, entendemos que existe um número limitado de cenários que serão solucionados por essas propostas, onde somente aplicações com foco em difusão serão atendidas. Nós trazemos a proposta de analisar a importância dos instantes de tempo dentro de um GVT de acordo com a sua conectividade, atingindo assim, uma nova gama de aplicações que não se enquadram nos processos anteriores. Por exemplo, nós buscamos identificar o instante mais ativo em um GVT, observando todas características dele em conjunto.

4 CENTRALIDADE DE TEMPO

Tradicionalmente, o conceito de Centralidade em Teoria dos Grafos está relacionado ao vértice. Isto é, um dado vértice será mais importante em um grafo dada uma determinada métrica. O foco tradicional é quem se destaca. A ideia de centralidade foi introduzida por (BAVELAS, 1950) aplicada à comunicação humana. Ele analisou a comunicação em pequenos grupos organizados de diferentes formas e concluiu que a centralidade estava relacionada à eficiência do grupo na resolução de problemas. Partindo dessa centralidade, surgiram conceitos de *Closeness*, que calcula quão próximo um vértice é dos demais em um grafo (SABIDUSSI, 1966), e de *Betweenness*, que identifica o vértice que melhor intermedeia a comunicação entre todos outros pares de vértice em um grafo (FREEMAN, 1977). Esses conceitos são baseados em cálculos de caminho mínimo entre vértices.

A Centralidade de Tempo em GVTs é similar às medidas de centralidade de vértice. Porém, sua finalidade é avaliar a importância relativa de um dado instante de tempo comparado com outros instantes de tempo de um mesmo GVT (COSTA et al., 2015). Aqui o foco está no quando. Diferentes centralidades podem ser construídas para analisar os instantes de tempo. A partir das centralidades baseadas em caminho mínimo, técnicas foram criadas para difusão de informação. As métricas Tempo de Cobertura (*Cover Time*) e Cobertura com Restrição de Tempo (*Time-Constrained Coverage*) propostas por (COSTA et al., 2015), são exemplos. A primeira analisa o tempo necessário para a difusão atingir um percentual predeterminado de vértices. Já a segunda, estima a quantidade (ou percentual) de nós alcançados dado um número de passos fixos.

Em algumas redes, o contexto é distinto da difusão de informação. Por exemplo, em uma Rede de Transporte Público é importante saber quantos eventos ocorrem em determinado instante de tempo, podendo observar seus picos e escassez durante os dias. Nesse exemplo de rede, sabemos que existem horários de picos das partidas e chegadas correspondentes ao início e final do expediente de trabalho. Portanto, é importante observar a topologia e conectividade da rede para analisar e classificar seus instantes de tempo.

A Centralidade de Tempo baseada em Conectividade é aquela que classifica os instantes de tempo em um GVT analisando sua topologia e conectividade. Diferentes métricas podem ser criadas para classificar esses instantes de tempo. De acordo com o objetivo surgirão novas métricas para analisar a rede de uma forma específica. Em nossa pesquisa focamos em métricas que classificam e identificam instantes de tempo com mais conexões. A seguir, descrevemos nossas métricas de Centralidades de Grau do Tempo e Centralidade de PageRank do Tempo.

4.1 CENTRALIDADE DE GRAU DO TEMPO

Em grafos tradicionais, a Centralidade de Grau do vértice foi definida inicialmente por (FREEMAN, 1978). Ele relaciona a centralidade de um vértice ao seu grau. O Grau do vértice corresponde ao número de ligações que ele possui com outros nós em uma rede. NEWMAN (2010) diz que o grau de um vértice em um grafo é o número de arestas conectadas a ele. Será mais central o vértice com maior grau.

Em grafos orientados, cada vértice v terá grau de entrada, grau de saída e grau total. O grau de entrada de um vértice v equivale ao número de arestas que chegam em v , isto é, o número de arestas onde o vértice v é a cabeça. O grau de saída de um vértice v equivale ao número de arestas que saem em v , isto é, o número de arestas onde o vértice v é a cauda. O grau total do vértice v é dado pela soma dos seus graus de entrada e de saída, que corresponde ao total de arestas que possuem o vértice v em uma de suas extremidades (BANG-JENSEN; GUTIN, 2007).

Para calcular o Grau do Tempo em um GVT, propomos gerar uma estrutura na qual os instantes de tempo sejam os vértices, mas que as demais características do GVT sejam mantidas na estrutura. Assim, conforme descrito na Seção 2.2.1, podemos representar um GVT utilizando a estrutura MAG; em seguida, subdeterminar o MAG considerando apenas o aspecto tempo –isto é, em seus vértices teremos apenas os instantes de tempo–; e por fim, gerar sua Matriz de Adjacência. Assim, seguimos os passos do Algoritmo 1 que receberá como parâmetros de entrada o GVT $G(V(G), E(G), T(G))$ – sendo $V(G)$ o conjunto de vértices, $T(G)$ o conjunto de instantes de tempo em que a rede está representada e $E(G)$ o conjunto de arestas – e a tupla ζ_t – tupla contendo o valor 1 na posição que indica o aspecto tempo, e 0 nas demais posições. Como saída, o Algoritmo retorna a Matriz de Adjacência M que representa o MAG subdeterminado em função do tempo e que será utilizada para o cálculo do Grau do Tempo.

Algoritmo 1: Geração do MAG S a partir de um GVT G .

```

Entrada:  $G, \zeta_t$  // GVT  $G(V(G), E(G), T(G))$  e tupla  $\zeta_t$  tendo tempo como aspecto único.
Saída:  $S, M$  // MAG subdeterminado  $S(A(S), E(S))$  e sua Matriz de Adjacência  $M$ .

GeraMAGS( $G, \zeta_t$ )
     $H \leftarrow \text{contraMAG}(G)$  // Constrói MAG  $H(A(H), E(H))$  a partir do GVT  $G$ .
     $S \leftarrow \text{subdetermineMAG}(H, \zeta_t)$  // Gera MAG subdeterminado  $S(A(S), E(S))$ .
     $M \leftarrow \text{geraMatrizAdjacencia}(S)$  // Constrói Matriz de Adjacência  $M$  do MAG  $S$ .
retorna  $S, M$ 

```

Assim, observando o Algoritmo 1, seja um MAG H que representa um GVT G hipotético e que possui o tempo como um de seus aspectos em $A(H)$. Para obter a sua Centralidade de Grau do Tempo, subdeterminamos o MAG H mantendo apenas o aspecto tempo. Teremos então, o MAG subdeterminado S , e sua lista de aspectos $A(S)$ com apenas um aspecto que corresponde ao conjunto de instantes de tempo da rede. Os instantes de tempo $i, j \in V(S)$, onde $V(S)$ é o conjunto dos vértices compostos do MAG S , sendo

que cada vértice composto possui apenas o instante de tempo. A Figura 15 representa o MAG S onde cada vértice é um instante de tempo.

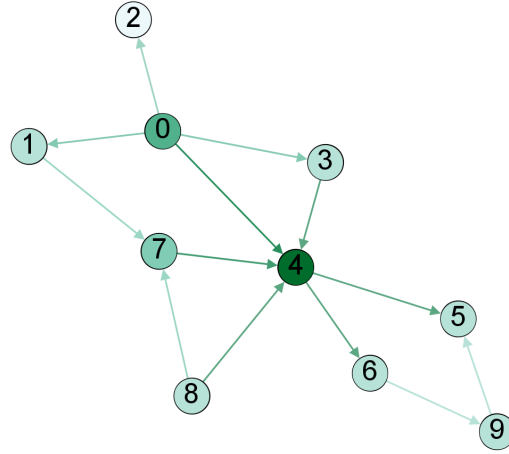


Figura 15 – Exemplo MAG S .

O grau de entrada, denotado por d_j^{in} , e grau de saída, denotado por d_i^{out} , se darão da seguinte forma:

$$d_j^{in} = \sum_{i=1}^N M_{ij}, \quad (4.1)$$

$$d_i^{out} = \sum_{j=1}^N M_{ij}, \quad (4.2)$$

onde M_{ij} é a posição i, j da matriz de adjacência do MAG S . O grau total do tempo i será a soma dos graus de entrada e saída, isto é, $d_i = d_i^{in} + d_i^{out}$. Assim, o Instante de Tempo 4 do MAG S representado pela Figura 15 terá $d_4^{in} = 4$, $d_4^{out} = 2$ e $d_4 = 6$.

O Algoritmo 2 traz os passos para calcular os Graus de Entrada e Saída do Tempo de um GVT, dada a Matriz de Adjacência que representa as conexões desse GVT considerando os instantes de tempo como vértices.

Algoritmo 2: Centralidade de Grau do Tempo para um GVT.

```

Entrada:  $M$  // Matriz de Adjacência  $M$ .
Saída:  $entrada, saida$  // Vetores com os Graus do Tempo de Entrada e de Saída.
CentralidadeGrauTempo( $M$ )
     $N \leftarrow |M[i]|$  // Armazena número de vértices.
    // Calcula Grau do Tempo de Entrada e de Saída para cada instante de tempo do GVT.
    para  $i \leftarrow 1; i \leq N; i \leftarrow i + 1$  faça
        para  $j \leftarrow 1; j \leq N; j \leftarrow j + 1$  faça
             $entrada[j] \leftarrow entrada[j] + M[i][j]$ 
             $saida[i] \leftarrow saida[i] + M[i][j]$ 
        fim
    fim
retorna  $entrada, saida$ 

```

4.2 CENTRALIDADE DE PAGERANK DO TEMPO

O PageRank, descrito em (BRIN; PAGE, 1998), foi o primeiro algoritmo utilizado na ferramenta de pesquisa do Google. Sua proposta inicial era utilizar uma métrica capaz de identificar e ranquear páginas importantes na web. Para isso, utiliza um processo de cálculo iterativo que envolve a simulação da navegação de forma aleatória dentro de uma rede de páginas web ligadas por hiperlinks. O objetivo é identificar páginas altamente recomendadas onde as pessoas tendem a acessar. Em outras palavras, o PageRank poderia quantificar a acessibilidade de locais potencialmente alcançados por pessoas.

A Figura 16 representa um exemplo numérico do PageRank em um sistema pequeno, sendo cada círculo um vértice e cada seta uma aresta que liga um vértice a outro. O vértice F tem um valor de PageRank mais elevado do que o vértice G , pois a aresta que chega em F vem de um vértice importante, enquanto as que chegam em G são de vértices com menor importância.

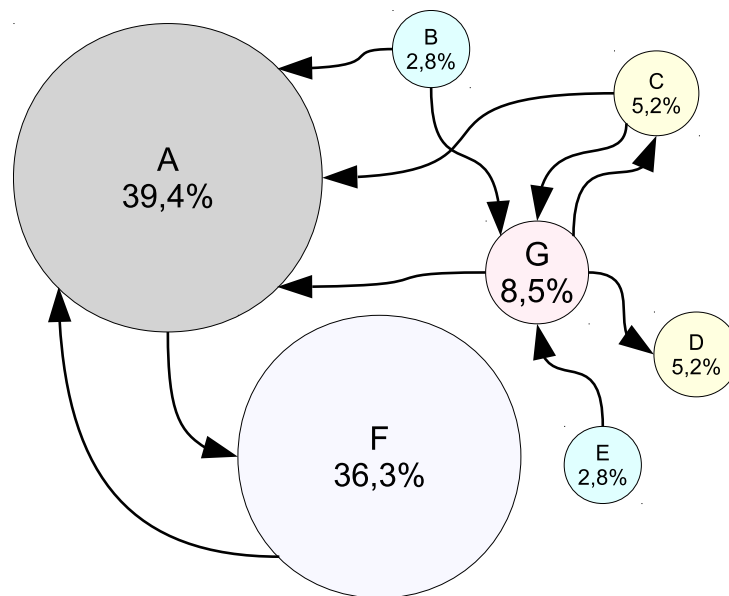


Figura 16 – Exemplo numérico do Pagerank em um sistema pequeno.

A centralidade de PageRank é derivada da Centralidade de Autovetor, que por sua vez é uma extensão da Centralidade de Grau (NEWMAN, 2010). Enquanto na Centralidade de Grau somente a quantidade de arestas ligadas a um vértice importa, no PageRank o valor da importância de cada vértice também importa. Assim, o PageRank gera uma distribuição que indica a probabilidade de um passeio aleatório chegar em um determinado vértice.

Para calcular o PageRank de um determinado instante de tempo, utilizamos a quantidade de arestas que chegam ao vértice e o valor proporcional de PageRank de todos os vértices que chegam a ele em um cálculo iterativo. Isto é, todos os vértices da rede começam com um mesmo valor inicial de PageRank, assim o cálculo iterativo será realizado

até que os valores de PageRank convirjam considerando um dado erro. Será mais central o vértice que tiver maior valor de PageRank.

Assim como na Centralidade de Grau do Tempo, executamos os passos do Algoritmo 1 antes de calcular a Centralidade de PageRank do Tempo.

Assim, observando o Algoritmo 1, seja um MAG H que representa um GVT hipotético e possui o tempo como um de seus aspectos em $A(H)$. Para obter a Centralidade de PageRank do Tempo, subdeterminamos o MAG H mantendo apenas o aspecto tempo, dado um ζ_t . Teremos então, o MAG subdeterminado S , sua lista de aspectos $A(S)$ com apenas um elemento que corresponde ao conjunto de instantes de tempo da rede, e seu conjunto de vértices compostos $V(S)$. Os instantes de tempo $i, j \in V(S)$ estão presentes nos vértices compostos.

O cálculo do PageRank de um instante de tempo $i \in V(S)$, denominado por pr_i , segue o proposto por NEWMAN (2010) e se dará da seguinte forma:

$$pr_i = \alpha \sum_j M_{ji} \frac{pr_j}{d_j^{out}} + \beta, \quad (4.3)$$

onde β é um termo constante que garante um valor mínimo para o PageRank; α é um parâmetro utilizado para equilibrar o primeiro termo com o termo constante da equação; M_{ji} é a posição j, i da matriz de adjacência do MAG S que retorna a quantidade de arestas que ligam esses dois vértices; pr_j é valor de PageRank do instante de tempo $j \in V(S)$; e d_j^{out} é o grau de saída do instante de tempo $j \in V(S)$.

Além disso, α e β são constantes positivas, que somadas valem 1. O β é a parte “livre”, o termo extra constante que todos os vértices recebem. Ao adicionar este segundo termo, mesmo os vértices com grau zero ainda obtêm a centralidade β , e uma vez que eles têm uma centralidade diferente de zero, então os vértices para os quais eles apontam derivam alguma vantagem de serem apontados. A Centralidade de PageRank difere da Centralidade do Autovetor comum por ter um parâmetro livre α , que governa o equilíbrio entre o termo do autovetor e o termo constante na Equação 4.3. Para utilizarmos a Centralidade de PageRank do Tempo, devemos primeiro escolher um valor para essas constantes. Ao fazer isso, é importante entender que α não pode ser arbitrariamente grande. Se $\alpha \rightarrow 0$, então resta apenas o termo constante na Equação 4.3 e todos os vértices têm a mesma centralidade β . À medida que aumentamos α , as centralidades aumentam e, eventualmente, chega um ponto em que elas divergem (NEWMAN, 2010).

O Algoritmo 3 trás os passos para calcular o PageRank do Tempo de um GVT, dada a Matriz de Adjacência que representa as conexões desse GVT considerando os instantes de tempo como vértices.

Algoritmo 3: Centralidade de PageRank do Tempo para um GVT.

Entrada: $S, M, saida, \alpha, \beta, erro$ // MAG subterminado $S(A(S), E(S))$, sua Matriz de Adjacência M , o Grau do Tempo de Saída $saida$, α e β .

Saída: pr // Vetor com PageRank do Tempo.

CentralidadePageRankTempo($S, M, saida, \alpha = 0.85, \beta = 0.15, erro = 0.0001$)

```

   $nv \leftarrow |saida|$  // Armazena número de vértices.
  para  $i \leftarrow 1; i \leq nv; i \leftarrow i + 1$  faça
  |  $pr[i] \leftarrow 1/nv$  // Calcula e armazena PageRank inicial.
  fim
  // Calcula e armazena o PageRank até que haja conversão dos valores, isto é, que o
  // valor de erro calculado seja menor que o estabelecido.
   $e = 1$ 
  enquanto  $e \geq erro$  faça
  |  $prOld \leftarrow pr$ 
  | para  $i \leftarrow 1; i \leq nv; i \leftarrow i + 1$  faça
  | |  $soma \leftarrow 0$ 
  | | para  $j \leftarrow 1; j \leq nv; j \leftarrow j + 1$  faça
  | | |  $soma \leftarrow M[j][i](pr[j]/sai[j])$ 
  | | fim
  | |  $pr[i] \leftarrow \alpha soma + \beta$ 
  | fim
  |  $e \leftarrow calculaNovoErro(pr, prOld)$ 
  fim
retorna  $pr$ 

```

5 APLICAÇÃO E ANÁLISE DE CENTRALIDADE DE TEMPO

Com a finalidade de avaliar e validar a Centralidade de Tempo baseada em Conectividade, aplicamos os conceitos propostos em um estudo de caso que usa um conjunto de dados real. Escolhemos um conjunto de dados que contém a característica temporal entre os seus atributos. Nesse sentido, escolhemos trabalhar com uma rede de transportes.

Redes de transportes são exemplo de Redes Complexas. Através da evolução das pesquisas em Redes Complexas tem sido possível representar e analisar mais detalhes de cada uma delas (STROGATZ, 2001). Compreender o funcionamento de uma rede de transportes nos permite conhecer e prever acontecimentos. Uma forma de representar essas redes é através da utilização de grafos.

Poucos conjuntos de dados de redes de transporte disponibilizados representam a realidade com muitas características. Por isso, escolhemos o conjunto de dados proposto por (GALLOTTI; BARTHELEMY, 2015) que representa todo o transporte público da Grã-Bretanha através de uma rede temporal multicamadas. Esse conjunto será descrito na Seção 5.1.

5.1 CONJUNTO DE DADOS

O conjunto de dados proposto por GALLOTTI; BARTHELEMY (2015) nos permitiu remodelar os dados de acordo com a estrutura MAG e avaliar as centralidades de tempo propostas. Seus autores extraíram os dados sobre ônibus urbano (“*Bus*”), ônibus interurbano (“*Coach*”), trem (“*Rail*”), balsa (“*Ferry*”) e metrô (“*Metro*”) do *National Public Transport Data Repository (NPTDR)* (UK, 2017) correspondentes a uma semana completa de Outubro de 2010, para evitar coincidir com feriados e períodos de férias. As informações sobre voos domésticos foram fornecidas pela *Innovata LLC* (USA, 2017) da semana de 18 a 24 de outubro de 2010. A Tabela 1 apresenta o conjunto de arquivos gerados no modelo proposto em (GALLOTTI; BARTHELEMY, 2015) com suas dimensões.

Tabela 1 – Dimensões do Conjunto de Dados.

Arquivo	Linhas	Tamanho
<i>layers.csv</i>	7	76 B
<i>nodes.csv</i>	267.032	12 MB
<i>edges.csv</i>	475.503	14 MB
<i>events.txt</i>	465.216	893 MB

O arquivo *layers.csv* traz as camadas do sistema de transporte sendo um identificador *layer* e o meio de transporte *layerLabel*. A coluna *layer* possui os valores de 0 a 5 e correspondem aos meios de transporte *Air*, *Ferry*, *Rail*, *Metro*, *Coach* e *Bus*, respectivamente, expressos na coluna *layerLabel*.

O arquivo *nodes.csv* traz informações sobre os nós da rede, chamados de pontos, como um simples ponto de ônibus, ou áreas de parada, como estações de trem e aeroporto. Suas colunas são *nodeid* que contém um identificador numérico para cada nó, *layerid* que indica a camada em que o nó aparece, *lat* que indica latitude, *lon* que indica longitude, e os campos *areacode* e *atccode* que são utilizados como identificador em outros arquivos que trazem mais informações sobre o nó.

O arquivo *edges.csv* traz informações sobre as arestas construídas pelos autores desse conjunto de dados. Suas colunas são *ori node* que indica o nó de origem, *des node* que indica nó de destino, *ori layer* que indica camada de origem, *des layer* indicando camada de destino, *minutes* que indica tempo mínimo de viagem entre a origem e o destino em minutos, e *km* que indica a distância euclidiana entre a origem e o destino em quilômetros. Para áreas de parada, nas quais o nó é o mesmo e altera a camada, eles calcularam o tempo de deslocamento de um ponto ao outro tendo como base a distância e a velocidade de 5 km/h . Na camada *Air*, tendo em vista os procedimentos realizados em aeroportos, eles acrescentaram 2 h ao tempo de deslocamento do embarque e 30 min ao tempo de deslocamento do desembarque.

O arquivo *events.csv* traz para cada aresta do arquivo *edges.csv* todos horários de partidas seguidos da duração da viagem. Na descrição do conjunto de dados é informado que os eventos dos ônibus urbanos são maioria, representando mais de 90% do total de eventos. *ori node* que indica o nó de origem, *des node* que indica nó de destino, *ori layer* que indica camada de origem, *des layer* indicando camada de destino, seguidos de um conjunto finito representado por $\{t_1, dt_1, t_2, dt_2, \dots, t_n, dt_n\}$, onde cada t_x , com $x \in \{1, 2, \dots, n\}$, representa cada horário de partida em minutos de *ori node* para *des node* dentro da semana, e cada dt_x , com $x \in \{1, 2, \dots, n\}$, representa a duração dessa viagem. Os instantes de partida tem início nas 0 h de segunda-feira totalizando 10.080 min em uma semana.

Além disso, a Figura 17 traz a distribuição das durações das viagens que utilizaremos. As camadas *Bus* e *Metro* não foram analisadas, pois o objetivo deste estudo de caso é realizar análises do comportamento do tempo no transporte entre cidades. Ainda, a Tabela 2 complementa as informações referentes a essas durações. Com esses dados podemos observar que grande maioria das viagens possuem menos de 100 min de duração. Dentro da camada *Air* temos uma rede de viagens mais longas, pois a maioria de suas viagens, cerca de 75% , possui duração acima de 65 min . Nas demais camadas, a maioria das viagens possui duração de viagem mais curtas. A camada *Rail* se destaca com as viagens mais curtas, cerca de 80% das viagens duram menos que 10 min . Essa última característica se dá por diversos fatores, tais como, velocidade do transporte, número de estações e proximidade entre elas.

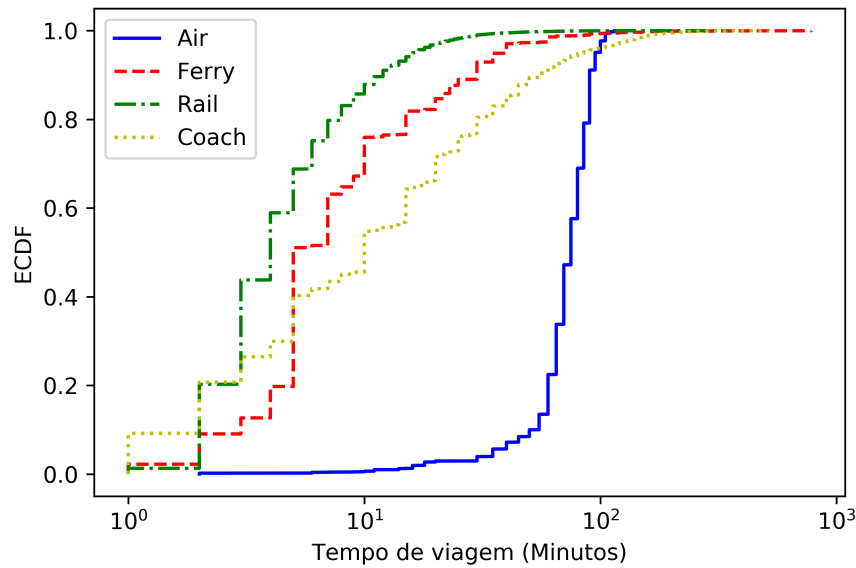


Figura 17 – ECDF das durações das viagens.

Tabela 2 – Informações das Durações das Viagens.

Camada	Qtde.	Média (min)	Desvio p. (min)	Mín. (min)	25% (min)	50% (min)	75% (min)	Máx. (min)
Air	4.357	72.4	18.3	2	65	75	85	115
Ferry	38.004	12.0	21.5	1	5	5	10	780
Rail	1.470.071	5.8	6.3	1	3	4	6	345
Coach	135.891	21.9	33.6	1	3	10	25	510

5.2 MODELO PROPOSTO

Conforme citado na Seção 2.3.1.6, utilizamos a estrutura MAG para modelar o GVT selecionado (Seção 5.1) para aplicação e análise das Centralidades de Tempo propostas.

Nessa seção trazemos a descrição do modelo construído para utilizar a estrutura MAG. Em seguida apresentamos a validação desse modelo. E, por fim, explicamos como realizamos a subdeterminação nessa estrutura.

5.2.1 Descrição do Modelo

Modelamos o conjunto de dados de acordo com a estrutura MAG, considerando três aspectos pertencentes a $A(H)$: o ponto ou área de parada, a camada (representando o meio de transporte) e o instante de tempo (representando horários de partida ou chegada), sempre nessa ordem. O produto cartesiano desses aspectos geram todos vértices compostos possíveis, e os nossos vértices compostos $V(H)$ estão contidos nesse conjunto, isto é,

$$V(H) \subseteq \prod_{n=1}^3 A(H)[n]. \quad (5.1)$$

Ainda nesse modelo, cada viagem, embarque, desembarque e mudança de meio de transporte é representado por uma aresta pertencente ao conjunto $E(H)$. Elas representam rotas entre dois locais, e serão descritas a seguir.

O aspecto ponto ou área de parada é um conjunto composto por 267.031 locais. No conjunto de dados utilizado (Sessão 5.1), esse aspecto eram os nós da rede de transporte e diversas informações estão associadas a ele. Porém, para esse modelo consideramos apenas o valor *nodeid*. São denominados pontos de parada os pontos simples, por exemplo, um terminal de balsa. E áreas de parada são aquelas que permitem dois ou mais meios de transporte chegarem nela, como estações de trem ou aeroportos.

O aspecto camada é o conjunto composto pelos meios de transporte, também descritos como camadas no conjunto de dados. Tendo em vista o objetivo de realizar análises do comportamento do tempo no transporte entre cidades, as camadas *Bus* e *Metro* não foram utilizadas nesse modelo. Para cada meio de transporte serão utilizadas duas camadas. Uma camada será usada para representar as viagens e outra camada será usada para representar esperas e conexões. Ou seja, a segunda camada de cada meio de transporte é uma camada de “trânsito”, por isso identificaremos com o sufixo “t”. Por exemplo, será utilizado o valor “2” para indicar o início ou final de uma viagem de trem, e será utilizado “2t” para indicar a permanência na plataforma de uma estação de trem. Assim, o aspecto camada será composto por oito itens. O conjunto desse aspecto está representado na Tabela 3.

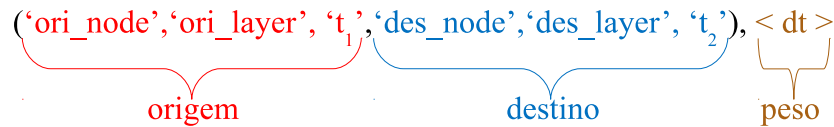
Tabela 3 – Valores do Aspecto Camada.

Meio de Transporte	Camada de transporte	Camada de conexões
Air	0	0t
Ferry	1	1t
Rail	2	2t
Coach	3	3t

O aspecto de instantes de tempo representa cada horário de partida ou chegada. Ele será um sub-conjunto de $\{0, 1, 2, \dots, 10.079\}$ que representa todos os minutos em uma semana a partir das 0h de segunda-feira. Representará o instante de chegada sempre que o vértice composto for o destino, e o instante de partida sempre que for origem.

Um vértice composto será expresso da seguinte forma $(node, layer, t)$, sendo *node* o local ou área de parada, *layer* a camada (meio de transporte ou de conexões) e *t* o instante de tempo. Seja um exemplo de vértice composto dado por $(1066, 0, 10060)$, onde o item 1066 pertence ao primeiro conjunto da lista de aspectos, e representa o Aeroporto de Barra em Eoligarry, Reino Unido no conjunto de dados. O item 0 pertence ao segundo conjunto da lista de aspectos e representa a camada *Air*. E o item 10060 pertence ao terceiro conjunto da lista de aspectos e representa domingo às 23 h 40 min.

Modelamos arestas ponderadas contendo o vértice composto de origem, o vértice composto de destino e o peso, que será a duração da aresta em minutos. Sua estrutura está ilustrada na Figura 18a, sendo os valores ori_node , ori_layer e t_1 correspondentes ao vértice composto de origem, os valores des_node , des_layer , t_2 correspondes ao vértice composto de destino, e o valor dt correspondente à duração da aresta em minutos entre a origem e o destino, isto é, $dt = t_2 - t_1$, que representa o peso da aresta. A Figura 18b trás um exemplo de uma aresta que representa um voo com origem no vértice (46311, 0, 00530), com destino no vértice (1066, 0, 00600) e com peso de 70 *min* correspondentes a duração do voo. Esse exemplo representa o voo partindo do Aeroporto de Glasgow em Paisley, Reino Unido às 8 h 50 *min* da segunda-feira, chegando no Aeroporto de Barra em Eoligarry, Reino Unido às 10 h da segunda-feira com duração de 70 *min*.



(a) Estrutura.

(**'46311'**, **'0'**, **'00530'**, **'1066'**, **'0'**, **'00600'**), < 70 >

(b) Exemplo.

Figura 18 – Estrutura de uma aresta ponderada.

Para construir o conjunto de arestas $E(H)$, observamos o funcionamento do sistema de transporte público. Não foram considerados translados até determinado ponto ou área de parada, mas foi considerado tudo que pode ocorrer dentro do sistema a partir de todos os pontos ou áreas de parada do conjunto de dados. Por exemplo, estando em uma estação de trem é possível viajar para outra estação que também tenha pontos de ônibus, e nessa é possível ir para o ponto de ônibus e viajar para outro local. Assim, foram construídos cinco tipos de arestas conforme a representação da Figura 19 que serão descritas a seguir.

Com efeito, observamos as etapas para uma viagem em apenas um meio de transporte, na qual uma pessoa está em um ponto ou área de parada, embarca no transporte, viaja, e desembarca no ponto destino. Para cada uma dessas etapas gastam-se tempos distintos. Nesse exemplo, observam-se três tipos de arestas representadas na Figura 19a, sendo uma para o embarque, outra para a viagem (deslocamento) e uma terceira para o desembarque. Sempre que uma pessoa desce em uma área de parada e decide continuar a viagem em outro meio de transporte, ela precisará caminhar do local de desembarque para o local de embarque do outro meio de transporte. Esse cenário está representado pelas arestas intermodais na Figura 19b. Por fim, foram modeladas as arestas temporais, representadas na Figura 19c, para o cenário em que uma pessoa aguarda por

semana para outra, usando esse processo de módulo.

As Arestas de Embarque (*Departure Edges*) representam o embarque no meio de transporte, isto é, representam a transição entre camadas de conexões e a camada de transporte. Por exemplo, o passageiro está na plataforma do trem que pretende viajar e entra no trem. As durações de embarque, que chamamos de dt_e , utilizados nessa aresta foram definidos considerando que a pessoa já está no local para realizar o embarque, seja plataforma ou portão. Isto é, na camada *Air* consideramos que o processo de *check-in* já foi realizado restando apenas o procedimento de embarque na aeronave, ao qual consideramos um tempo de 30 *min.* E, para as demais camadas foram considerados 5 *min.* Essas durações são valores arbitrários, uma vez que não aparecem na base original. Para construir essas arestas, utilizamos as Arestas de Deslocamento construídas anteriormente. Assim, ordenamos as Arestas de Deslocamento por meio de transporte (*ori_layer*), por ponto ou área de origem (*ori_node*) e por horário de partida. Para cada vértice composto único de origem das Arestas de Deslocamento foi criada uma Aresta de Embarque seguindo a seguinte estrutura

$$(\text{ori_node}, \text{ori_layer}+t, t_1 - dt_e, \text{ori_node}, \text{ori_layer}, t_1), < dt_e >,$$

sendo que a parcela (*ori_node*, *ori_layer*+*t*, $t_1 - dt_e$) representa o vértice composto de origem da aresta, a parcela (*ori_node*, *ori_layer*, t_1) representa o vértice composto de destino da aresta e a parcela $< dt_e >$ representa o peso da aresta. Para contruí-la, começamos pelo vértice composto de destino, que será o mesmo vértice composto de origem da Aresta de Deslocamento, isto é, a Aresta de Embarque tem como destino a origem da Aresta de Deslocamento (Figura 19a). Em seguida, ajustamos o vértice composto de origem que terá o mesmo valor de *ori_node*, a camada será a camada de origem da Aresta de Deslocamento acrescida da letra *t* e o instante de tempo será calculado considerando o instante de tempo de origem da Aresta de Deslocamento menos a duração do embarque ($t_1 - dt_e$). Por fim, o peso será a duração do embarque dt_e . Caso o instante de tempo de origem calculado fique negativo, consideraremos que o embarque foi na semana anterior, e ajustaremos somando 10.080 ao valor.

É importante observar que, para cada tupla (*local*, *meio de transporte*, *horário de partida*) distinta das Arestas de Deslocamento, foi construída apenas uma Aresta de Embarque. Ou seja, existem alguns locais, como aeroportos, nos quais há mais de um voo saindo no mesmo horário. Nesse caso, só será gerada uma Aresta de Embarque para todos os voos que tiverem suas partidas no mesmo horário. Por exemplo, a Aresta de Embarque da Figura 20a foi gerada a partir das Arestas de Deslocamento da Figura 20b.

As Arestas de Desembarque (*Arrival Edges*) representam o desembarque do meio de transporte, isto é, a saída do meio de transporte para a plataforma, por exemplo. Nessa aresta, as durações de desembarque, que chamamos de dt_d , foram arbitradas considerando

$$(40906,0t,00355,40906,0,00385), < 30 > \qquad (40906,0,00385,31874,0,00465), < 80 >$$

$$\qquad\qquad\qquad (40906,0,00385,41487,0,00480), < 95 >$$

(a) Aresta de Embarque. (b) Arestas de Deslocamento.

Figura 20 – Exemplos de Arestas.

que a pessoa sairá do meio de transporte para plataforma ou portão. Para camada *Air* foram considerados 30 *min* e para as demais 5 *min*. Como primeira etapa para essa construção, ordenamos os vértices compostos de destino das Arestas de Deslocamento. Similar ao processo de criação das Arestas de Embarque, para todo vértice composto único de destino das Arestas de Deslocamento, uma Aresta de Desembarque é criada. Para construí-las, fizemos um processo similar ao processo de criação das Arestas de Embarque seguindo a seguinte estrutura

$$(des_node, des_layer, t_2, des_node, des_layer+t, t_2 + dt_d), < dt_d >,$$

sendo que a parcela $(des_node, des_layer, t_2)$ representa o vértice composto de origem da aresta, a parcela $(des_node, des_layer+t, t_2 + dt_d)$ representa o vértice composto de destino da aresta e a parcela $< dt_d >$ representa o peso da aresta. O vértice composto de origem da Aresta de Desembarque será o mesmo vértice composto de destino da Aresta de Deslocamento, isto é, a Aresta de Deslocamento tem como destino a origem da Aresta de Desembarque (Figura 19a). Já o vértice composto de destino da Aresta de Desembarque terá o mesmo valor de *des_node*, a mesma camada de destino da Aresta de Deslocamento acrescida da letra *t* e o instante de tempo será calculado considerando o instante de tempo do destino da Aresta de Deslocamento somado a duração do desembarque $(t_2 + dt_d)$. Por fim, o peso será a duração do desembarque dt_d . Caso o instante de tempo do vértice composto de destino calculado ultrapasse o valor 10.079, consideraremos que o desembarque foi na próxima semana, e ajustaremos com a operação de módulo $(t_2 + dt_d)\%10080$.

A Aresta Intermodal (*Intermodal Edges*), representada na Figura 19b, foi definida para os locais que são área de parada, nos quais existem a possibilidade de trocar o meio de transporte, por exemplo um aeroporto que possui pontos de ônibus ou estação de metrô integrados aos portões de embarque. Nesse caso, teremos o deslocamento de uma camada para outra, chamado de dt_i . O tempo para este deslocamento vem do conjunto de dados, e foi arbitrado pelos autores do conjunto de dados original considerando a distância percorrida entre as camadas e uma velocidade média de 5 *Km/h*. Sempre que a camada destino era *Air* foram acrescidos ao tempo de deslocamento 90 *min*, pois, somando com 30 *min* que estão na Aresta de Embarque, é o tempo padrão para realização de *check-in* nos aeroportos. Para construir essas arestas utilizamos as Arestas de Desembarque e de Embarque, e representamos assim

$$(des_node_D, des_layer_D, t_{2D}, ori_node_E, des_layer_E, t_{1E} \geq (t_{2D} + dt_i)), < t_{1E} - t_{2D} >.$$

Para cada Aresta de Desembarque em locais com mais de uma camada, foi criada uma Aresta Intermodal para cada camada diferente do desembarque. Esse cenário está exemplificado na Figura 21, que representa um aeroporto com pontos de ônibus e plataformas de trem. Nele, sempre que um passageiro chega de ônibus ao aeroporto, ele poderá ir para um portão de embarque aéreo ou para uma plataforma de trem. Assim, o vértice composto de origem da Aresta Intermodal é o vértice composto de destino da Aresta de Desembarque (Figura 19b). Já o vértice composto de destino da Aresta Intermodal é o vértice composto de origem da próxima Aresta de Embarque em cada camada existente no local, considerado o tempo de deslocamento dt_i .

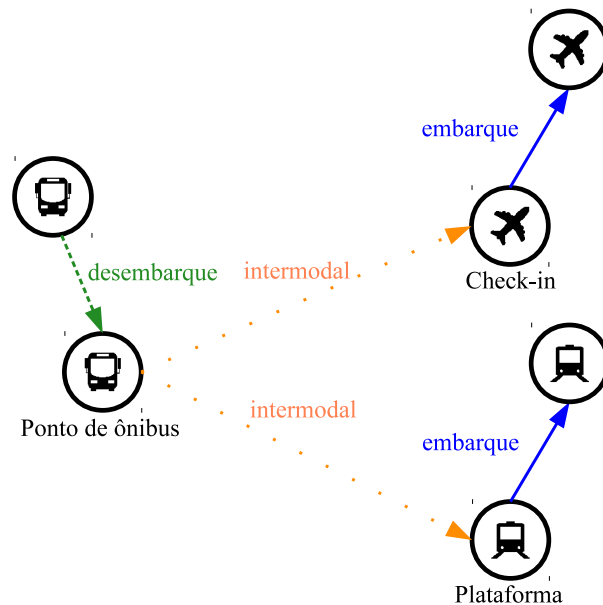


Figura 21 – Arestas Intermodais.

A Aresta Temporal (*Temporal Edges*), representada na Figura 19c, possibilita realizar conexões de viagens em um mesmo meio de transporte, interligando todas as chegadas e partidas de um mesmo local (ponto ou área de parada) e camada. Assim, ao desembarcar em um local será possível aguardar por uma nova partida que seja de interesse do passageiro. Podemos dizer que essas arestas costuram todas as chegadas e partidas em determinado local e camada. Para construí-las, utilizamos as Arestas de Embarque e Desembarque. As Arestas Intermodais não foram utilizadas, pois os vértices compostos que deram origem a essas arestas são os existentes nas Arestas de Embarque e Desembarque. Primeiro, para cada local de cada camada, separamos todos os vértices compostos de origem das Arestas de Embarque e todos os vértices compostos de destino das Arestas de Desembarque. Ordenamos esses vértices compostos pelo aspecto dos instantes de tempo e construímos arestas interligando os vértices compostos. Isto é, ligamos o primeiro vértice

composto no segundo, o segundo no terceiro, e assim sucessivamente até ligarmos o último no primeiro. O último liga-se ao primeiro para gerar o ciclo semanal do transporte público.

O tamanho de cada conjunto da lista de aspectos e de cada conjunto de arestas gerados estão expressos na Tabela 4. A primeira linha dessa tabela expressa o cenário completo do sistema de transporte público intercidades trabalhado. As demais linhas expressam cada meio de transporte sendo visto isoladamente. Esclarecemos que os valores expressos na primeira linha não são somatório dos demais, tendo em vista que existem interseções entre os conjuntos da lista de aspectos (locais e instantes de tempo) nas diferentes camadas. O conjunto do aspecto camada para cada meio de transporte sempre terá apenas dois elementos, sendo o identificador do meio de transporte e o mesmo identificador acompanhado da letra t . Por exemplo, ao analisar apenas o meio de transporte *Air* o conjunto do aspecto camada será $\{0, 0t\}$.

Tabela 4 – Arestas Construídas.

Meios de transporte	Arestas	Locais	Camadas	Tempos
Air, Ferry, Rail e Coach	7.877.731	4.302	8	10.079
Air	19.062	48	2	1.803
Ferry	177.342	216	2	8.170
Rail	6.765.993	2.490	2	10.044
Coach	615.014	1.864	2	10.038

5.2.2 Validação do Modelo

A fim de garantir uma estrutura sólida, várias validações foram realizadas durante a criação das arestas em formato texto e após a geração do MAG. É importante garantir que as arestas estejam interligas, que os instantes de tempo pertençam ao intervalo de uma semana e que os instantes de tempo não sejam negativos.

Durante a criação das arestas eram verificados o tempo de origem, o tempo de destino e o peso. Tempos de origem e destino deveriam estar dentro do conjunto $\{0, \dots, 10.079\}$. Sempre que o tempo inicial gerado ficasse negativo, consideramos o início da aresta na semana anterior calculando $t_1 + 10.080$. Quando o tempo final ficasse superior a 10.079, consideramos o final da aresta na semana seguinte calculando o resto da divisão, $t_2 \% 10.080$. O peso deveria ser sempre maior que 1, pois o tempo mínimo suportado por este modelo é 1 min. Assim, sempre que a duração da aresta fosse menor que 1, essa não era gerada.

Após a geração do MAG, foram verificadas a sequência embarque, deslocamento e desembarque. Assim, analisamos se para toda Aresta de Embarque havia no mínimo uma Aresta de Deslocamento, sendo o vértice composto de destino da primeira o mesmo vértice composto de origem da segunda. Para toda Aresta de Deslocamento se havia uma e somente uma Aresta de Desembarque, sendo o vértice composto de destino da primeira

o mesmo vértice composto de origem da segunda. Para cada Aresta de Desembarque se havia ao menos uma Aresta de Deslocamento, sendo o vértice composto de origem da primeira o mesmo vértice composto de destino da segunda.

As Arestas Temporais devem fechar um ciclo em cada local e camada. Cada Aresta Temporal deverá apontar para uma e somente uma Aresta Temporal, isto é, o vértice composto de destino de uma Aresta Temporal deverá ser o vértice composto de origem da próxima Aresta Temporal na linha do tempo. Foi verificado se existia uma e somente uma próxima Aresta Temporal, se a sequência dessas arestas era monótona crescente, e se a última ligava-se na primeira aresta.

Durante a validação das Arestas Temporais, foram detectados quatro locais com apenas uma Aresta de Embarque ou de Desembarque, conforme a Tabela 5, que além da Aresta de Embarque ou de Desembarque possui a Aresta de Deslocamento que gerou o embarque/desembarque. Essas viagens de fato existiram na semana em que as informações foram extraídos para construção do conjunto de dados original. Tendo em vista que a quantidade de arestas era ínfima, mantivemos elas.

Tabela 5 – Locais com somente uma Aresta de Embarque ou de Desembarque.

Meio de Transporte	<i>nodeid</i>	Descrição	Arestas
<i>Air</i>	163007	<i>Coventry Baginton Airport</i>	(163007, 0, 00570, 163007, 0t, 00600), < 30 > (38731, 0, 00555, 163007, 0 , 00570), < 15 >
<i>Air</i>	35785	<i>Gloucestershire Airport</i>	(35785, 0, 00510, 35785, 0t, 00540), < 30 > (38731, 0, 00495, 35785, 0 , 00510), < 15 >
<i>Ferry</i>	254019	<i>North Ronaldsay Orkney Ferry Terminal</i>	(640, 1, 06120, 254019, 1 , 06280), < 160 > (254019, 1, 06280, 254019, 1t, 06285), < 5 >
<i>Rail</i>	29424	<i>Okehampton Rail Station</i>	(29424, 2 , 08360, 22054, 2, 08367), < 7 > (29424, 2t, 08355, 29424, 2, 08360), < 5 >

Dentre o conjunto de arestas foram identificados um grupo inferior a 0,1 % do total de arestas que possuem duração superior a 20 h. Parte dessas arestas são Temporais e outra parte Intermodais. Esses casos ocorrem em lugares com poucos embarques e desembarques.

Além disso, também identificamos um grupo de 0,3 % do total de arestas em que o local de origem e destino possui o mesmo valor de identificação, com mesmo meio de transporte, com intervalos pequenos de tempo (maiores que 1 *min*). Alguns casos ocorrem em rios, onde os terminais de balsas das duas margens recebem o mesmo identificador. Outros são em aeroportos, grandes estações de trem, ou qualquer local que possua mais de uma parada para um meio de transporte, podendo ser mais de um ponto de ônibus, por exemplo.

5.2.3 Subdeterminação

Conforme descrito na Seção 2.2.4, o MAG subdeterminado tem as arestas do MAG original projetadas sobre a estrutura de aspecto reduzido dada pela subdeterminação conforme as configurações escolhidas. Como nosso objetivo é validar a Centralidade de Tempo baseada em Conectividade, subdeterminamos o MAG gerado considerando o aspecto único instante de tempo ($\zeta = (0, 0, 1)$) e optamos por multigrafos sem laços (*loops*) para preservar todas arestas com tempos de origem e destino distintas. Como nenhuma aresta do modelo tem peso zero, não seria gerado nenhum laço (*loop*) e todas as arestas permaneceram. Exceto as 4 arestas temporais que possuem duração de 10080 *min*, que foram geradas pelos locais descritos na Tabela 5. A Tabela 6 traz o quantitativo de arestas resultantes da subdeterminação.

Tabela 6 – Arestas após a subdeterminação.

Meios de Transporte	Arestas	Instantes de Tempo
Air, Ferry, Rail e Coach	7.877.727	10.079
Air	19.060	1.803
Ferry	177.341	8.170
Rail	6.765.992	10.044
Coach	615.014	10.038

5.3 AVALIAÇÃO DAS CENTRALIDADES DE TEMPO

Nessa seção trazemos a metodologia para calcular as Centralidades de Grau do Tempo e de PageRank do Tempo, bem como a discussão dos resultados encontrados.

5.3.1 Metodologia

Para execução e construção desse modelo utilizando a estrutura MAG, foi utilizado o ferramental disponível em (WEHMUTH et al., 2017). Este foi construído utilizando a linguagem Python. Tendo em vista a quantidade de arestas no conjunto de dados escolhido, foram utilizados computadores com no mínimo 16 GB de memória RAM para a execução das ferramentas.

Em um primeiro momento, extraiu-se as informações necessárias do conjunto de dados para a criação das arestas. Em seguida, executamos a ferramenta para geração da estrutura MAG com todas as arestas.

Com a estrutura MAG construída, foi possível manipulá-la de diferentes maneiras de acordo com o objetivo de avaliação. Primeiro avaliamos a estrutura completa, e em seguida, geramos sub-MAGs para cada meio de transporte, e para algumas combinações de meios de transporte.

Com a estrutura MAG a ser analisada em mãos, subdeterminamos utilizando apenas o aspecto instantâneo de tempo, considerando multigrafo direcionado (múltiplos arcos sem laço *loop*). Pois, como nosso conjunto de dados é uma rede de transporte, era interessante manter todas as viagens entre dois instantes de tempo.

Após a subdeterminação, realizamos os cálculos das centralidades desejadas. A seguir apresentamos os resultados encontrados com suas respectivas análises.

5.3.2 Centralidade de Grau do Tempo

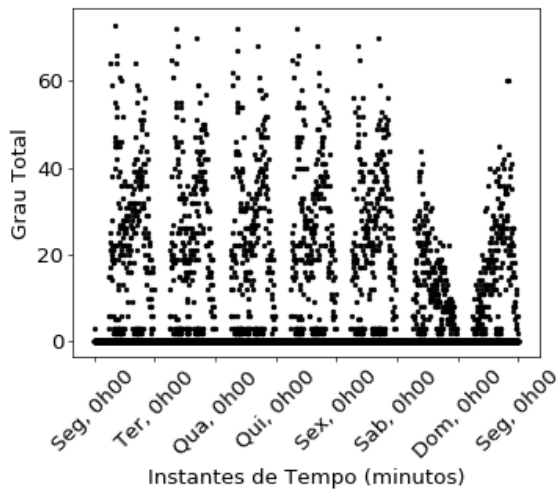
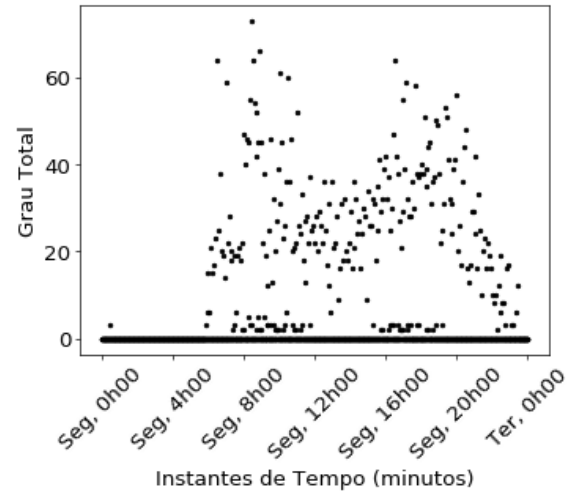
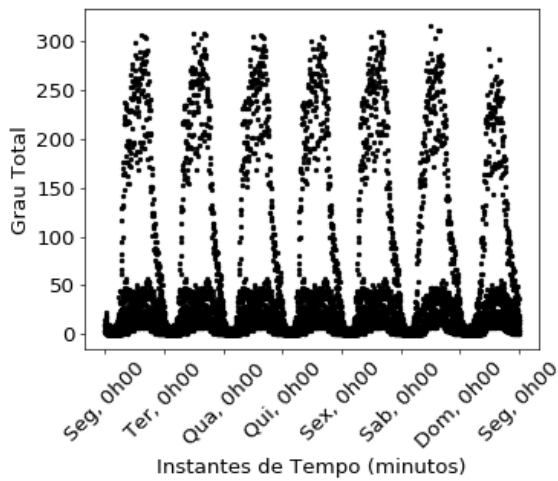
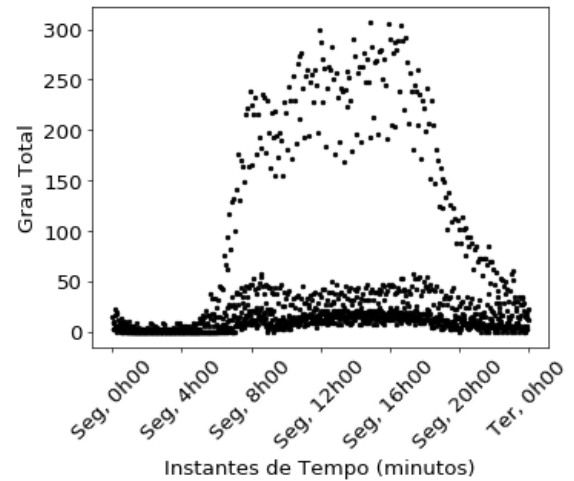
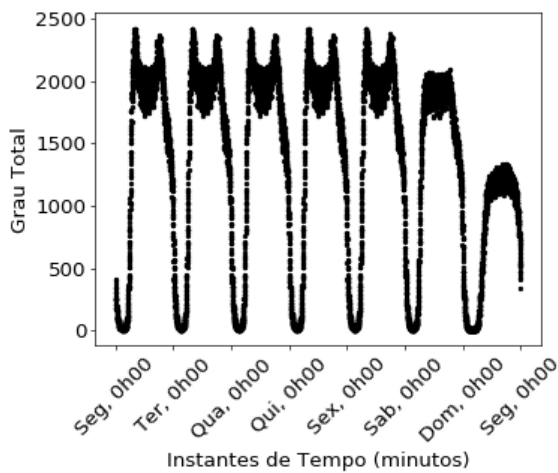
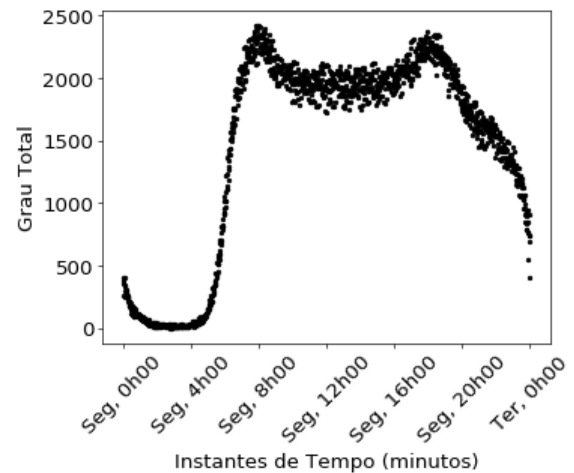
Nós calculamos grau de entrada, grau de saída e grau total para diferentes cenários. Todavia, nós colocamos no corpo da dissertação apenas os gráficos referentes ao grau total, pois os três graus possuem um comportamento similar. Ainda assim, todos os detalhes que destacam diferenças entre eles serão ressaltados no texto. Além disso, todos os gráficos para grau de entrada e grau de saída estão no Apêndice-A.

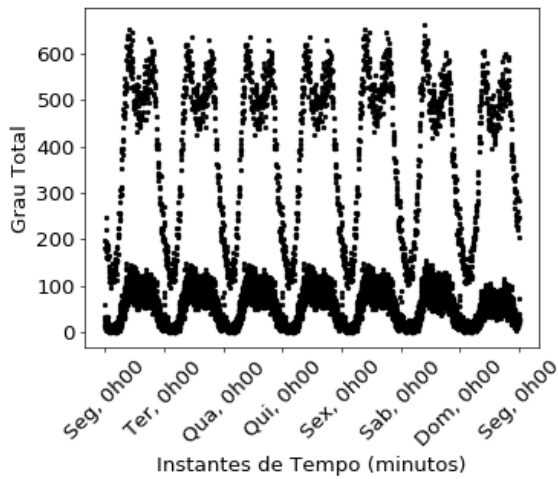
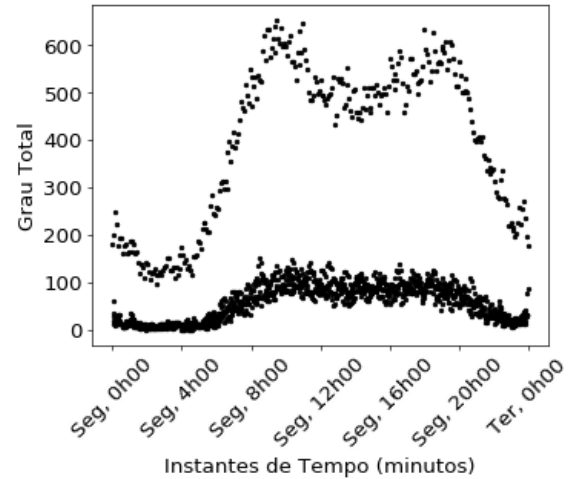
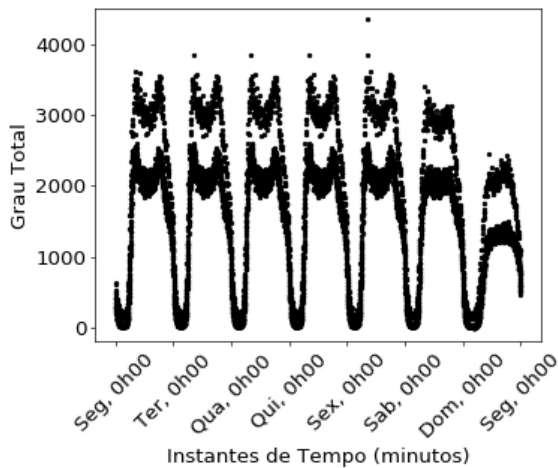
Em um primeiro momento, nós analisamos cada meio de transporte separadamente, isto é, calculamos os graus do tempo para os meios de transporte *Air* (Figuras 22a e 22b), *Ferry* (Figuras 22c e 22d), *Rail* (Figuras 22e e 22f), e *Coach* (Figuras 23a e 23b). Em seguida, nós observamos a rede de transporte completa, com todos esses meios de transporte juntos (Figuras 23c e 23d). Por fim, nós observamos os graus da rede combinando esses meios de transporte três a três (Figura 24).

Todos os gráficos dos Graus do Tempo apresentam a relação do grau calculado para cada instante de tempo, isto é, o grau de entrada é o total de chegadas em cada minuto (Apêndice-A – Figuras 32a, 32c, 32e, 33a, 33c e 34); o grau de saída é o total de partidas em cada minuto (Apêndice-A – Figuras 35a, 35c, 35e, 36a, 36c e 37); e o grau total é a soma das partidas e chegadas em cada minuto (Figuras 22a, 22c, 22e, 23a, 23c e 24). Nos três casos foram calculados durante uma semana, a partir das 0 h da segunda-feira. Além disso, os minutos em que não há partidas e chegadas também foram representados.

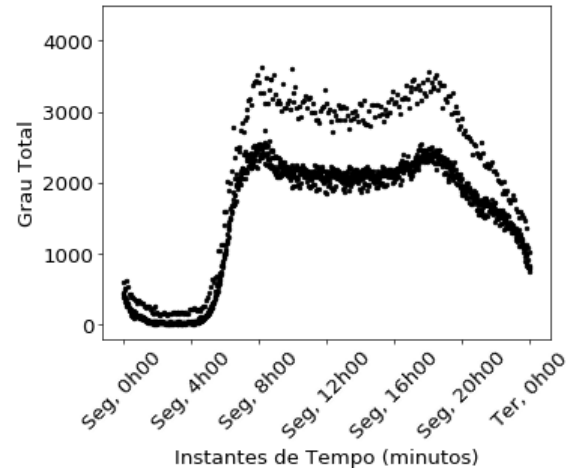
É importante ressaltar que nossa análise é referente aos acontecimentos em cada instante de tempo. Na análise em questão, observamos o número de partidas e chegadas em cada minuto em toda Grã-Bretanha. Isso significa que, se o grau de entrada igual a 10, não são necessariamente 10 chegadas em um mesmo local. Ou, se existe alguma periodicidade nos horários de partida, por exemplo a cada 5 *min*, ela pode ser um padrão em todo ou parte do território.

A primeira característica observada em todos esses cálculos e cenários é comportamento diário da rede. Pois, em cada gráfico semanal das Figuras 22, 23 e 24 é possível visualizar sete agrupamentos dos graus totais por instantes de tempo. Ao observar os instantes de tempo que iniciam um novo agrupamento, verificamos que correspondem de fato a um novo dia dentro da semana. Assim, visualizamos o comportamento da rede nos

(a) *Air* – Durante a semana.(b) *Air* – Segunda-feira.(c) *Ferry* – Durante a semana.(d) *Ferry* – Segunda-feira.(e) *Rail* – Durante a semana.(f) *Rail* – Segunda-feira.Figura 22 – Grau do Tempo – *Air*, *Ferry* e *Rail*.

(a) *Coach* – Durante a semana.(b) *Coach* – Segunda-feira.

(c) Todos juntos – Durante a semana.



(d) Todos juntos – Segunda-feira.

Figura 23 – Grau do Tempo – *Coach* e Todos juntos.

sete dias da semana, iniciando na segunda-feira e indo até domingo. Esse comportamento também ocorre nos graus de entrada (Apêndice-A – Figuras 32, 33 e 34) e de saída (Apêndice-A – Figuras 35, 36 e 37).

Para cada cenário calculado, existe um padrão de acontecimentos da segunda-feira à sexta-feira, isto é, o total de partidas e chegadas em cada minuto do dia é similar nesse período. O sábado e o domingo nos meios de transporte *Ferry* (Figura 22c) e *Coach* (Figura 23a) seguem o comportamento semanal, porém difere nos demais cenários. É importante ressaltar que este comportamento condiz com a realidade em uma rede de transporte, na qual existem padrões de partidas e chegadas diárias, com diferenças aos finais de semana. Assim, considerando esse padrão diário e com a finalidade de visualizar melhor os minutos dentro de um dia, selecionamos a segunda-feira e fizemos uma ampliação dela em todos cenários, representados nas Figuras 22b, 22d, 22f, 23b e 23d.

Ao observar as redes dos meios de transporte com uma quantidade menor de

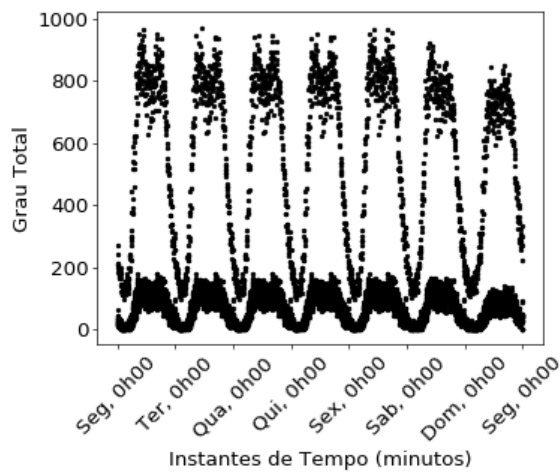
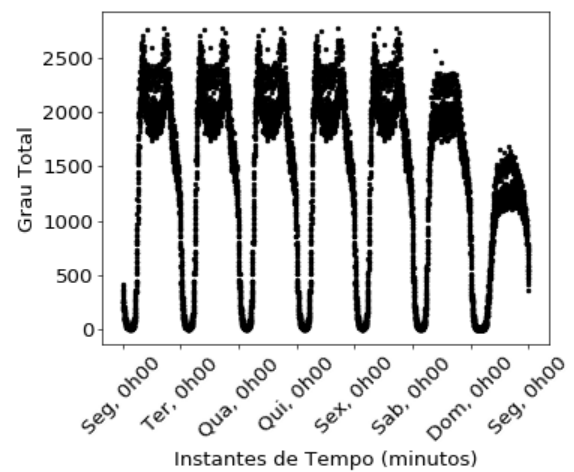
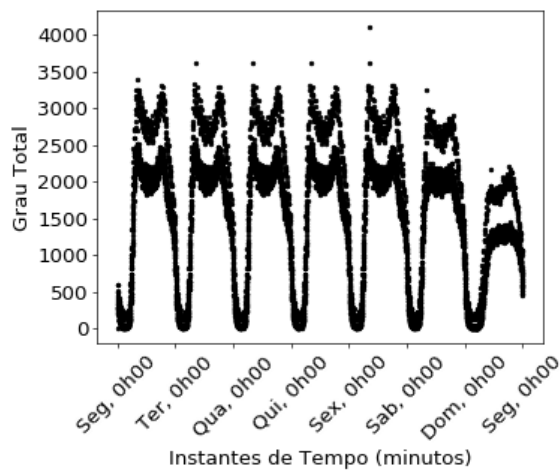
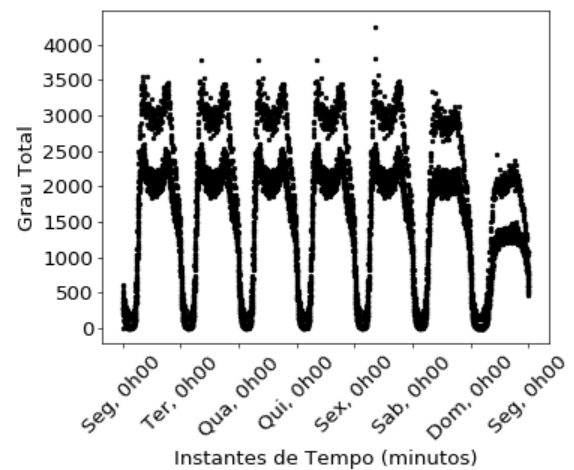
(a) Rede com *Air*, *Ferry* e *Coach*.(b) Rede com *Air*, *Ferry* e *Rail*.(c) Rede com *Air*, *Rail* e *Coach*.(d) Rede com *Ferry*, *Rail* e *Coach*.

Figura 24 – Grau Total do Tempo combinando meios de transporte três a três.

partidas e chegadas (Seção 5.1 – Tabela 2), como o *Air* (Figura 22a), visualizamos vários instantes com grau igual a zero. Isso ocorre pelo fato de ter menos viagens na rede e, nesse caso, por ter um padrão nos horários de partidas e chegadas que são a cada 5 *min*, por exemplo. Em contrapartida, ao analisarmos as redes com mais viagens encontramos poucos minutos na semana não possuem ao menos uma partida ou chegada. Por exemplo, na rede do *Rail* (Figura 22e) apenas 36 minutos dentro da semana não possuem ao menos uma partida ou chegada. Veja que a rede de cada meio de transporte possui um conjunto específico de minutos dentro de uma semana que possuem grau maior que zero. Ao analisarmos a rede com todos os meios de transporte em conjunto (Figura 23c), unimos esses conjuntos de minutos e identificamos apenas 1 minuto dentro da semana que não possui ao menos uma chegada ou partida. Esse fato demonstra um ponto importante em analisar as redes em conjunto, pois trás de uma vez todos instantes nos quais a rede completa está ativa.

Uma outra característica presente em cada gráfico semanal são os vales entre um dia e outro, que possuem poucas partidas e chegadas em cada minuto. Eles representam os períodos noturno (madrugada) entre os dias e estão bem evidenciados nas Figuras 22e e 23c. Essas representações são condizentes com a realidade já conhecida em redes de transporte públicas, nas quais é comum ter um fluxo maior de partidas e chegadas durante o dia em relação as madrugadas, as quais quase não possuem viagens.

Assim como os vales, também identificamos picos em todos os gráficos semanais. Exceto na rede do *Ferry*, todas demais possuem uma característica bimodal, com dois picos bem demarcados em cada dia. Isto é, elas são caracterizadas por grandes vales nas madrugadas, um grande crescimento nas manhãs, um pequeno decaimento durante o dia, um novo crescimento a tarde, e um grande decaimento a noite. Observamos que os horários desses picos são, geralmente, os mesmos em todos os dias da semana. Porém, seus valores de grau não são os mesmos, a quantidade de partidas e chegadas não são as mesmas mas ocorrem no mesmo horário na maioria dos dias (geralmente, de segunda-feira a sexta-feira). No Apêndice- B trazemos esses horários de picos para as redes analisadas organizados em três tabelas, sendo Grau de Entrada do Tempo na Tabela 9, Grau de Saída do Tempo na Tabela 10 e Grau Total do Tempo na Tabela 11. Consideramos instantes de tempo mais centrais por dia, os instantes em que ocorrem os picos diários dos Graus do Tempo. Por exemplo, a rede do *Air* possui dois horários importantes, sendo o horário da manhã sempre às 8 h 25 min e o da tarde sempre às 16 h 30 min. Ao visualizar cada rede isolada e a rede completa, verificamos que os horários de pico não são semelhantes para nenhuma rede e que os picos da rede completa se aproximam mais da rede *Rail* sofrendo influência das demais redes. O instante de tempo mais central do Grau Total do Tempo dentro da semana, considerando todos os meios de transporte, é sexta-feira às 8 h 25 min com um total de 4.357 partidas e chegadas. É possível visualizar esse ponto mais central da semana no gráfico da Figura 23c, pois está destacado da curva.

Ainda analisando os picos do Grau Total do Tempo, observamos que os instantes de pico são aqueles que interligam viagens, isto é, esses picos são aqueles que possuem um grande número de chegadas e partidas, indicando o fluxo de pessoas nesses horários. Por exemplo, seja uma pessoa que mora na periferia de um grande centro e trabalha nesse grande centro. Para essa pessoa chegar ao local de trabalho pode precisar pegar mais de um transporte. O horário de saída da casa dessa pessoa vai depender o quão distante ela mora da cidade destino, ao chegar em um local intermediário, ela pegaria outro transporte. As chegadas e partidas em locais intermediários como do exemplo representam os picos do Grau Total, onde muitos transportes chegam em um mesmo horário e partem em seguida.

Também analisamos os picos dos Graus de Entrada e Saída do Tempo. Os dados referentes ao Grau de Entrada do Tempo estão no Apêndice- A nas Figuras 32, 33 e 34 e no Apêndice- B na Tabela 9. Já os dados referentes ao Grau de Saída do Tempo estão no

Apêndice- A nas Figuras 35, 36 e 37 e Apêndice- B na Tabela 10). Observamos que os picos matutinos dos Graus de Entrada do Tempo estão vinculados ao horário comercial, no qual as pessoas chegam para o trabalho ou estudo. E os picos vespertinos dos Graus de Saída do Tempo estão vinculados ao final do expediente, no qual as pessoas saem do trabalho ou estudo. Essa característica é mais marcante ao analisar a rede com todos os meios de transporte.

Uma linha dupla pode ser observada nos gráficos que representam as redes *Ferry* (Figura 22c), *Coach* (Figura 23a) e o todos juntos (Figura 23c). Cada linha representa o número de partidas e chegadas por minuto durante a semana. Sendo que, a linha superior indica que em determinados minutos existem um número maior dessas partidas e chegadas ocorrendo. Esses minutos geralmente são aqueles que possuem uma periodicidade de ocorrência, por exemplo, partidas a cada 15 minutos a partir das 5 h. Já linha inferior representa as partidas e chegadas em horários menos comuns, por exemplo, uma partida às 5 h 03 min. No caso das redes *Ferry* e *Coach* a linha superior chega a ter seis vezes mais partidas e chegadas que a inferior, e que a superior de fato possui horários com uma periodicidade em vários pontos do território. Essa característica está presente na rede conjunto de todos os meios de transporte (Figura 23c), influenciada principalmente pelos meios *Ferry* e *Coach*. Observe que a rede *Rail* (Figura 22e) possui um comportamento diferenciado das demais, na qual o número de partidas e chegadas vai crescendo ao longo do dia de uma forma unificada, representada por apenas uma linha no gráfico. Fato esse decorrente das partidas e chegadas a cada minuto nessa rede, na qual é comum ter horário de partida como 8 h 48min. Esse comportamento da rede *Rail* faz com que a linha inferior da na rede com todos meios de transporte fique mais próxima a linha superior. Importante ressaltar, que essa característica não seria visualizada analisando apenas os meios de transporte isoladamente. Olhar o conjunto todo trás informações nem sempre observadas.

Nas redes *Air* (Figura 22a) e *Coach* (Figura 23a), o sábado é marcado por uma concentração maior de chegadas e partidas pela manhã. E, nas redes *Air* (Figura 22a) e *Rail* (Figura 22e), o domingo é caracterizado pelas chegadas e partidas se concentram a tarde. Essas duas características aparecem no conjunto completo e representam o diferencial dos finais de semana. Aos sábados as pessoas estão indo para suas residências e aos domingos já iniciam o processo inverso em direção aos locais de trabalho.

Com a finalidade de comparar a análise de cada rede isolada com um agrupamento delas, fizemos a análise de todas em conjunto (Figura 23c) e também fizemos o agrupamento delas considerando os meios de três a três (Figura 24). Ao comparar as quatro redes geradas desses agrupamentos, podemos perceber que o comportamento do Grau Total do Tempo segue principalmente as redes dos meios de transporte com maior número de partidas e chegadas. Observe que as características da rede *Air* (Figura 22a) não se

destacam em nenhum agrupamento (Figuras 23c e 24), inclusive o resultado do Grau Total para o agrupamento das redes *Ferry*, *Rail* e *Coach* (Figura 24d) é similar ao agrupamento com todos os meios (Figura 23c). Outro fato notório, novamente, é o impacto que a rede *Rail* gera nos agrupamentos, onde as redes que possuem o meio de transporte *Rail* (Figuras 24b, 24c, 24d e 23c) seguem as características da rede *Rail* isolada (Figura 22e). E a rede com os meios de transporte *Air*, *Ferry* e *Coach* (Figura 24a) segue as características das redes *Ferry* (Figura 22c) e *Coach* (Figura 23a).

Construímos um ranqueamento dos valores de Grau Total do Tempo por instantes de tempo, a fim de observar sua distribuição. A Figura 25 apresenta esse ranqueamento, no qual também conseguimos observar o impacto que os meios de transporte causam no conjunto. É notório que a rede *Rail* gera um impacto maior na rede completa, porém, as demais influenciam principalmente em dois momentos. Isto é, nos instantes de tempo com maiores valores de Grau Total do Tempo que é impactado por todas as redes, e nos instantes com menor Grau Total do Tempo onde o impacto vem principalmente da rede *Coach*. Afirmamos que a análise isolada de cada meio de transporte não condiz com o funcionamento da rede como um todo.

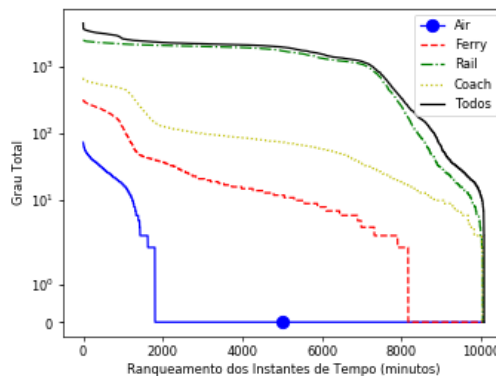


Figura 25 – Classificação dos Graus do Tempo em todos os tempos.

5.3.3 Centralidade de PageRank do Tempo

A métrica de PageRank do Tempo ranqueia os tempos considerando dois fatores: o número de instantes de tempo que apontam para o instante de tempo avaliado e a reputação que esses instantes de tempo carregam. Ou seja, um instante de tempo será bem ranqueado sempre que for apontado por uma combinação de vários instantes de tempo e por instantes com alta reputação.

Nós calculamos o PageRank para os mesmos cenários calculados na Seção 5.3.2. Assim temos gráficos do PAGERANK para rede composta pelo meio de transporte *Air* (Figura 26a), pelo *Ferry* (Figura 26c), pelo *Rail* (Figura 26e), pelo *Coach* (Figura 27a) e pela combinação de todos (Figura 27c). Todos os gráficos dessa seção possuem a relação

do valor de PageRank do Tempo pelo instante de tempo (minuto dentro de uma semana ou de um dia), considerando a partir das 0 h de Segunda-feira. Os instantes de tempo dentro da semana que não possuem atividade na rede foram representados com PageRank igual a 0 (zero).

Assim como observamos nos Graus do Tempo (Seção 5.3.2), os gráficos dos PageRanks do Tempo exibem sete agrupamentos bem definidos referente aos sete dias da semana a partir da segunda-feira até o domingo.

Além disso, também ocorre um padrão diário dos valores de PageRank durante toda a semana para cada cenário. Por essa razão, selecionamos as segundas-feiras e geramos ampliações dela em cada gráfico para visualizar mais detalhes do dia. Essas ampliações estão expressas nas Figuras 26b (*Air*), 26d (*Ferry*), 26f (*Rail*), 27b (*Coach*) e 27d (Todos os meios).

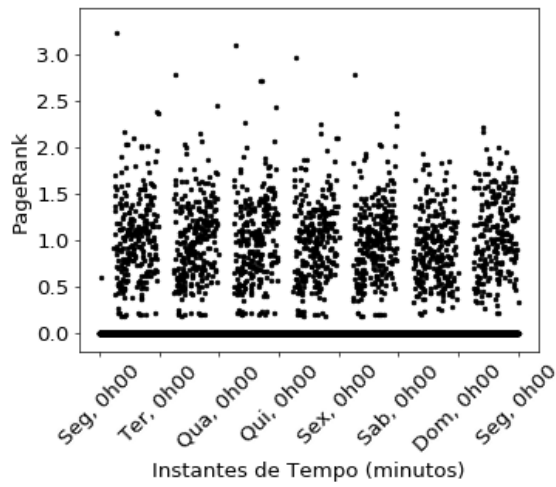
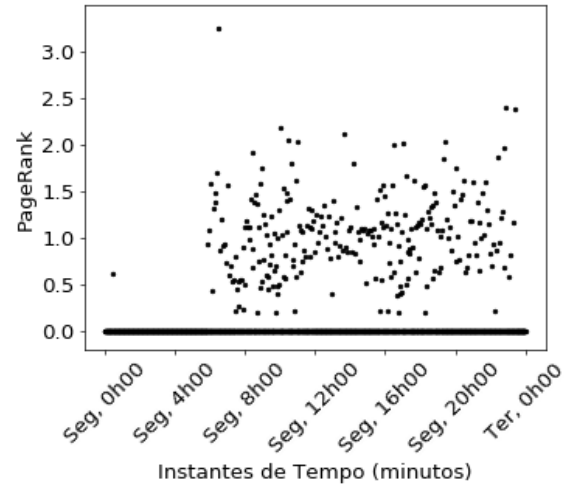
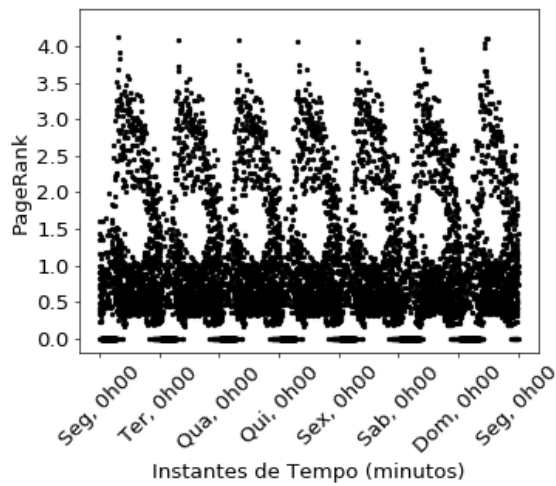
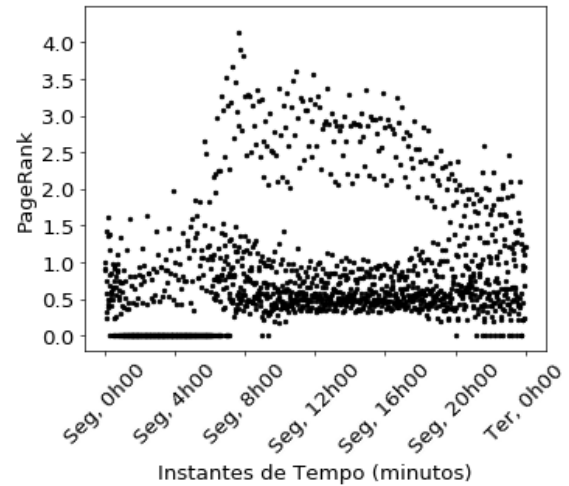
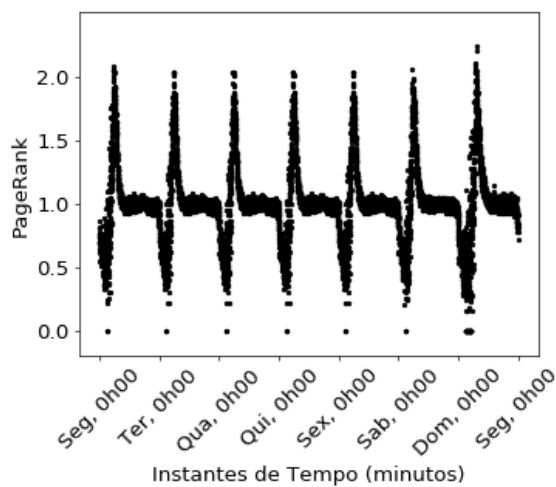
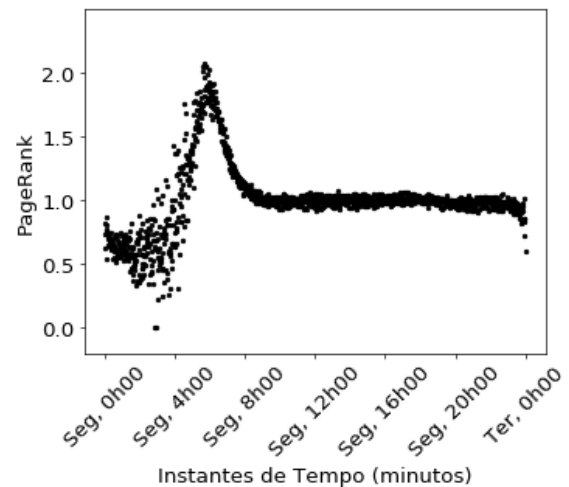
Para todo cenário é notório um pico matutino diário para o valor de PageRank do Tempo. O pico equivale ao mesmo horário em todos os dias durante a semana, de acordo com o cenário avaliado. Porém, os valores de PageRank desses picos não são os mesmos. Descrevemos os horários desses picos e seus valores de PageRank do Tempo na Tabela 7. Por exemplo, o horário 6 h 10 min é o pico de segunda-feira ao sábado na rede com todos os meios de transporte. De acordo com o ranqueamento feito pelo PageRank do Tempo, chamamos os picos de instantes de tempo mais centrais em cada dia, e o maior valor de PageRank do Tempo durante toda semana, de instante de tempo mais central da rede.

Tabela 7 – Maiores valores de PageRank diários por Meio de Transporte.

Dia	Air		Ferry		Rail		Coach		Todos	
	Horário	PR ¹	Horário	PR ¹	Horário	PR ¹	Horário	PR ¹	Horário	PR ¹
Seg	6h30	3.24	7h40	4.13	5h40	2.07	7h55	3.31	6h10	3.57
Ter	6h30	2.79	7h40	4.07	5h40	2.03	7h55	3.22	6h10	3.28
Qua	6h30	3.11	7h40	4.08	5h40	2.03	7h55	3.23	6h10	3.27
Qui	6h30	2.97	7h40	4.07	5h40	2.03	7h55	3.23	6h10	3.27
Sex	6h30	2.79	7h40	4.06	5h40	2.03	7h55	3.21	6h10	3.28
Sab	9h30	1.93	8h55	3.95	5h40	2.05	7h55	3.36	6h10	2.93
Dom	10h10	2.22	10h55	4.11	7h24	2.23	7h55	3.54	8h35	3.76

¹ PageRank

Analisando esses horários de picos nas redes que possuem apenas um meio de transporte, observamos que a rede com o *Rail* é a única com os horários de pico anteriores ao horário de pico da rede com todos os meios de transporte (Tabela 7). E como ela é uma rede com mais viagens (Seção 5.1 – Tabela 2), tem um peso maior, fazendo com que o horário do pico com todos os meios de transporte seja um horário anterior ao das demais redes dos outros meios (*Air*, *Ferry* e *Coach*). É importante observar que, a análise do PageRank em redes isoladas leva a um horário diferente da rede em conjunto. Assim, a forma tradicional de avaliar cada meio de transporte separadamente pode não trazer

(a) *Air* – Durante a semana.(b) *Air* – Segunda-feira.(c) *Ferry* – Durante a semana.(d) *Ferry* – Segunda-feira.(e) *Rail* – Durante a semana.(f) *Rail* – Segunda-feira.Figura 26 – PageRank do Tempo – *Air*, *Ferry* e *Rail*.

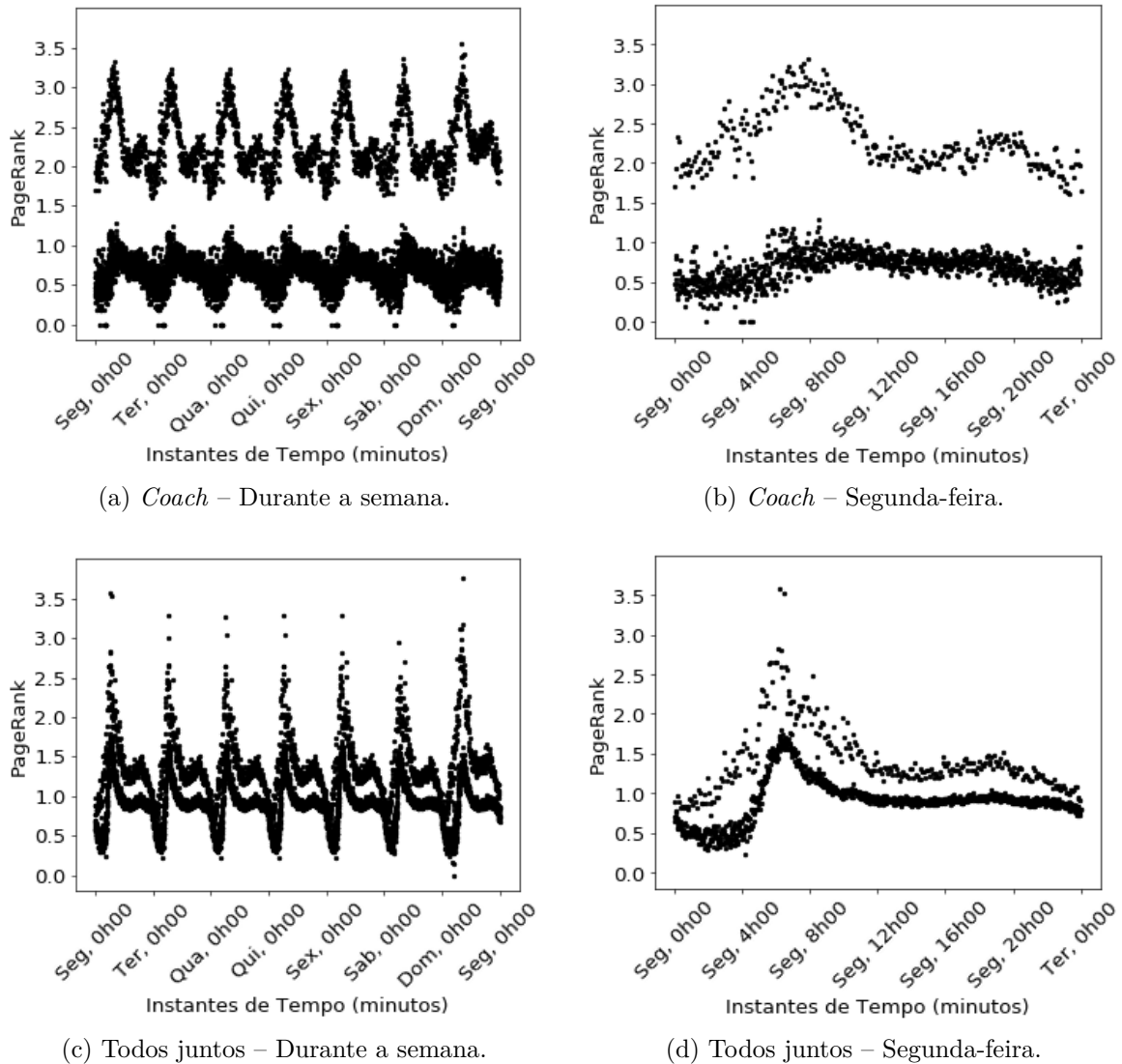


Figura 27 – PageRank do Tempo – *Coach* e Todos juntos.

horários vistos nesse trabalho, sendo importante agrupar todos os meios de transporte participantes da rede.

Observe que os valores de PageRank na rede do meio de transporte *Air* (Figura 26a) são bem distribuídos, porém eles também seguem a característica do pico matutino diário. Na rede do meio de transporte *Rail* (Figura 26e) dentro do padrão diário, visualizamos um comportamento similar a gráficos de pulsos de batimentos cardíacos. Esse gráfico explicita um decaimento abrupto durante a noite (madrugada), um crescimento também abrupto pela manhã, um decaimento suave a tarde, voltando a cair a noite. Essa característica está presente na rede com todos meios de transporte reunidos (Figura 27c). Esse pico matutino ocorre por ter um grupo significativo de instantes de tempo seguidos, com uma reputação alta, no qual um instante de tempo aponta para todos próximos dentro desse conjunto. Esse fato tem menor ocorrência no período da tarde, e não acontece a noite (madrugada).

A fim de ilustrar melhor esse cenário de como um desses picos ganha valor, apresentamos a Figura 28a que simula um grafo que tem como nó principal o pico de segunda-feira da rede com todos os meios de transporte (6 h 10 min), considerando apenas os instantes de tempo com os maiores valores de PageRank que possuem arestas para esse pico, e os instantes de tempo com os maiores valores de PageRank que possuem arestas para esses anteriores. Fizemos essa simulação para todos os dias da semana (veja os demais grafos nas figuras do Apêndice- C), e observamos que existe um padrão no qual os instantes presentes no conjunto $\{5h10, 5h15, 5h20, \dots, 6h00, 6h05\}$ apontam para os instantes superiores do conjunto e para o pico 6 h 10 min. Esse comportamento faz com que os valores de PageRank sejam acumulados até o horário de pico. Relacionamos esse fato a maior probabilidade de pessoas estarem viajando no instante de tempo desses picos.

Também geramos um grafo para um instante de tempo com valor de PageRank do Tempo intermediário (11 h 55 min) na Figura 28b. O comportamento não é similar ao pico do dia, o que reforça a ideia do acúmulo de reputação.

Além disso, é importante observar que nas figuras de ilustração dos grafos apresentadas foi representada apenas uma aresta entre dois instantes de tempo, porém podem existir várias conforme detalhado na Tabela 8, pois estamos trabalhando com um multigrafo direcionado. Nessa tabela trazemos o quantitativo de cada tipo de aresta de três horários, sendo *Seg, 4 h 11* o instante de tempo com menor valor de PageRank do Tempo diário, *Seg, 6 h 10* o pico superior diário e *Seg, 11 h 55* o instante com valor de PageRank do Tempo diário intermediário. Esse quantitativo de arestas também é importante para o ganho de reputação.

Tabela 8 – Tipos de Arestas que conectam ao instante.

Tempo	Embarque	Desembarque	Intermodal	Temporal	Deslocamento	Total
<i>Seg, 4 h 11</i>	0 (0.0%)	0 (0.0%)	0 (0.0%)	4 (80.0%)	1 (20.0%)	5
<i>Seg, 6 h 10</i>	211 (19.5%)	131 (12.1%)	259 (24.0%)	309 (28.6%)	171 (15.8%)	1.081
<i>Seg, 11 h 55</i>	281 (16.1%)	292 (16.8%)	310 (17.8%)	561 (32.2%)	299 (17.2%)	1.743

Na Figura 29 apresentamos a distribuição dos valores de PageRank do Tempo dos instantes de tempo que apontam para os três horários descritos na Tabela 8, utilizando uma ECDF (*empirical cumulative distribution function*) para cada conjunto. O instante de tempo *Seg, 4 h 11 min* com menor valor de PageRank é o destino de apenas 5 arestas. Já o instante de tempo *Seg, 6 h 10 min* com maior valor de PageRank, é o destino de 1.081 arestas. E, por fim, o instante de tempo *Seg, 11 h 55 min* que possui um valor intermediário de PageRank, é o destino de 1.743 arestas. Observe que esse último possui mais ligações que o instante de tempo pico, porém o instantes que ligam no intermediário não são tão importantes quanto os que ligam no mais central, isto é, não carregam uma reputação alta, um valor de PageRank alto.

Na rede do meio de transporte *Coach* (Figura 27a) existe um segundo pico do

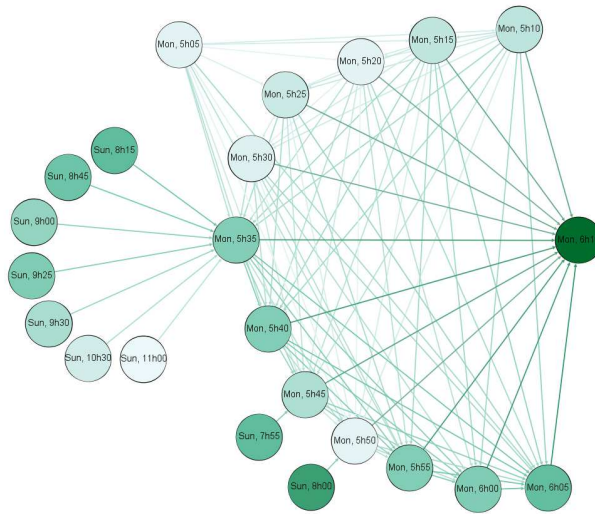
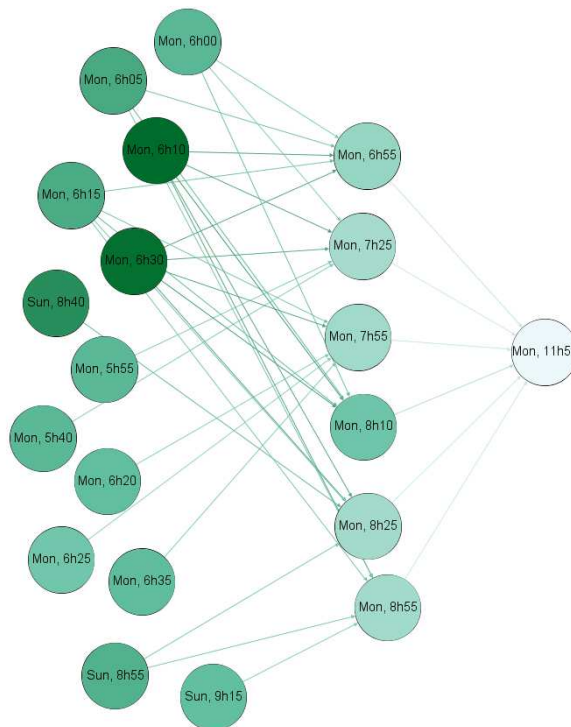
(a) Maiores PageRanks que chegam em *Seg, 6h10min*.(b) Maiores PageRanks que chegam em *Seg, 11h55min*.

Figura 28 – Grafos.

valor de PageRank presente no turno vespertino. Esse pico recebe uma importância (reputação) dos horários anteriores a ele indica o horário final de expediente comercial. Essa característica aparece na rede que combina todos os meios de transporte (Figura 27c). Além disso, o gráfico dos valores de PageRank da rede *Coach* também trás duas linhas assim como nos gráficos de Grau do Tempo (Figuras 23a, 33a e 36a). A linha superior é

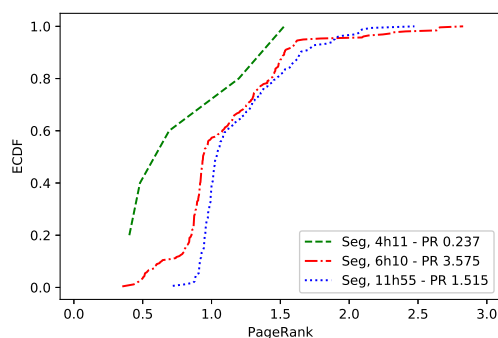


Figura 29 – PageRanks dos instantes de tempo antecessores.

caracterizada pelos horários com mais viagens se interligando e repassando suas reputações, já a inferior são dos instantes que possuem menos viagens interligadas e menores reputações. Assim, visualizamos que os horários mais padrões de partidas e chegadas dentro território do conjunto de dados, tais como 6h10 ou 10h30, terão uma reputação maior, e estarão presentes na linha superior indicando maior atividade da rede nesse instante, tanto na rede *Coach* quanto na rede com todos juntos. Em contrapartida, os horários mais incomuns, tais como 4h11 ou 5h01, terão uma reputação menor e estarão na linha inferior indicando menor número de viagens, tanto na rede *Coach* quanto na rede com todos juntos.

Os instantes de tempo com menores valores de PageRank aparecem nos vales dos gráficos, e indicam menor fluxo de viagens, menor atividade da rede. Eles estão presentes principalmente nos períodos noturnos (madrugadas), quando há um menor fluxo de viagens.

A Figura 27c apresenta o PageRank do tempo da rede composta por todos os meios de transporte citados a cima. Conseguimos visualizar o comportamento diário, a característica da rede *Rail*, e a pequena separação assim como ocorre na rede *Coach*. Porém, o horário de pico diário não é igual a nenhum meio de transporte isolado. Ao juntar todos os meios de transporte para analisar, observamos que o horário pico ficou entre o meio Rail e todos os outros. Entendemos que o Rail fez com que o horário de pico fosse mais cedo, tendo em vista o grande peso que possui na rede. O maior PageRank do Tempo dentro da rede é o horário de *Domingo, 8 h 35 min* com valor de 3.762, sendo assim considerado o mais central da semana.

Assim como na análise de ranqueamento dos valores de Grau Total do Tempo (Seção 5.3.2 - Figura 25), construímos um ranqueamento dos valores de PageRanks do Tempo por instantes de tempo para cada rede, Figura 30, a fim de observar sua distribuição. É notório que a rede *Rail* e a rede com todos seguem uma distribuição similar, porém, as demais redes influenciam na rede com todos particularmente em cerca de 20% dos instantes de tempo com maior valor de PageRank em cada. A rede *Rail* possui menos de 10% dos instantes de tempo com valores de PageRank acima de 2. Já as redes *Ferry* e

Coach possuem cerca de 20% dos instantes de tempo com valores de PageRank acima de 2. Esse fato faz com que a rede completa não acompanhe apenas a rede *Rail*, é possível observar nesse ranqueamento que a linha da rede com todos os meios se distancia da linha que representa a rede *Rail* nesse ponto.

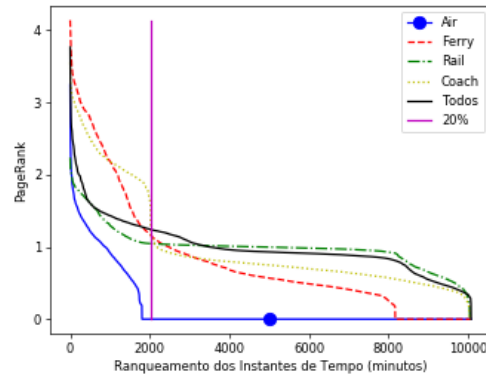
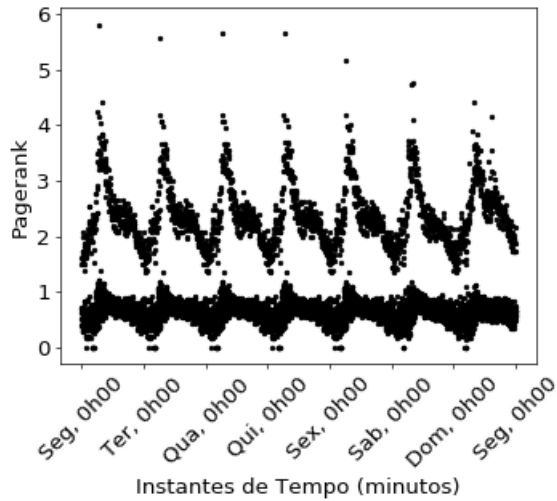
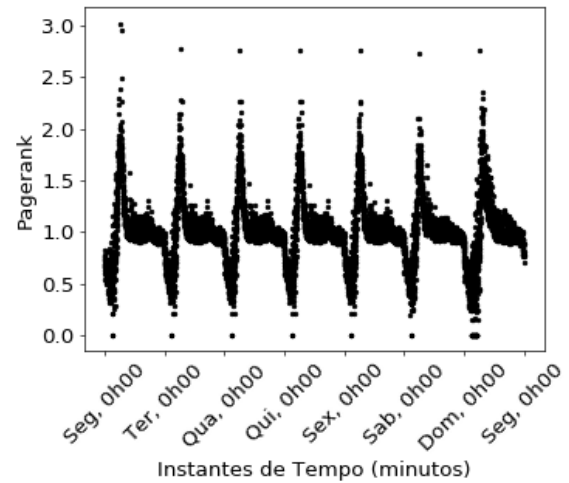


Figura 30 – Classificação dos PageRank do Tempo.

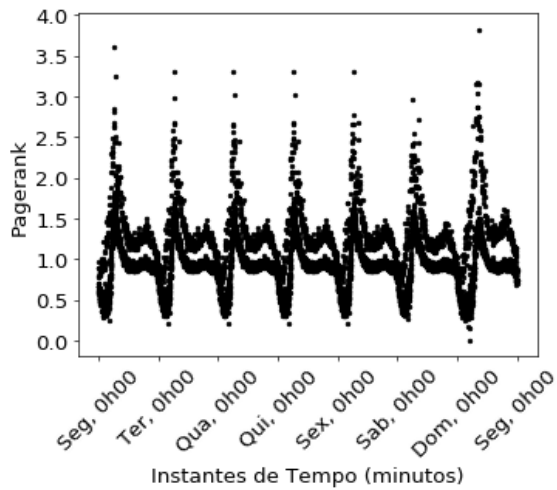
A fim de observar melhor a importância do agrupamento dos meios de transporte no sistema, geramos sub-Mags do MAG inicial combinando as camadas três a três, subdeterminamos seguindo os mesmos passos descritos na Seção 5.3.1 e calculamos o PageRank para cada combinação. A Figura 31 trás os gráficos desses PageRanks, sendo as combinações dispostas na seguinte ordem “*Air, Ferry e Rail*”, “*Air, Ferry e Coach*”, “*Air, Rail e Coach*” e “*Ferry, Rail e Coach*”. É importante observar que a presença do *Rail* faz os resultados acompanharem o seu comportamento, como visto nas Figuras 26 e 27.



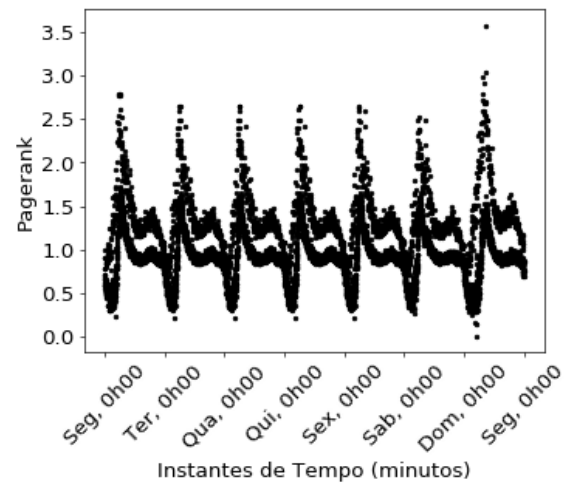
(a) Air, Ferry and Coach.



(b) Air, Ferry and Rail.



(c) Air, Rail and Coach.



(d) Ferry, Rail and Coach.

Figura 31 – PageRank do Tempo combinando meios de transporte três a três.

6 CONCLUSÕES

Neste trabalho, ampliamos a definição de Centralidade de Tempo em GVTs. De maneira similar a centralidade de nó, a centralidade de tempo avalia a importância de um dado instante de tempo em um GVT, de acordo com o objetivo a ser analisado. Os trabalhos anteriores buscavam identificar apenas os nós centrais no decorrer do tempo, e também identificar o instante de tempo para processos de difusão. Porém, também é importante identificar instantes de tempo centrais para tomadas de decisão, tendo como base a estrutura e conectividade da rede.

Até o momento, não foram identificados muitos trabalhos referentes a centralidade de tempo em GVTs. Sendo que, os localizados, tem como objetivo a identificação de instantes de tempo para realização de processos de difusão. Os demais, estão mais interessados em qualificar a importância dos nós em uma rede, do que em determinar os instantes de tempo mais importantes em um GVT.

A Centralidade de Tempo pode ser analisada de diferentes formas, de acordo com a finalidade do estudo. Nesse trabalho, propomos duas métricas baseadas em conectividade: *Grau do Tempo* e *PageRank do Tempo*. Ambas com intuito de identificar instantes da rede mais ativos. Para avaliar nossas métricas, aplicamos a um conjunto de dados que representa uma semana de viagens em uma Rede de Transporte Público real, considerando os meios de transporte aéreo, trem, balsa e ônibus interurbano. Com as métricas foi possível elencar os momentos dos dias durante a semana que possuem maiores e menores atividades, além de identificar os horários mais ativos de cada dia.

Sabendo que muitas redes modeladas por GVTs são notoriamente cíclicas, a Centralidade de Tempo pode ser útil para construção de previsões de comportamentos futuros em GVTs, tanto pelos padrões de conexões, quanto pelo aparecimento ou movimentação de nós ao longo do tempo. Características dos instantes de tempo do passado podem gerar um modelo de previsão de instantes de tempo importante em potencial, uma vez que as características podem tornar a acontecer.

Os resultados avaliados na rede de transporte corroboram com a realidade desse tipo de rede. Os instantes de tempo elencados como mais centrais (importantes) são de fato os instantes que possuem mais atividade de partidas e chegadas na rede. Além disso, as métricas expressam o comportamento diário dentro de uma semana, com horários de pico correspondentes a início e fim de expedientes de trabalho, bem como os instantes de tempo com menores avaliações indicam horários com menor fluxo de viagens.

Além disso, mostramos a importância em se avaliar a rede considerando todos os meios de transporte envolvidos. Pois, as diferentes características de cada meio de transporte impactam na rede completa, com todos os meios de transporte envolvidos. Vimos que, ao analisar a rede em conjunto, encontramos instantes de tempo importantes os quais

ainda não haviam sido elencados, ao analisar cada meio de transporte individualmente.

A identificação dos instantes de tempo importantes demandam um conhecimento prévio das características da rede, bem como de sua evolução ao longo do tempo. Porém, esse fato não é um limitador, pois sabemos que muitas redes modeladas por GVTs são certamente cíclicas. Vimos isso no nosso estudo de caso aplicado em uma rede de transporte.

Dentro de centralidade do tempo ainda há muito para ser pesquisado. Como principais trabalhos futuros, trazemos a proposta de analisar o impacto de interrupções em determinados instantes de tempo, realizando perturbações na rede a fim de observar o comportamento do sistema a partir do cálculo das centralidades. Nesse caso, levaria em consideração algo que gerasse um impacto na rede completa no mesmo instante de tempo, por exemplo, algo que impactaria uma grande parte da rede de transporte seria condições climáticas, tais como fortes nevascas.

Outra proposta seria a continuidade das avaliações e criações de métricas para Centralidade do Tempo em GVTs, assim como existem diversas métricas de centralidade tradicionais, buscando qualificar instantes de tempo importantes a partir de diferentes redes e objetivos.

REFERÊNCIAS

- BANG-JENSEN, J.; GUTIN, G. Z. **Digraphs - Theory, Algorithms and Applications.**, 2007.
- BAVELAS, A. Communication patterns in task-oriented groups. **The Journal of the Acoustical Society of America**, v. 22, n. 6, p. 725–730, 1950.
- BONDY, J. A.; MURTY, U. S. R. **Graph Theory**, 2008. (Graduate Texts in Mathematics, v. 244).
- BRIN, S.; PAGE, L. The anatomy of a large-scale hypertextual web search engine. **Computer Networks and ISDN Systems**, v. 30, n. 1, p. 107 – 117, 1998. Proceedings of the Seventh International World Wide Web Conference.
- BUI-XUAN, B.-M.; FERREIRA, A.; JARRY, A. Evolving graphs and least cost journeys in dynamic networks. In: **WiOpt'03: Modeling and Optimization in Mobile, Ad Hoc and Wireless Networks**, 2003. p. 10 pages.
- CASTEIGTS, A.; FLOCCHINI, P.; QUATTROCIOCCHI, W.; SANTORO, N. Time-varying graphs and dynamic networks. **International Journal of Parallel, Emergent and Distributed Systems**, Taylor & Francis, v. 27, n. 5, p. 387–408, 2012.
- COSTA, E. C.; VIEIRA, A. B.; WEHMUTH, K.; ZIVIANI, A.; SILVA, A. P. C. D. Time centrality in dynamic complex networks. **Advances in Complex Systems (ACS)**, v. 18, n. 07n08, 11 and 12 2015.
- DIESTEL, R. **Graph Theory**. 4. ed., 2010. (Graduate Texts in Mathematics, v. 173).
- EVANS, J. R.; MINIEKA, E. **Optimization Algorithms for Networks and Graphs - 2nd ed., Revised and Expanded**. 2. ed., 1992.
- FERREIRA, A. Building a reference combinatorial model for manets. **Netwrk. Mag. of Global Internetwkg.**, IEEE Press, Piscataway, NJ, USA, v. 18, n. 5, p. 24–29, set. 2004.
- FREEMAN, L. C. A set of measures of centrality based on betweenness. **Sociometry**, American Sociological Association, Sage Publications, Inc., v. 40, n. 1, p. 35–41, 1977.
- FREEMAN, L. C. Centrality in social networks conceptual clarification. **Social Networks**, p. 215, 1978.
- GALLOTTI, R.; BARTHELEMY, M. The multilayer temporal network of public transport in great britain. **Scientific Data**, v. 2, 2015.

- HOLME, P. Network reachability of real-world contact sequences. **Phys. Rev. E**, American Physical Society, v. 71, p. 046119, Apr 2005.
- HOLME, P.; SARAMÄKI, J. Temporal networks. **Physics Reports**, v. 519, n. 3, p. 97 – 125, 2012. Temporal Networks.
- KIM, H.; ANDERSON, R. Temporal node centrality in complex networks. **Phys. Rev. E**, American Physical Society, v. 85, p. 026107, Feb 2012.
- KIVELÄ, M.; ARENAS, A.; BARTHELEMY, M.; GLEESON, J. P.; MORENO, Y.; PORTER, M. A. Multilayer networks. **Journal of Complex Networks**, v. 2, n. 3, p. 203–271, 2014.
- KOSTAKOS, V. Temporal graphs. **Physica A: Statistical Mechanics and its Applications**, v. 388, n. 6, p. 1007 – 1023, 2009.
- NEWMAN, M. **Networks: An Introduction**, 2010.
- NICOSIA, V.; TANG, J. K.; MASCOLO, C.; MUSOLESI, M.; RUSSO, G.; LATORA, V. Graph metrics for temporal networks. **CoRR**, abs/1306.0493, 2013.
- SABIDUSSI, G. The centrality index of a graph. **Psychometrika**, v. 31, n. 4, p. 581–603, Dec 1966.
- STROGATZ, S. H. Exploring complex networks. **Nature**, v. 410, p. 268–276, 2001.
- TANG, J.; MUSOLESI, M.; MASCOLO, C.; LATORA, V. Temporal distance metrics for social network analysis. In: **Proceedings of the 2ND ACM SIGCOMM Workshop on Online Social Networks (WOSN '09)**, 2009. p. 31–36.
- TANG, J.; MUSOLESI, M.; MASCOLO, C.; LATORA, V.; NICOSIA, V. Analysing information flows and key mediators through temporal centrality metrics. In: **Proceedings of the 3rd Workshop on Social Network Systems**, 2010. (SNS '10), p. 3:1–3:6.
- TANG, J.; SCELLATO, S.; MUSOLESI, M.; MASCOLO, C.; LATORA, V. Small-world behavior in time-varying graphs. **Physical review. E, Statistical, nonlinear, and soft matter physics**, v. 81, p. 055101, 05 2010.
- UK. **National Public Transport Data Repository (NPTDR)**. 2017. Data do acesso: 06 de setembro de 2017. Disponível em: <<https://data.gov.uk/dataset/nptdr>>.
- USA. **Innovata LLC**. 2017. Data do acesso: 06 de setembro de 2017. Disponível em: <<http://www.innovata-llc.com>>.
- WANG, Z.; PEI, X.; WANG, Y.; YAO, Y. Ranking the key nodes with temporal degree deviation centrality on complex networks. In: **2017 29th Chinese Control And Decision Conference (CCDC)**, 2017. p. 1484–1489.

- WEHMUTH, K.; FLEURY, E.; ZIVIANI, A. On multiaspect graphs. **Theoretical Computer Science**, v. 651, p. 50–61, 2016.
- WEHMUTH, K.; FLEURY, E.; ZIVIANI, A. Multiaspect graphs: Algebraic representation and algorithms. **Algorithms, MDPI**, v. 10, n. 1, p. 1–36, 2017.
- WEHMUTH, K.; ZIVIANI, A. Avoiding spurious paths in centralities based on shortest paths in high order networks. In: **2018 Eighth Latin-American Symposium on Dependable Computing (LADC)**, 2018.
- WEHMUTH, K.; ZIVIANI, A.; FLEURY, E. A unifying model for representing time-varying graphs. In: **2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)**, 2015. p. 1–10.

A GRÁFICOS DOS GRAUS DO TEMPO

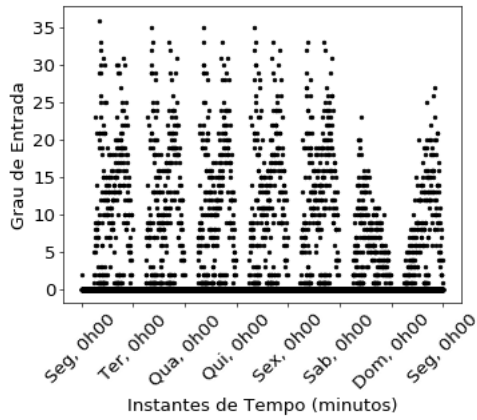
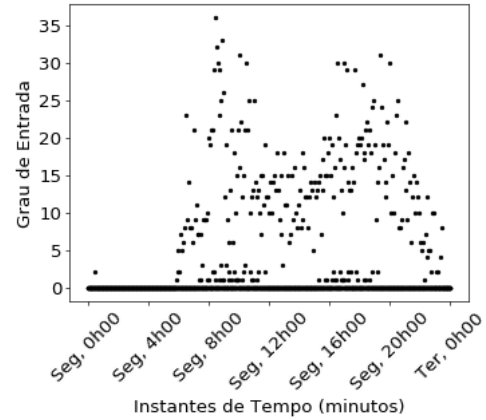
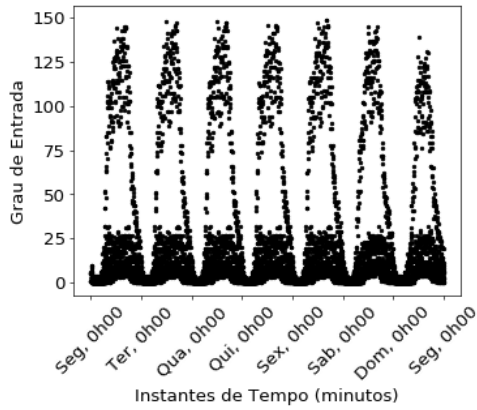
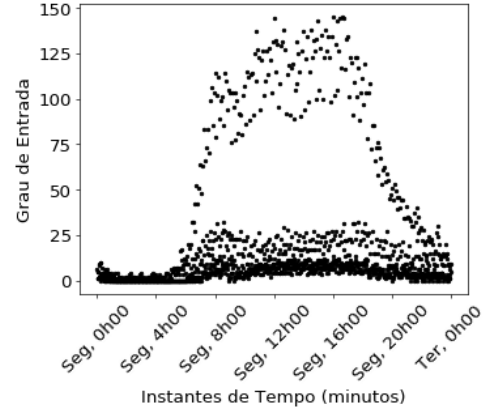
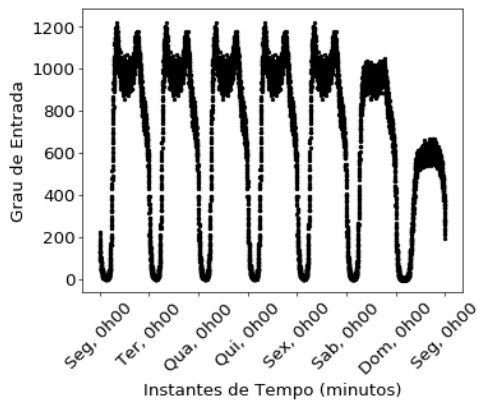
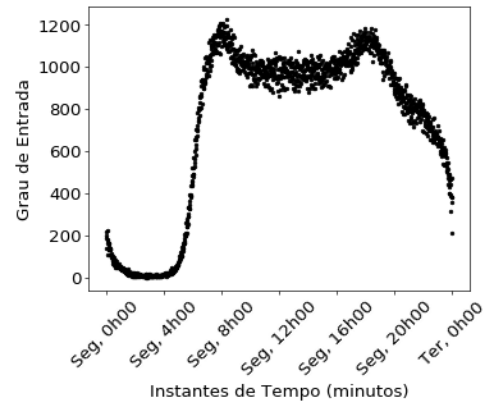
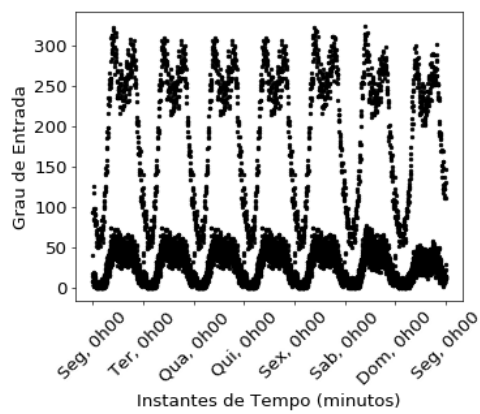
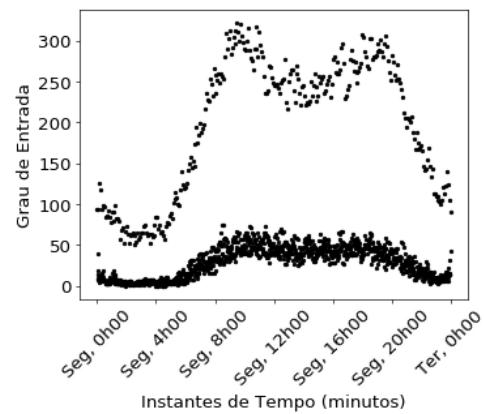
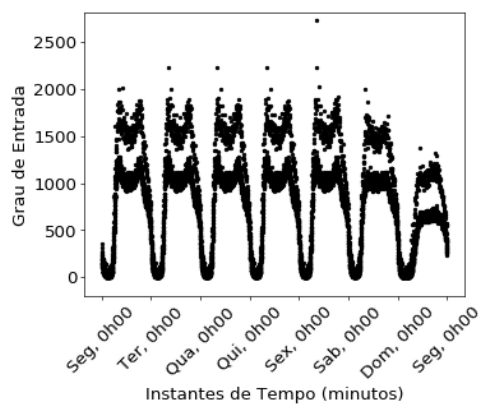
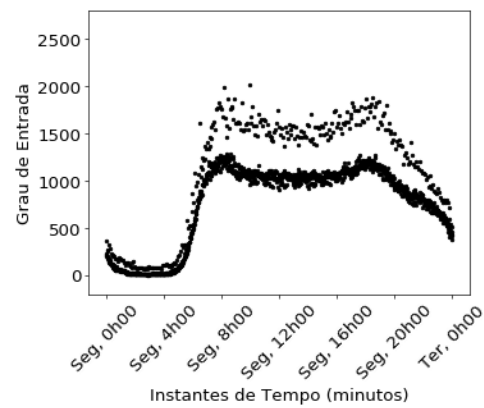
(a) *Air* – Durante a semana.(b) *Air* – Segunda-feira.(c) *Ferry* – Durante a semana.(d) *Ferry* – Segunda-feira.(e) *Rail* – Durante a semana.(f) *Rail* – Segunda-feira.

Figura 32 – Grau de Entrada do Tempo – *Air*, *Ferry* e *Rail*.

(a) *Coach* – Durante a semana.(b) *Coach* – Segunda-feira.

(c) Todos juntos – Durante a semana.



(d) Todos juntos – Segunda-feira.

Figura 33 – Grau de Entrada do Tempo – *Coach* e Todos juntos.

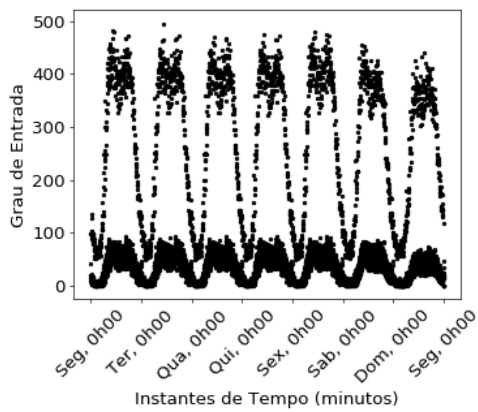
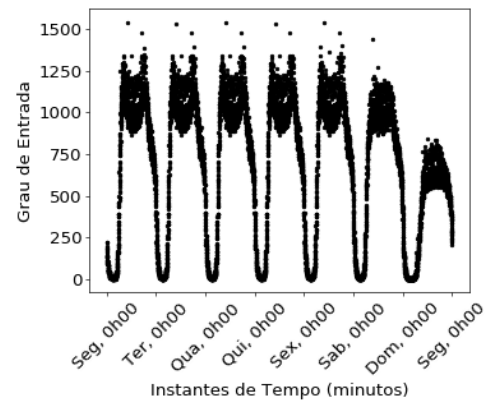
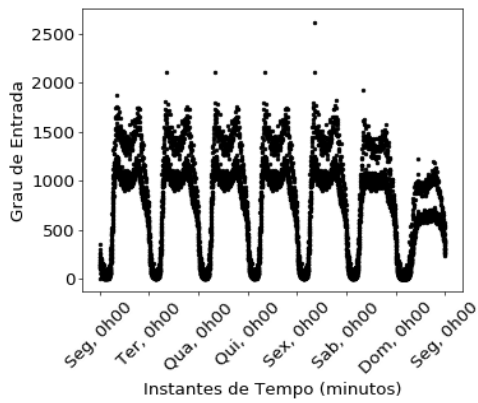
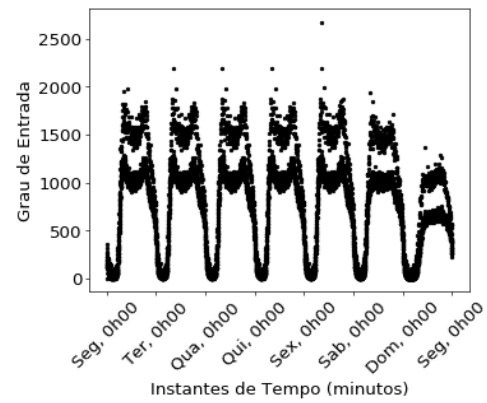
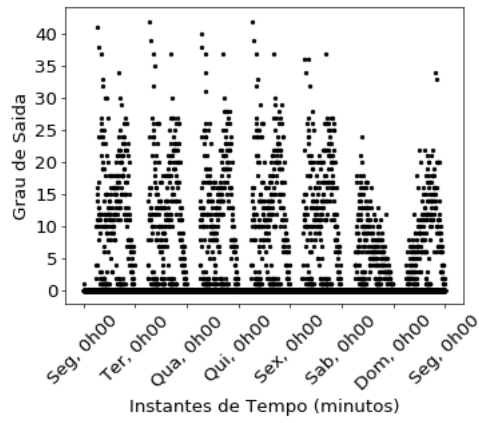
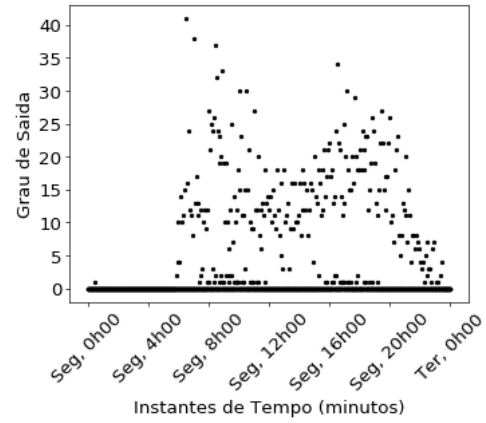
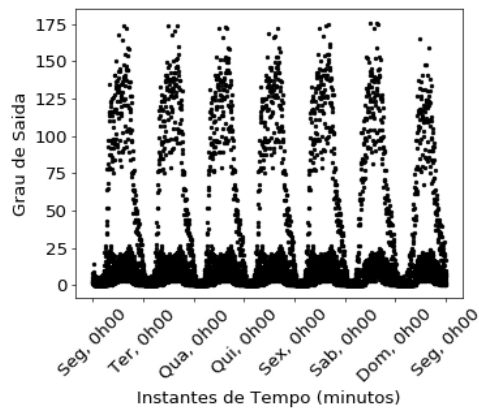
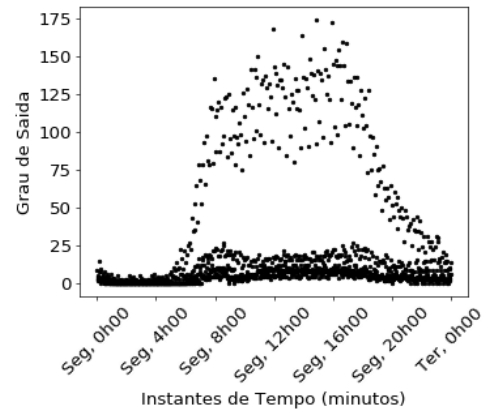
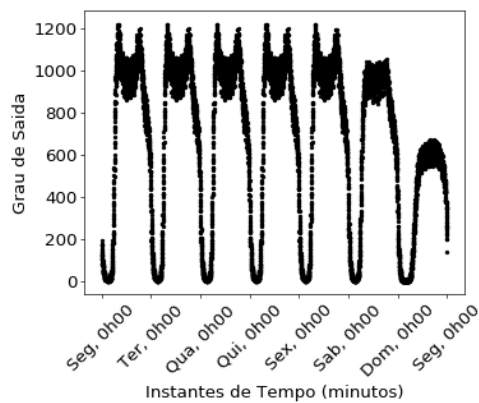
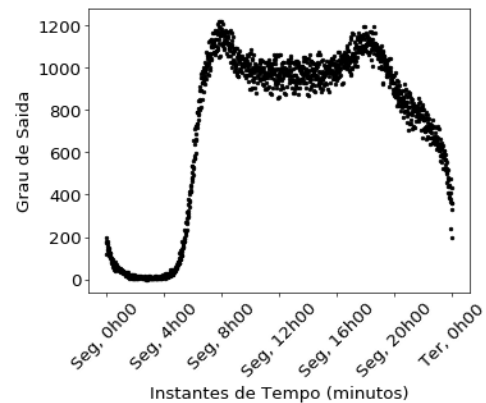
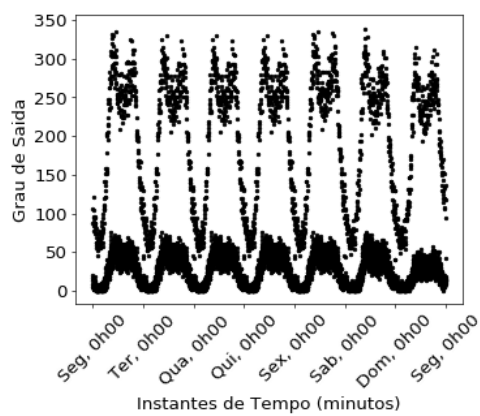
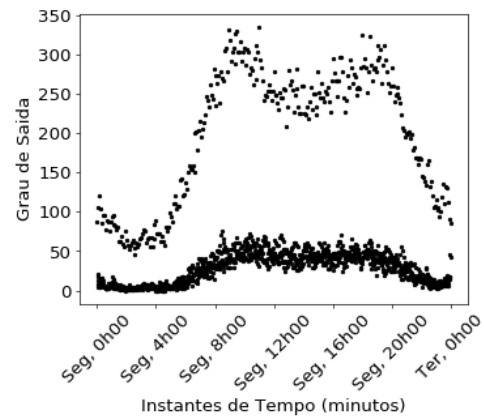
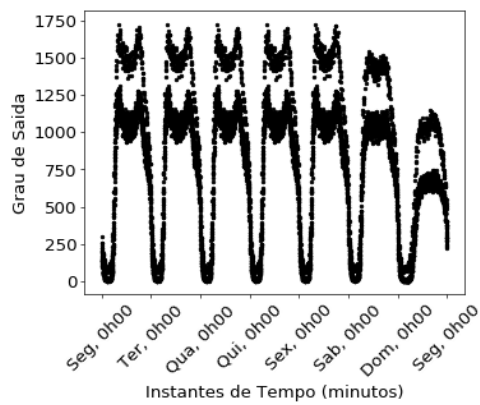
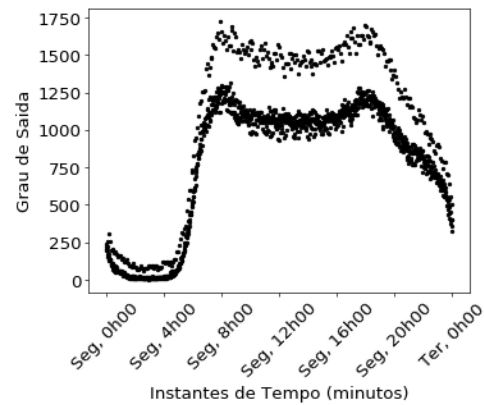
(a) *Air, Ferry e Coach.*(b) *Air, Ferry e Rail.*(c) *Air, Rail e Coach.*(d) *Ferry, Rail e Coach.*

Figura 34 – Grau de Entrada do Tempo combinando meios de transporte três a três.

(a) *Air* – Durante a semana.(b) *Air* – Segunda-feira.(c) *Ferry* – Durante a semana.(d) *Ferry* – Segunda-feira.(e) *Rail* – Durante a semana.(f) *Rail* – Segunda-feira.Figura 35 – Grau de Saída do Tempo – *Air*, *Ferry* e *Rail*.

(a) *Coach* – Durante a semana.(b) *Coach* – Segunda-feira.

(c) Todos juntos – Durante a semana.



(d) Todos juntos – Segunda-feira.

Figura 36 – Grau de Saída do Tempo – *Coach* e Todos juntos.

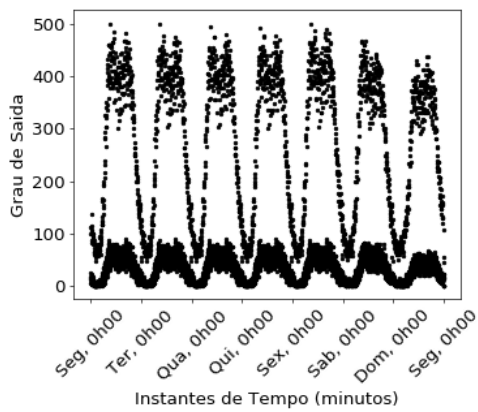
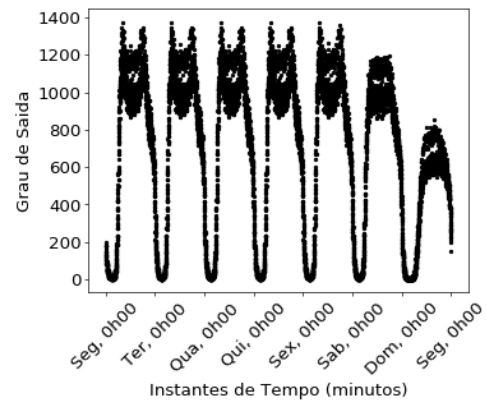
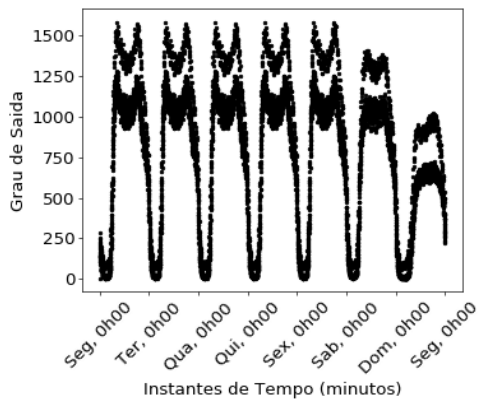
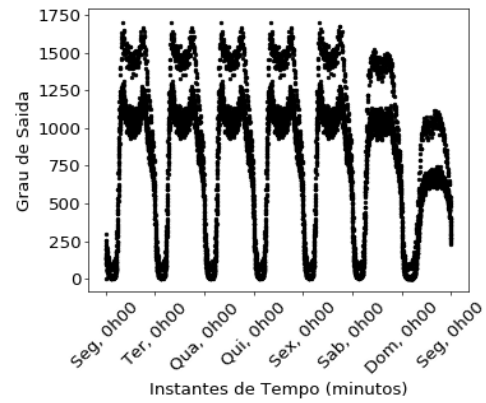
(a) *Air, Ferry e Coach.*(b) *Air, Ferry e Rail.*(c) *Air, Rail e Coach.*(d) *Ferry, Rail e Coach.*

Figura 37 – Grau de Saída do Tempo combinando meios de transporte três a três.

B TABELAS DOS PICOS DOS GRAUS DO TEMPO

Tabela 9 – Graus de Entrada do Tempo.

Grau de Entrada do Tempo	Instante de Tempo Rede <i>Air</i>	Grau de Entrada do Tempo	Instante de Tempo Rede <i>Air</i>
36*	Segunda-feira, 08h25	31	Segunda-feira, 19h20
35	Terça-feira, 08h25	33	Terça-feira, 16h30
35	Quarta-feira, 08h25	33	Quarta-feira, 17h00
35	Quinta-feira, 08h25	33	Quinta-feira, 17h00
33	Sexta-feira, 08h55	33	Sexta-feira, 16h30
23	Sábado, 10h00	16	Sábado, 13h00

Grau de Entrada do Tempo	Instante de Tempo Rede <i>Ferry</i>	Grau de Entrada do Tempo	Instante de Tempo Rede <i>Ferry</i>
144	Segunda-feira, 12h00	145	Segunda-feira, 16h40
148	Terça-feira, 12h00	147	Terça-feira, 16h40
148	Quarta-feira, 12h00	147	Quarta-feira, 16h15 e 16h30
146	Quinta-feira, 12h00	146	Quinta-feira, 16h15
147	Sexta-feira, 12h00	149*	Sexta-feira, 16h00 e 16h15
145	Sábado, 12h00	145	Sábado, 16h30

Grau de Entrada do Tempo	Instante de Tempo Rede <i>Rail</i>	Grau de Entrada do Tempo	Instante de Tempo Rede <i>Rail</i>
1224*	Segunda-feira, 08h19	1180	Segunda-feira, 17h55 e 18h25
1224*	Terça-feira, 08h19	1180	Terça-feira, 17h55 e 18h25
1224*	Quarta-feira, 08h19	1180	Quarta-feira, 17h55 e 18h25
1224*	Quinta-feira, 08h19	1180	Quinta-feira, 17h55 e 18h25
1224*	Sexta-feira, 08h19	1181	Sexta-feira, 17h55
1042	Sábado, 10h10	1052	Sábado, 18h28

Grau de Entrada do Tempo	Instante de Tempo Rede <i>Coach</i>	Grau de Entrada do Tempo	Instante de Tempo Rede <i>Coach</i>
322	Segunda-feira, 09h30	308	Segunda-feira, 18h00
309	Terça-feira, 10h15	306	Terça-feira, 18h00
309	Quarta-feira, 10h15	306	Quarta-feira, 18h00
309	Quinta-feira, 10h15	306	Quinta-feira, 18h00
322	Sexta-feira, 09h30	312	Sexta-feira, 18h00 e 18h05
325*	Sábado, 09h30	298	Sábado, 19h35

Grau de Entrada do Tempo	Instante de Tempo Rede com todos	Grau de Entrada do Tempo	Instante de Tempo Rede com todos
2013	Segunda-feira, 09h55	1872	Segunda-feira, 18h30
2219	Terça-feira, 08h10	1884	Terça-feira, 18h30
2220	Quarta-feira, 08h10	1889	Quarta-feira, 18h30
2218	Quinta-feira, 08h10	1870	Quinta-feira, 18h30
2731*	Sexta-feira, 08h25	1914	Sexta-feira, 18h55
2000	Sábado, 08h10	1716	Sábado, 18h55

Tabela 10 – Graus de Saída do Tempo.

Grau de Saída do Tempo	Instante de Tempo Rede <i>Air</i>	Grau de Saída do Tempo	Instante de Tempo Rede <i>Air</i>
41	Segunda-feira, 06h30	34	Segunda-feira, 16h30
42*	Terça-feira, 06h30	37	Terça-feira, 16h30
40	Quarta-feira, 06h30	37	Quarta-feira, 16h30
42*	Quinta-feira, 06h30	37	Quinta-feira, 16h30
36	Sexta-feira, 6h30	37	Sexta-feira, 16h30
24	Sábado, 09h30	16	Sábado, 13h30
18	Domingo, 11h00	34	Domingo, 19h25

Grau de Saída do Tempo	Instante de Tempo Rede <i>Ferry</i>	Grau de Saída do Tempo	Instante de Tempo Rede <i>Ferry</i>
168	Segunda-feira, 11h55	174	Segunda-feira, 14h55
174	Terça-feira, 11h55	174	Terça-feira, 15h55
172	Quarta-feira, 11h55	173	Quarta-feira, 14h55
169	Quinta-feira, 11h55	172	Quinta-feira, 15h55
172	Sexta-feira, 11h55	175	Sexta-feira, 15h55
176*	Sábado, 11h55	176*	Sábado, 14h55

Grau de Saída do Tempo	Instante de Tempo Rede <i>Rail</i>	Grau de Saída do Tempo	Instante de Tempo Rede <i>Rail</i>
1220	Segunda-feira, 07h50	1198	Segunda-feira, 18h35
1221*	Terça-feira, 07h50	1198	Terça-feira, 18h35
1221*	Quarta-feira, 07h50	1198	Quarta-feira, 18h35
1221*	Quinta-feira, 07h50	1198	Quinta-feira, 18h35
1221*	Sexta-feira, 07h50	1196	Sexta-feira, 18h35
1042	Sábado, 8h40	1053	Sábado, 18h23

Grau de Saída do Tempo	Instante de Tempo Rede <i>Coach</i>	Grau de Saída do Tempo	Instante de Tempo Rede <i>Coach</i>
335	Segunda-feira, 11h00	325	Segunda-feira, 18h00
329	Terça-feira, 11h00	323	Terça-feira, 18h00
329	Quarta-feira, 11h00	323	Quarta-feira, 18h00
329	Quinta-feira, 11h00	323	Quinta-feira, 18h00
335	Sexta-feira, 11h00	334	Sexta-feira, 18h00
339*	Sábado, 09h30	309	Sábado, 18h30

Grau de Saída do Tempo	Instante de Tempo Rede com todos	Grau de Saída do Tempo	Instante de Tempo Rede com todos
1725*	Segunda-feira, 07h55	1701	Segunda-feira, 17h55
1721	Terça-feira, 07h55	1698	Terça-feira, 17h55
1719	Quarta-feira, 07h55	1697	Quarta-feira, 17h55
1721	Quinta-feira, 07h55	1699	Quinta-feira, 17h55
1721	Sexta-feira, 07h55	1713	Sexta-feira, 17h55
1539	Sábado, 10h10	1514	Sábado, 16h40

Tabela 11 – Graus Totais do Tempo.

Grau Total do Tempo	Instante de Tempo Rede <i>Air</i>	Grau Total do Tempo	Instante de Tempo Rede <i>Air</i>
73*	Segunda-feira, 08h25	64	Segunda-feira, 16h30
72	Terça-feira, 08h25	70	Terça-feira, 16h30
72	Quarta-feira, 08h25	68	Quarta-feira, 16h30
72	Quinta-feira, 08h25	68	Quinta-feira, 16h30
68	Sexta-feira, 08h25	70	Sexta-feira, 16h30
44	Sábado, 09h30	31	Sábado, 13h30
31	Domingo, 11h30	60	Domingo, 19h25 e 19h55

Grau Total do Tempo	Instante de Tempo Rede <i>Ferry</i>	Grau Total do Tempo	Instante de Tempo Rede <i>Ferry</i>
299	Segunda-feira, 11h55	307	Segunda-feira, 14h55
309	Terça-feira, 11h55	309	Terça-feira, 15h55
306	Quarta-feira, 11h55	307	Quarta-feira, 14h55
301	Quinta-feira, 11h55	306	Quinta-feira, 15h55
306	Sexta-feira, 11h55	310	Sexta-feira, 14h55 e 15h55
317*	Sábado, 11h55	312	Sábado, 14h55

Grau Total do Tempo	Instante de Tempo Rede <i>Rail</i>	Grau Total do Tempo	Instante de Tempo Rede <i>Rail</i>
2422*	Segunda-feira, 07h55 e 07h59	2373	Segunda-feira, 17h55
2422*	Terça-feira, 07h55 e 07h59	2373	Terça-feira, 17h55
2422*	Quarta-feira, 07h55 e 07h59	2373	Quarta-feira, 17h55
2422*	Quinta-feira, 07h55 e 07h59	2373	Quinta-feira, 17h55
2422*	Sexta-feira, 07h55 e 07h59	2376	Sexta-feira, 17h55
2074	Sábado, 10h10	2089	Sábado, 18h28

Grau Total do Tempo	Instante de Tempo Rede <i>Coach</i>	Grau Total do Tempo	Instante de Tempo Rede <i>Coach</i>
652	Segunda-feira, 09h30	633	Segunda-feira, 18h00
636	Terça-feira, 11h00	629	Terça-feira, 18h00
636	Quarta-feira, 11h00	629	Quarta-feira, 18h00
636	Quinta-feira, 11h00	629	Quinta-feira, 18h00
652	Sexta-feira, 09h30	646	Sexta-feira, 18h00
664*	Sábado, 09h30	602	Sábado, 18h30

Grau Total do Tempo	Instante de Tempo Rede com todos	Grau Total do Tempo	Instante de Tempo Rede com todos
3621	Segunda-feira, 08h10	3562	Segunda-feira, 18h00
3850	Terça-feira, 08h10	3547	Terça-feira, 18h00
3852	Quarta-feira, 08h10	3540	Quarta-feira, 18h30
3849	Quinta-feira, 08h10	3551	Quinta-feira, 18h00
4357*	Sexta-feira, 08h25	3572	Sexta-feira, 18h00
3413	Sábado, 08h10	3142	Sábado, 18h55

C OUTROS GRAFOS SIMULADOS

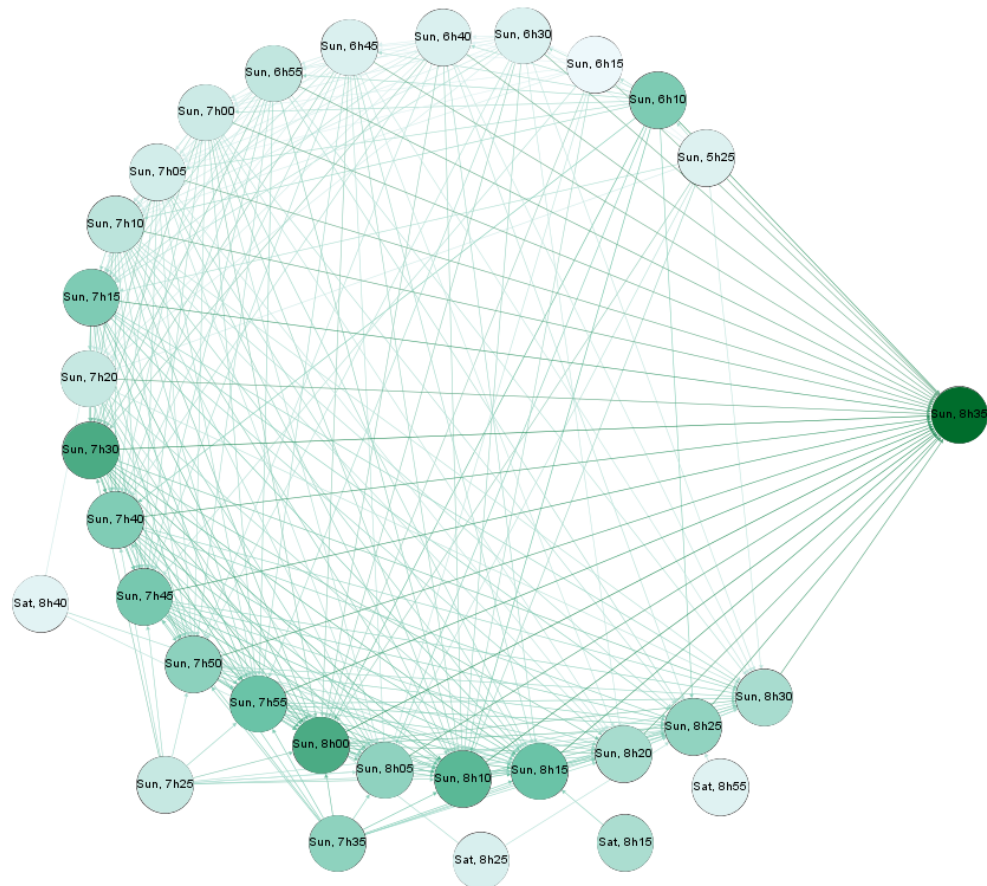


Figura 38 – Maiores PageRanks que chegam no *Domingo*, 8 h 35 min.

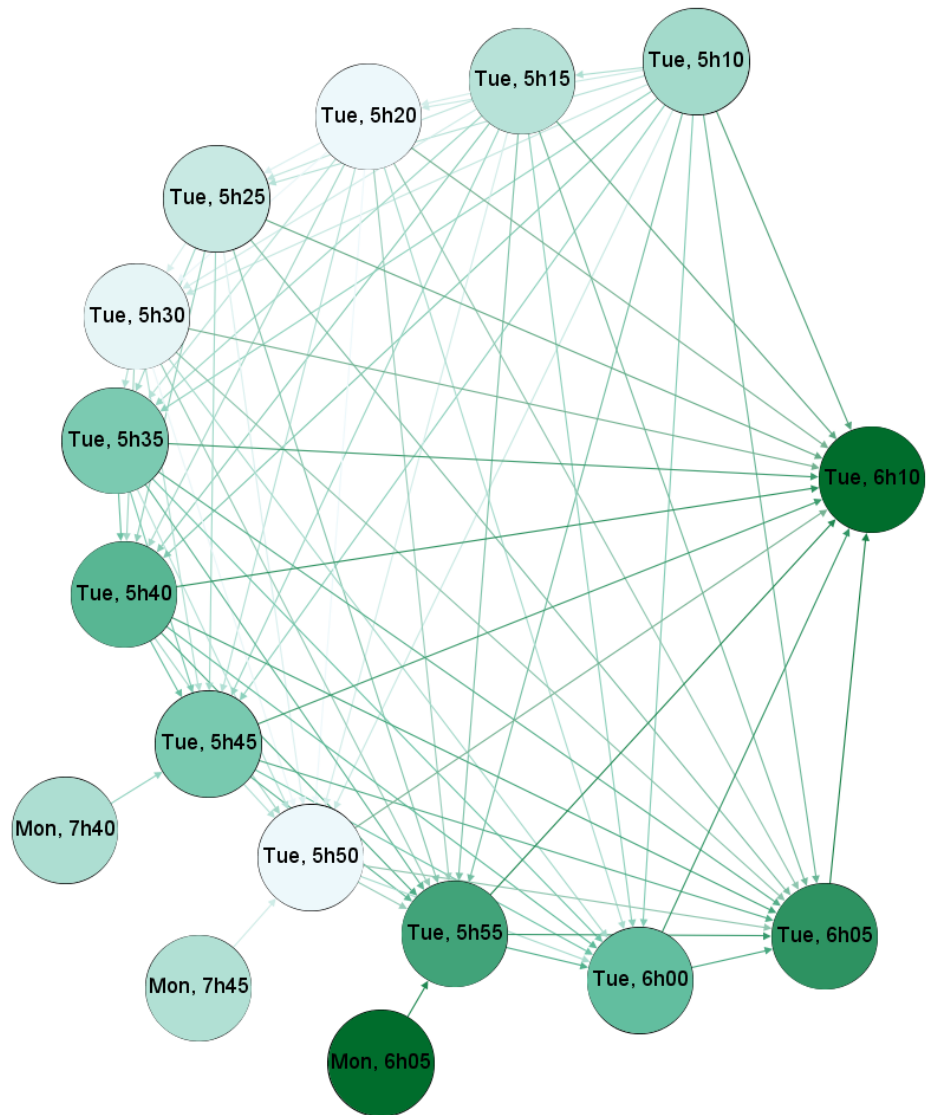


Figura 39 – Maiores PageRanks que chegam na *Terça – feira, 6 h 10 min.*

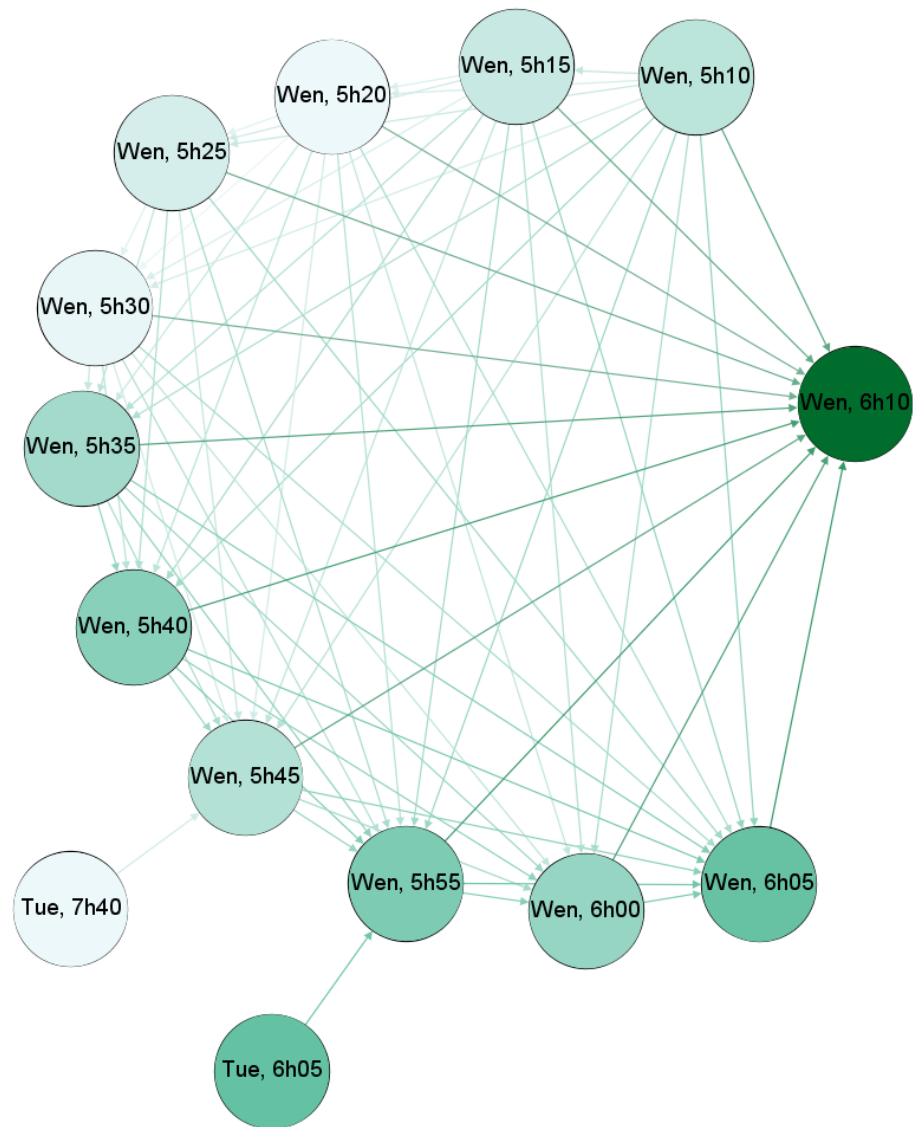


Figura 40 – Maiores PageRanks que chegam na *Quarta – feira, 6 h 10 min.*

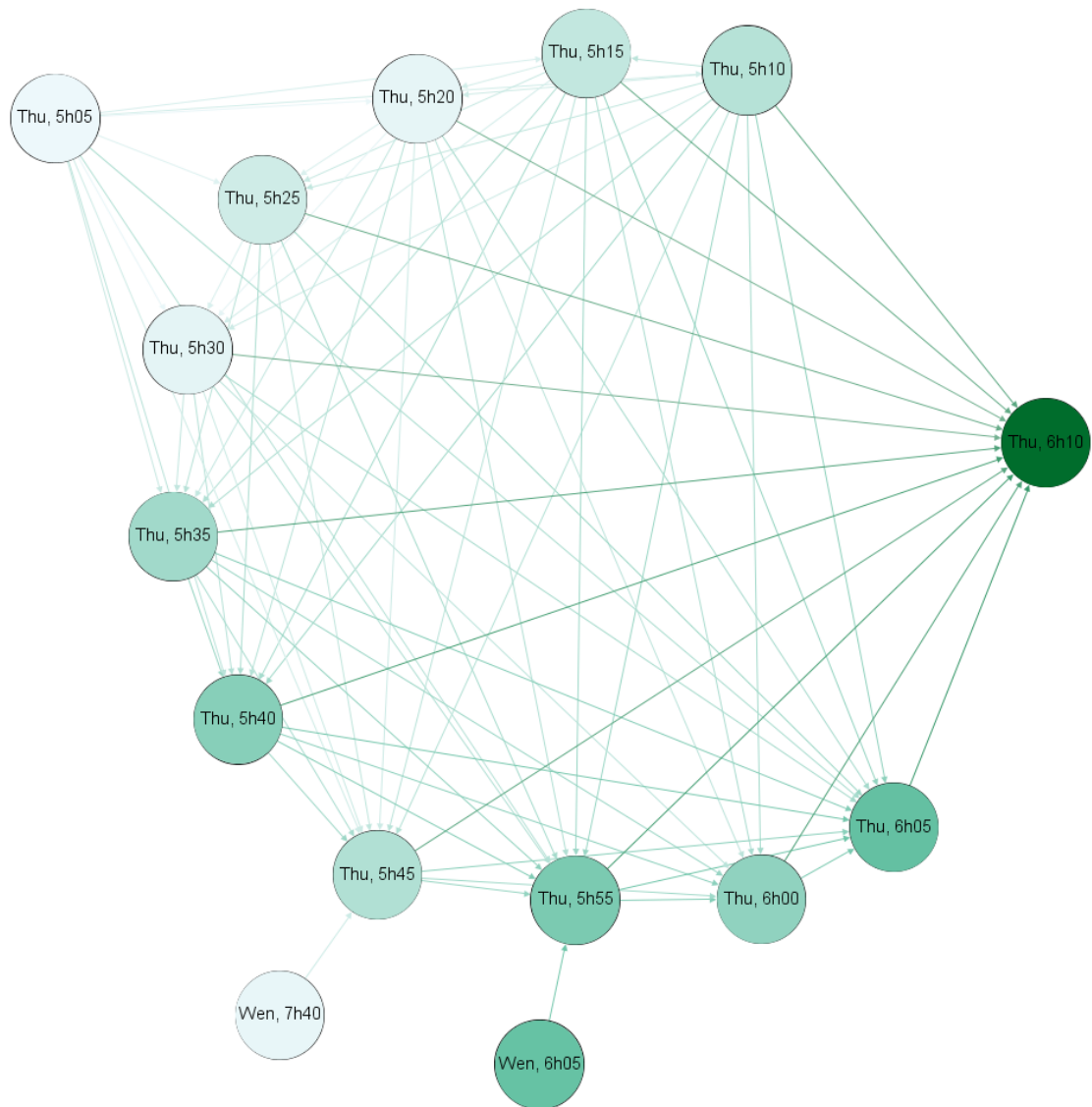


Figura 41 – Maiores PageRanks que chegam na *Quinta – feira, 6 h 10 min.*

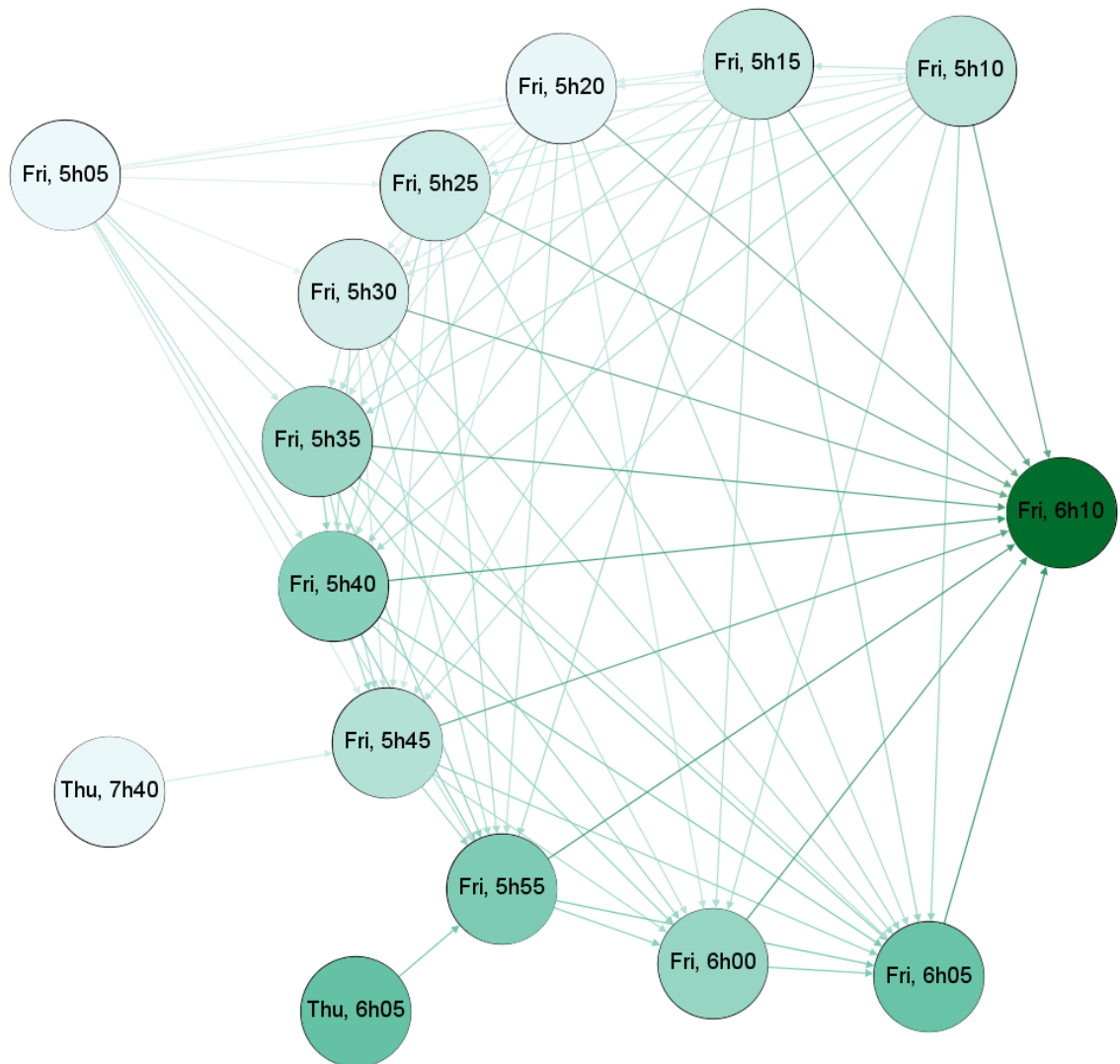


Figura 42 – Maiores PageRanks que chegam na *Sexta – feira, 6 h 10 min.*

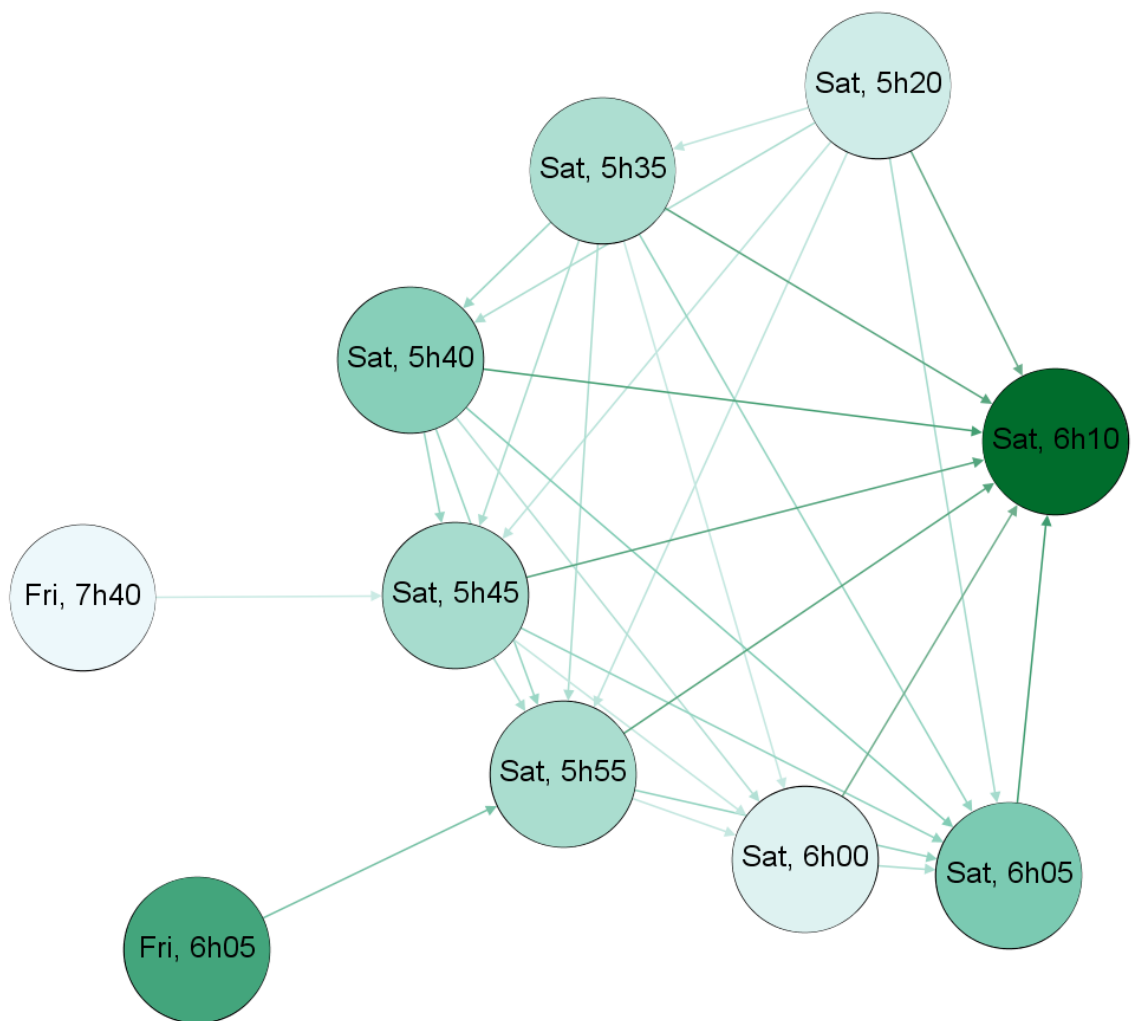


Figura 43 – Maiores PageRanks que chegam no *Sabado, 6 h 10 min.*