

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL
INSTITUTO DE INFORMÁTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM COMPUTAÇÃO

RICARDO ANDRADE CAVA

**Abordagens heterogêneas para a
exploração interativa de grafos
multivariados**

Tese apresentada como requisito parcial para a
obtenção do grau de Doutor em Ciência da
Computação

Profa. Dra. Carla Maria Dal Sasso Freitas
Orientadora

Prof. Dr. Marco Winckler
Coorientador

Porto Alegre
2017

CIP – CATALOGAÇÃO NA PUBLICAÇÃO

Cava, Ricardo Andrade

Abordagens heterogêneas para exploração interativa de grafos multivariados / Ricardo Andrade Cava. – 2017.

145 f.:il.

Orientador: Profa. Dra. Carla Maria Dal Sasso Freitas ; Co-orientador: Prof. Dr. Marco Winckler .

Tese (doutorado) – Universidade Federal do Rio Grande do Sul. Programa de Pós-Graduação em Computação. Porto Alegre, BR – RS, 2017.

1.Visualização de informações. 2.Grafos Multivariados
3.Visualização de agrupamentos. I. Freitas, Carla Maria Dal Sasso, II. Wincker,Marco III. Título.

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL

Reitor: Prof. Rui Vicente Oppermann

Vice-Reitora: Profa. Jane Fraga Tutikian

Pró-Reitor de Pós-Graduação: Prof. Celso Loureiro Chaves

Diretor do Instituto de Informática: Profa. Carla Maria Dal Sasso Freitas

Coordenador do PPGC: Prof. João Luiz Dihl Comba

Bibliotecária-Chefe do Instituto de Informática: Beatriz Regina Bastos Haro

AGRADECIMENTOS

À minha família pelo apoio e compreensão nos momentos mais difíceis.

À minha orientadora, professora Carla Maria Dal Sasso Freitas, pela confiança em mim depositada e cujo apoio e sugestões foram fundamentais na realização desse trabalho.

Ao meu co-orientador, professor Marco Winckler que, além de me receber no estágio de doutoramento no *Institut de Recherche en Informatique de Toulouse*, contribuiu com inúmeras ideias que propiciaram uma evolução significativa desse trabalho.

Ao Instituto Federal de Educação Ciência e Tecnologia Sul-rio-grandense por ter me concedido a licença capacitação para a realização do doutorado.

À Capes pelo apoio financeiro na realização do estágio em Toulouse.

RESUMO

Muitas aplicações tratam dados estruturados na forma de grafos, como, por exemplo, redes sociais, redes de computação e comunicação, redes epidemiológicas, entre outras. Essas aplicações são baseadas em grafos multivariados representando itens e relacionamentos caracterizados por múltiplos atributos. A maioria das técnicas descritas na literatura para lidar com grafos multivariados concentram-se em problemas associados com visualização da topologia ou em problemas associados com a visualização de múltiplos atributos de itens separados da topologia do grafo. Durante a exploração de grafos multivariados, os usuários podem se beneficiar da combinação de diversas técnicas de visualização. A fim de apoiar os usuários durante essa exploração, esta tese propõe uma abordagem que permite ao usuário combinar diversas técnicas de visualização, mantendo o controle da história das visualizações encadeando-as de uma maneira integrada. Os usuários são capazes de comparar os resultados fornecidos por diferentes técnicas de visualização, o que proporciona o sinergismo necessário para a compreensão mais completa do conjunto de dados. São propostas também três técnicas para a exploração de grafos multivariados. A primeira técnica (*ClusterVis*) fornece a visualização das relações entre atributos de nodos pertencentes a agrupamentos. A segunda, denominada *GlyphMatrix*, explora o uso de glifos e matriz de adjacência, para visualizar a relação entre atributos associados às arestas. E, finalmente, a terceira (*Iris*) permite a visualização de atributos associados às arestas de nodos adjacentes.

Palavras-chave: Visualização de informações, Grafos multivariados, Visualização de agrupamentos.

Heterogeneous approaches for interactive exploration of multivariate graphs

ABSTRACT

Many computing applications imply dealing with network data, for example, social networks, communications and computing networks, epidemiological networks, among others. These applications are based on multivariate graphs representing items and relationships characterized by multiple attributes. Most of the visualization techniques described in the literature for dealing with multivariate graphs focus either on problems associated with the visualization of topology or on problems associated with the visualization of multiple attributes of items, separated from the graph topology. During the exploration of multivariate graphs, users might get benefit of combining these diverse visualization techniques. In order to support users during that exploration, this thesis proposes an approach that allows users to combine diverse visualization techniques while keeping track of the history of chained visualizations in an integrated way. Users are able to compare results provided by different visualization techniques, and thus the tools provide the synergism one needs to fully comprehend the data set. Three techniques were embedded in the approach. The first one emphasizes the visualization of relations between the attributes of nodes belonging to clusters, and thus is called ClusterVis. The second one is named GlyphMatrix, and explores the use of glyphs and adjacency matrices as an alternative representation of the relation between the attributes of edges. Finally, a third technique (Iris) provides features for the visualization of attributes of edges of adjacent nodes.

Keywords: Information visualization, Multivariate graphs, Cluster visualization

LISTA DE FIGURAS

Figura 1.1 – Componentes de um grafo multivariado	14
Figura 2.1 – <i>RelVis</i>	22
Figura 2.2 – <i>PivotGraph</i>	23
Figura 2.3 – <i>NVSS: Network Visualization by Semantic Substrates</i>	24
Figura 2.4 – <i>NoodleView</i>	25
Figura 2.5 – <i>GraphScape</i>	26
Figura 2.6 – <i>AttriGraph</i>	26
Figura 2.7 – <i>Paired Parallel Coordinates</i>	27
Figura 2.8 – <i>GraphDice</i>	28
Figura 2.9 – <i>PEx-Graph</i>	29
Figura 2.10 – <i>P-SPLOM: Parallel Scatterplot Matrix</i>	29
Figura 2.11 – <i>Network Lens</i>	30
Figura 2.12 – <i>CluMa-GO</i>	31
Figura 2.13 – <i>GraphTrail</i>	31
Figura 2.14 – <i>ViNCent: Visualization of Network Centralities</i>	32
Figura 2.15 – <i>enRoute</i>	33
Figura 2.16 – <i>SocialNetSense</i>	33
Figura 2.17 – <i>JauntyNets</i>	34
Figura 2.18 – <i>PNLB: Parallel Node-Link Bands</i>	35
Figura 2.19 – <i>PIWI</i>	35
Figura 2.20 – Visualização de <i>cluster</i> de documentos	36
Figura 2.21 – <i>OnionGraph</i>	37
Figura 2.22 – Sistema proposto por Elzen e Wijk (2014)	37
Figura 2.23 – (a) <i>Petal</i> (b) <i>Threads</i>	38
Figura 2.24 – (a) Grafo original (b) <i>Heatmaps</i> (c) <i>Edges Bundled</i> (d) Com nodos agregados	39
Figura 2.25 – (a) <i>Hierarchy Explorer</i> (b) <i>Group Explorer</i> (b) <i>Candidate Explorer</i> (d) Glifos	39
Figura 2.26 – (a) Comunidades (b) Nodos de fronteiras	40
Figura 2.27 – (a) <i>Data Overview</i> (b) <i>Timeline View</i> (c) <i>Alter Timeline View</i>	41
Figura 2.28 – Área de aplicação e existência de atributos nas arestas	42
Figura 2.29 – Classificação das técnicas por abordagem	44
Figura 2.30 – Representação visual dos atributos e da topologia	47
Figura 2.31 – Tarefas para exploração de grafos multivariados	49
Figura 2.32 – Técnicas de interação - Visões	52
Figura 2.33 – Técnicas de interação – Estruturas visuais	53
Figura 2.34 – Técnicas de interação – Dados	53
Figura 2.35 – Desenvolvimento das técnicas no período de 2005 a 2016	54
Figura 2.36 – Quantidade de citações entre as técnicas selecionadas	54
Figura 3.1 – Modelo para <i>design</i> e validação de ferramentas de visualização	57
Figura 3.2 – Múltiplas instâncias de visualização	58
Figura 3.3 – Instâncias de visualização interligadas	59
Figura 3.4 – Tabelas de autores, artigos e relação de coautoria	61
Figura 3.5 – Grafos: (a) Coautoria (b) Eventos em comum (c) coautoria no <i>VAST</i>	61
Figura 3.6 – Quantidade de ferramentas x quantidade de requisitos atendidos	64
Figura 3.7 – Quantidade de requisitos atendidos em cada ferramenta	64
Figura 3.8 – Premissas atendidas em cada ferramenta	65
Figura 3.9 – Quantidade de ferramentas que atendem os requisitos	66
Figura 3.10 – Quantidade de ferramentas x quantidade de técnicas distintas	66

Figura 3.11 – Quantidade de técnicas distintas utilizadas pelas ferramentas.....	67
Figura 3.12– Quantidade de técnicas distintas utilizadas pelas ferramentas.....	67
Figura 3.13 – (a) Seleção de 1 nível (b) Seleção de 2 níveis.....	68
Figura 3.14 – (a) Caminho mais curto (b) Todos os caminhos simples.....	69
Figura 3.15 –Utilização de área de seleção.....	69
Figura 3.16 –Seleção recursiva de um caminho.....	70
Figura 3.17 –(a) <i>Clusters</i> gerados automaticamente (b) Supernodos e superarestas.....	70
Figura 3.18 – Processo de exploração.....	72
Figura 4.1 – <i>ClusterVis</i> – Atributos de nodos em um <i>cluster</i>	74
Figura 4.2 – (a) Árvore geradora (b) Árvore geradora e <i>B-splines</i>	74
Figura 4.3 – Ordenação dos valores de atributos.....	75
Figura 4.4 – Sobreposição de arestas.....	76
Figura 4.5 – <i>GlyphMatrix</i> – atributos de arestas.....	77
Figura 4.6– Glifos disponíveis na <i>GlyphMatrix</i>	77
Figura 4.7– <i>Tooltips</i> que exibem glifos.....	78
Figura 4.8– Matriz 106 x 106 com 15 linhas visíveis.....	79
Figura 4.9– <i>Iris</i> – Um atributo das arestas que conectam um nodo de interesse.....	80
Figura 4.10– <i>Iris</i> com 55 nodos visíveis de um total de 67.....	81
Figura 5.1– <i>MGExplorer</i> (a) Ligações entre janelas (b) Janelas ocultas (c) Histórico.....	84
Figura 5.2– E-R da base de dados de coautoria INF-UFRGS.....	85
Figura 5.3– Representação inicial do grafo com dados de 2011.....	86
Figura 5.4– Grafo clusterizado por linha de pesquisa.....	86
Figura 5.5– Ativação da <i>GlyphMatrix</i>	87
Figura 5.6– Coautores do pesquisador Pimenta.....	87
Figura 5.7– Coautorias na linha de pesquisa Engenharia de Software.....	88
Figura 5.8– Coautorias em Computação Gráfica no ano de 2011.....	88
Figura 5.9 – Subgrafo visualizado com a <i>ClusterVis</i>	89
Figura 5.10 – <i>GlyphMatrix</i> : – Mostra todas as conexões entre os coautores.....	89
Figura 5.11 – Oliveira e seus coautores.....	90
Figura 5.12 – Coautorias dos pesquisadores Comba e Maciel.....	90
Figura 5.13 – Coautorias na linha de pesquisa Computação Gráfica.....	91
Figura 6.1 – Botões <i>Tools</i> e <i>Hide</i>	93
Figura 6.2 – Painel de ferramentas da <i>ClusterVis</i>	94
Figura 6.3 – Estado inicial da <i>ClusterVis</i> para a tarefa 1.....	95
Figura 6.4 – Estado inicial da <i>ClusterVis</i> para a tarefa 2.....	97
Figura 6.5 – Estado inicial da <i>ClusterVis</i> para a tarefa 3.....	98
Figura 6.6 – Aplicação usada na avaliação da <i>ClusterVis</i>	100
Figura 6.7 – Painel de ferramentas da <i>GlyphMatrix</i>	102
Figura 6.8 – Estado inicial da <i>GlyphMatrix</i> para as tarefas 1 e 2.....	103
Figura 6.9 – Visualização dos resultados de partidas de futebol com a <i>GlyphMatrix</i>	105
Figura 6.10 – Painel de ferramentas da <i>Iris</i>	108
Figura 6.11 – Estado inicial da <i>Iris</i> para a tarefa 1.....	109
Figura 6.12 – Estado inicial da <i>Iris</i> para a tarefa 2.....	110
Figura 6.13 – Estado inicial da <i>Iris</i> para a tarefa 3.....	111
Figura 6.14 – <i>Iris</i> inserida em uma página <i>Web</i>	113
Figura 6.15 – Painel de ferramentas da técnica que exhibe o diagrama de nodos e arestas....	116
Figura 6.16 – Barra de seleção.....	117
Figura 6.17 – Menu de contexto para ativações de outras técnicas de visualização.....	117
Figura 6.18 – Representação de uma janela oculta.....	117
Figura 6.19 –Estado inicial da <i>MGExplorer</i> para a tarefa 1.....	118

Figura 6.20 –Estado inicial da <i>MGExplorer</i> para a tarefa 2.....	120
Figura 6.21 –Estado inicial da <i>MGExplorer</i> para a tarefa 2.....	121

LISTA DE TABELAS

Tabela 2.1 – Mecanismos de busca utilizados.....	20
Tabela 2.2 – Técnicas classificadas	21
Tabela 2.3 – Resultados da análise das questões QP1 e QP2.....	42
Tabela 2.4 – Técnicas classificadas pela abordagem	44
Tabela 2.5 – Representação visual dos atributos e da topologia	46
Tabela 2.6 – Tarefas para exploração de grafos multivariados	49
Tabela 2.7 - Técnicas de interação para exploração de grafos multivariados	52
Tabela 3.1 – Requisitos atendidos pelas ferramentas	63
Tabela 6.1 – Detalhamento das ações para execução da tarefa 1 na <i>ClusterVis</i>	99
Tabela 6.2 – Detalhamento das ações para execução da tarefa 2 na <i>ClusterVis</i>	99
Tabela 6.3 – Detalhamento das ações para execução da tarefa 3 na <i>ClusterVis</i>	100
Tabela 6.4 – Análise das tarefas	101
Tabela 6.5 – Detalhamento das ações para execução da tarefa 1 na <i>GlyphMatrix</i>	105
Tabela 6.6 – Detalhamento das ações para execução da tarefa 2 na <i>GlyphMatrix</i>	105
Tabela 6.7 – Análise do resultado geral dos jogos	107
Tabela 6.8 – Análise das estatísticas dos jogos	107
Tabela 6.9 – Detalhamento das ações para execução da tarefa 1 na <i>Iris</i>	112
Tabela 6.10 – Detalhamento das ações para execução da tarefa 2 na <i>Iris</i>	112
Tabela 6.11 – Detalhamento das ações para execução da tarefa 3 na <i>Iris</i>	113
Tabela 6.12 –Técnica de visualização para exibir resultados de uma consulta.....	114
Tabela 6.13 – Uso da <i>Iris</i> para a visualização das informações de coautoria	115
Tabela 6.14 – Detalhamento das ações para execução da tarefa 1 na <i>MGExplorer</i>	123
Tabela 6.15 – Detalhamento das ações para execução da tarefa 2 na <i>MGExplorer</i>	124
Tabela 6.16 – Detalhamento das ações para execução da tarefa 3 na <i>MGExplorer</i>	125

LISTA DE ABREVIATURAS E SIGLAS

<i>CW</i>	<i>Cognitive Walkthrough</i>
<i>DOSA</i>	<i>Detail to Overview via Selections an Aggregations</i>
<i>GOMS</i>	<i>Goal, Operators, Methods, Selection rules</i>
<i>MGExplorer</i>	<i>Multivariate Graph Explorer</i>
<i>NES</i>	<i>Network Exploring Space</i>
<i>NVSS</i>	<i>Network Visualization by Semantic Substrates</i>
<i>P-SPLOM</i>	<i>Parallel Scatterplot Matrix</i>
<i>PNLB</i>	<i>Parallel Node-Link Bands</i>
<i>RBS</i>	<i>Representation Building Space</i>
<i>SUS</i>	<i>System Usability Scale</i>
<i>UFRGS</i>	<i>Universidade Federal do Rio Grande do Sul</i>
<i>ViNCent</i>	<i>Visualization of Network Centralities</i>

SUMÁRIO

1	INTRODUÇÃO	13
1.1	Visualização de grafos multivariados	13
1.2	Motivação	15
1.3	Objetivos e contribuições esperadas	17
1.3.1	Objetivo geral	17
1.3.2	Objetivos específicos	18
1.3.3	Contribuições.....	18
1.4	Organização do texto	18
2	ESTADO DA ARTE	19
2.1	Método utilizado na revisão	Erro! Indicador não definido.
2.1.1	Questões de pesquisa	19
2.1.2	Processo de busca	20
2.1.3	Critérios de exclusão	20
2.1.4	Realização da análise.....	21
2.2	Descrição das técnicas selecionadas para análise	22
2.3	Resultados da análise	41
2.3.1	Área de aplicação e atributos de arestas	41
2.3.2	Abordagens de integração da topologia e seus atributos	42
2.3.3	Representação visual	45
2.3.4	Tarefas para a exploração de grafos multivariados	47
2.3.5	Técnicas de interação para exploração de grafos multivariados	49
2.4	Comentários finais	53
3	EXPLORAÇÃO INTERATIVA DE GRAFOS MULTIVARIADOS	56
3.1	Premissas para a exploração interativa de grafos multivariados	58
3.2	Requisitos para aderência à abordagem	62
3.3	Caminhos possíveis para a exploração de grafos multivariados	68
3.4	Conexões entre visões e histórico de exploração de grafos multivariados	71
4	TÉCNICAS DESENVOLVIDAS	73
4.1	ClusterVis	73
4.1.1	Representação visual	73
4.1.2	Interação	75
4.1.3	Limitações	76
4.2	GlyphMatrix	76
4.2.1	Representação visual	77
4.2.2	Interação	78
4.2.3	Limitações	79
4.3	Iris	79
4.3.1	Representação visual	80
4.3.2	Interação	81
4.3.3	Limitações	82
5	MULTIVARIATE GRAPH EXPLORER	83

5.1	Descrição da <i>MGExplorer</i>	83
5.2	Estudo de caso	84
5.2.1	Dados visualizados	84
5.2.2	Exemplos de exploração	85
6	AVALIAÇÃO	92
6.1	Avaliação da <i>ClusterVis</i>	94
6.1.1	Utilidade	95
6.1.2	Performance	99
6.1.3	Satisfação	100
6.1.3.1	<i>Participantes</i>	100
6.1.3.2	<i>Procedimento</i>	101
6.1.3.3	<i>Resultados</i>	101
6.2	Avaliação da <i>GlyphMatrix</i>	102
6.2.1	Utilidade	103
6.2.2	Performance	104
6.2.3	Satisfação	105
6.2.3.1	<i>Participantes</i>	106
6.2.3.2	<i>Procedimento</i>	106
6.2.3.3	<i>Resultados</i>	106
6.3	Avaliação da <i>Iris</i>	108
6.3.1	Utilidade	108
6.3.2	Performance	112
6.3.3	Satisfação	113
6.3.3.1	<i>Participantes</i>	114
6.4	Avaliação da <i>MGExplorer</i>	115
6.4.1	Utilidade	115
6.4.1.1	<i>Mecanismos de interação com o diagrama de nodos e arestas:</i>	116
6.4.1.2	<i>Mecanismos de interação com o histórico:</i>	116
6.4.1.3	<i>Mecanismos de interação de uso geral:</i>	116
6.4.1.4	<i>Avaliação</i>	117
6.4.2	Performance	123
6.5	Comentários finais	125
7	CONCLUSÃO	128
7.1	Publicações	129
7.2	Trabalhos futuros	130
	REFERÊNCIAS	131
	APÊNDICE A – DOCUMENTOS USADOS NAS AVALIAÇÕES	138

1 INTRODUÇÃO

Há vários anos que estimativas apontam para uma produção mundial de dados da ordem de cinco *exabytes*/ano provenientes de várias fontes (LYMAN; VARIAN, 2003) com uma previsão de crescimento de 30% ao ano dessa produção. Apesar dessa enorme disponibilidade de informação parecer positiva, o excesso pode resultar em um efeito conhecido como “sobrecarga de informação” que ocorre quando o volume de informações é tão grande que um indivíduo não é capaz de localizar e fazer uso do que ele necessita. Uma abordagem para contornar as dificuldades de selecionar informações relevantes dentre os resultados de buscas em sistemas de recuperação de informações e de grandes bases de dados é utilizar técnicas de visualização de informações (GERSHON; EICK, 1997, MUNZNER, 2014) por meio das quais o usuário obtém uma representação visual que, se por um lado abstrai detalhes do conjunto de informações, por outro propicia uma organização desse conjunto segundo algum critério. Assim, representações gráficas tradicionalmente empregadas apenas para divulgação e apresentação de dados e resultados passaram a ser usadas como ferramentas de exploração, no que Tukey (1977) pioneiramente denominou análise exploratória visual.

Vários campos do conhecimento tratam com dados que são representados computacionalmente como grafos. De fato, qualquer domínio que pode ser modelado como uma coleção de elementos com relacionamentos entre si pode ser representado por um grafo. Muito frequentemente, esses elementos são caracterizados por atributos diversos, o que permite que essas coleções sejam também passíveis de representação como um conjunto de dados multivariados ou multidimensionais. Dessa combinação, resulta o que se convencionou chamar de grafos multivariados, cujo tratamento através de técnicas de visualização é objeto do presente trabalho.

1.1 Visualização de grafos multivariados

Um grafo $G=(V,E)$ pode ser definido por dois conjuntos finitos V e E , cujos elementos são denominados respectivamente de vértices (ou nodos) e arestas (ou arcos). Cada aresta é formada por um par de vértices v_1 e v_2 pertencentes ao conjunto V . Esse par pode ser ordenado definindo um sentido para cada aresta (grafo dirigido) ou não ordenado (grafo não dirigido). Dois vértices v_1 e v_2 são considerados adjacentes se o par (v_1,v_2) pertencer ao conjunto E .

Um grafo multivariado consiste de um grafo G no qual cada nodo ou aresta pode ter associado a si um conjunto de atributos (KERREN et al., 2014). Assim, um grafo

multivariado pode ser tratado como (i) um conjunto de dados multidimensionais, composto de itens que são descritos por vários atributos (usualmente mais que três) e (ii) como um grafo, onde as conexões entre nodos (a topologia) desempenham o papel central (Figura 1.1).

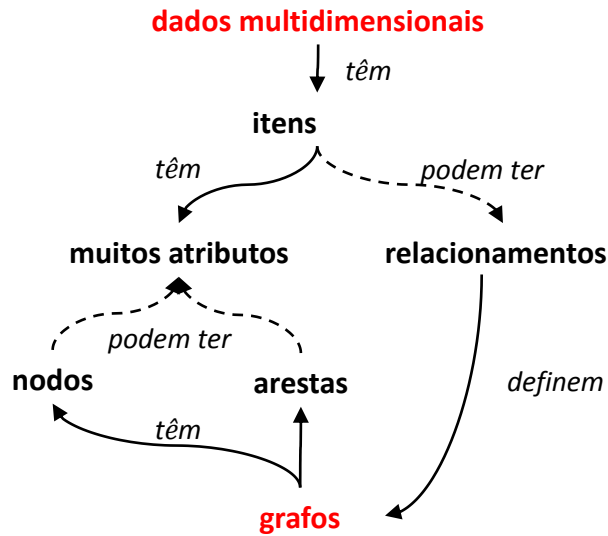


Figura 1.1 – Componentes de um grafo multivariado

Por exemplo, em uma rede social os elementos que representam pessoas podem conter atributos como nome, sexo, data de nascimento, e os relacionamentos, informações como tipo e tempo de existência dos mesmos entre outros. Redes de coautoria podem conter diversas informações sobre cada autor e os seus relacionamentos, diversos atributos associados às publicações em comum. Na área da biologia molecular pode-se representar uma rede de interação entre proteínas, armazenando atributos como nome, função e reações entre elas.

A visualização de grafos refere-se a um conjunto de técnicas desenvolvidas com o objetivo de gerar uma representação visual integrada a mecanismos de interação, de forma a facilitar a percepção das relações contidas na estrutura. Tais técnicas apresentam duas características importantes para a visualização de informações: i) representação de relações entre itens de informações e ii) possibilidade de uma visão topológica da rede de relações representada pelo grafo. Entretanto, Card et al. (1999b) relaciona alguns problemas importantes na visualização de grafos, como, por exemplo, a definição do posicionamento de nodos. Com esse propósito diversos algoritmos de *layout* (BATTISTA et al., 1999, TAMASSIA, 2013) têm sido propostos, levando em conta diferentes critérios estéticos, que, muitas vezes, funcionam apenas para um determinado tipo de grafo.

Já o desafio da visualização de dados multidimensionais é mapear as múltiplas dimensões em uma única estrutura visual 1D, 2D ou 3D de forma a obter uma apresentação clara e compreensível, permitindo assim uma satisfatória análise visual dos dados. Técnicas de visualização de dados multidimensionais permitem mostrar vários atributos da informação ao mesmo tempo contudo, o número de atributos que podem ser visualizados é limitado pelo conjunto de marcas e canais visuais utilizados pela técnica de visualização. Cada aumento no número de elementos visuais usados para representar atributos tem como impacto direto uma sobrecarga visual da representação. Embora as técnicas de visualização desenvolvidas possam ser empregadas com sucesso para pequenos conjuntos de dados, tendem a apresentar problemas para conjuntos maiores com alta dimensionalidade, como, por exemplo, a dificuldade em identificar satisfatoriamente grupos de dados correlacionados. A visualização de dados multidimensionais levanta um dilema entre os benefícios de mostrar o máximo possível de uma só vez, para minimizar a necessidade de navegação e exploração, e os custos de mostrar muito de uma só vez, onde o usuário é sobrecarregado pela desordem visual (MUNZNER, 2014).

Ao longo dos anos, pesquisadores desenvolveram várias técnicas para a visualização de grafos (HERMAN et al., 2000, LANDESBERGER et al., 2011, BECK, et al., 2014) e conjunto de dados multidimensionais (OLIVEIRA; LEVKOWITZ, 2003; WONG; BERGERON, 1994). No entanto, aplicações envolvendo grafos multivariados tornaram-se bastante comum ultimamente demandando a exploração de dados multidimensionais em conjunto com a sua topologia. Pode-se citar como exemplos disso aplicações nas áreas da engenharia de software (DIEHL et al., 2014), redes sociais (MUELDER et al., 2014) e ciências da vida (KOHLBACHER et al., 2014).

1.2 Motivação

A visualização de grafos multivariados é um tema desafiador para a pesquisa em visualização de informações devido à riqueza, variedade e quantidade de informações que podem ser representadas por eles. Nesse tipo de estrutura, nodos e arestas não são mais entidades atômicas, já que são repositórios de dados que possuem uma quantidade qualquer de atributos. Encontrar representações visuais que possam ajudar os usuários a compreender tal estrutura é um claro desafio. Levando-se em conta os limites de resolução de tela e os limites da percepção humana, muitas vezes é impossível mostrar claramente toda a informação de uma forma útil.

Para obter uma visão global das soluções propostas pela comunidade científica para a visualização de grafos multivariados, foi realizada uma análise comparativa das características de diversas técnicas desenvolvidas nos últimos dez anos (Capítulo 2). Elas foram analisadas considerando cinco grupos de critérios. O primeiro grupo enfoca algumas características gerais de cada técnica, o segundo utiliza a classificação adotada por Jusufi et al. (2013), que agrupa as técnicas de visualização segundo a forma como as representações são construídas, levando em conta a integração da estrutura topológica com a visualização dos atributos. O terceiro refere-se à representação visual utilizada, enquanto o quarto engloba as tarefas suportadas para a análise de grafos multivariados. Finalmente, o quinto grupo de critérios refere-se aos mecanismos de interação disponibilizados pelas técnicas. Após a realização dessa análise foram identificados alguns aspectos que poderiam ser explorados.

As técnicas, em sua grande maioria, dão uma maior ênfase à visualização de atributos de nodos. Embora boa parte delas também disponibilize algum tipo de visualização de atributos de arestas, muitas vezes, a representação está limitada ao uso da cor e espessura da linha que as representam. Essa estratégia dificulta a exploração dos múltiplos atributos e a investigação da relação entre eles.

Outro ponto a ser considerado é o fato de que grande parte das técnicas não suporta tarefas de exploração baseadas na estrutura do grafo. Tarefas como as relacionadas com conectividade e acessibilidade, que permitem obter informações sobre nodos que podem ser acessados a partir de outro, ainda precisam ser investigadas.

Quando o grafo possui uma grande quantidade de nodos ou torna-se mais denso devido à existência de muitas ligações entre eles, o diagrama de nodos e arestas torna-se pouco legível com nodos sobrepostos e muitos cruzamentos de arestas. Essa sobrecarga visual é agravada se adicionarmos aos grafos a representação dos atributos de nodos e arestas. Entretanto, filtragem e agregação são estratégias comuns que podem ser usadas para amenizar esse problema, gerando uma representação simplificada com tamanho e densidade reduzidos. Na filtragem, nodos ou arestas são removidos do grafo segundo algum critério definido pelo usuário, reduzindo o seu tamanho e densidade, melhorando a legibilidade do diagrama. A agregação tem por objetivo reunir um conjunto de nodos em um super-nodo. Nessa operação, as arestas também são agrupadas, reduzindo o tamanho e a densidade do grafo. Embora essa estratégia também contribua para a legibilidade da representação, pode reduzir a percepção geral da topologia. Portanto, encontrar uma forma de mostrar a relação entre os múltiplos atributos de nodos e arestas sem que ocorra uma sobrecarga visual, explorando a visualização de nodos pertencentes a agrupamentos, é outro aspecto a ser investigado.

Considerando o grande número de possíveis técnicas de visualização tanto para a visualização de grafos como a visualização de dados multidimensionais, é impossível encontrar uma única combinação de técnicas que apoie plenamente a exploração de tais conjuntos de dados. Assim, a solução é fornecer ao usuário diversas técnicas para serem usadas de uma forma sinérgica na visualização e exploração combinadas de grafos multivariados. Ao fazê-lo, é colocado em prática o framework *what-why-how* (MUNZNER, 2014) em um sentido mais amplo porque há também a necessidade de manter o controle de quais ações de exploração o usuário executa em busca da compreensão dos dados e quais técnicas de visualização foram empregadas na exploração. Dessa forma, este trabalho propõe uma abordagem para a exploração de grafos multivariados que não só integra diferentes técnicas de visualização, mas também representa as conexões entre as visualizações criadas ao longo do caminho de exploração. Essa abordagem foi concebida com base no estudo de várias técnicas descritas na literatura para a exploração de grafos multivariados.

1.3 Objetivos e contribuições esperadas

Durante a investigação mencionada na seção anterior, foram identificadas três premissas para o desenvolvimento de técnicas de visualização para grafos multivariados:

- 1) Durante o processo de exploração de um grafo multivariado pode ser necessário que os usuários façam uso de diferentes técnicas de visualização;
- 2) Os itens de informação armazenados em grafos multivariados não podem ser facilmente exibidos todos de uma única vez; portanto, um processo exploratório deve permitir que os usuários inspecionem e comparem subconjuntos de dados por meio de visualizações interligadas;
- 3) A exploração de grafos multivariados não exige apenas a exploração de nodos, mas também pode exigir a inspeção de arestas, subgrafos e relacionamentos criados devido ao valor dos atributos dos nodos e arestas.

1.3.1 Objetivo geral

Diante dessas premissas, estabeleceu-se como objetivo geral deste trabalho “propor, desenvolver e avaliar uma abordagem para apoiar a exploração de grafos multivariados que integre diferentes técnicas de visualização possibilitando a exibição simultânea, o registro e a representação da sequência das visualizações utilizadas na investigação de subconjuntos de dados de interesse obtidos a partir do grafo a ser explorado”.

1.3.2 Objetivos específicos

Foram definidos os seguintes objetivos específicos para este trabalho:

- 1) Propor e desenvolver uma técnica de representação da sequência de visualizações usadas na exploração para grafos multivariados.
- 2) Propor e desenvolver duas técnicas de visualização para a exploração de atributos de arestas em grafos multivariados.
- 3) Propor e desenvolver uma técnica para a visualização das relações entre atributos de nodos pertencentes a agrupamentos em grafos multivariados.

1.3.3 Contribuições

A seguir estão relacionadas as contribuições deste trabalho:

- 1) Definição de uma nova abordagem a ser adotada para a exploração de grafos multivariados.
- 2) Desenvolvimento de uma técnica para a visualização das relações entre atributos de nodos pertencentes a agrupamentos em grafos multivariados (*ClusterVis*).
- 3) Desenvolvimento de uma técnica para a exploração de atributos de arestas pertencentes a grafos multivariados (*GlyphMatrix*).
- 4) Desenvolvimento de uma técnica para a visualização de atributos de arestas associadas a um determinado nodo de interesse (*Iris*).
- 5) Uma análise comparativa das técnicas para visualização de grafos multivariados desenvolvidas nos últimos anos.

As técnicas e a abordagem proposta foram implementadas em uma ferramenta como prova de conceito, denominada *Multivariate Graph Explorer*, detalhada no Capítulo 5.

1.4 Organização do texto

Após este capítulo introdutório, o texto está estruturado da seguinte maneira. O Capítulo 2 apresenta uma análise comparativa das características de diversas técnicas desenvolvidas nos últimos anos utilizadas na exploração de grafos multivariados. O Capítulo 3 descreve de forma conceitual a abordagem proposta para a exploração de grafos multivariados. O Capítulo 4 descreve a proposta de três novas técnicas que foram desenvolvidas durante a realização deste trabalho. O Capítulo 5 apresenta o protótipo utilizado como prova de conceito e um estudo de caso com a exploração de redes de coautoria dos pesquisadores do Instituto de Informática da Universidade Federal do Rio Grande do Sul. O Capítulo 6 relata o processo de avaliação executado para validar as técnicas propostas. E finalmente, o Capítulo 7 contém as principais conclusões e trabalhos futuros.

2 ESTADO DA ARTE

Um grafo multivariado é um tipo de estrutura que apresenta desafios à comunidade de pesquisadores da área de visualização de informações. Ele proporciona a representação de estruturas complexas de informações constituídas de elementos e seus relacionamentos que, por sua vez, podem estar associados a múltiplos atributos, o que representa uma complicação adicional ao desenvolvimento de técnicas para sua visualização. Este capítulo tem por objetivo apresentar uma análise comparativa das características de diversas técnicas (ou sistemas) desenvolvidas nos últimos anos para a exploração visual de grafos multivariados. Essa análise foi concluída no início do ano de 2016.

2.1 Método utilizado na revisão

Para a realização desta análise foi adotado um protocolo de revisão *ad hoc* (KITCHENHAM; CHARTERS, 2007) cujas etapas que o constituem são descritas a seguir.

2.1.1 Questões de pesquisa

A primeira etapa do protocolo é constituída pelas seguintes questões de pesquisa que deverão ser respondidas com esta revisão:

QP1 As técnicas foram projetadas para uma aplicação específica ou podem ser usadas na exploração de grafos multivariados genéricos?

QP2 Quais técnicas proporcionam a representação de múltiplos atributos de arestas?

QP3 De que forma as técnicas integram a representação da topologia do grafo e seus respectivos atributos?

QP4 Quais são as primitivas gráficas empregadas para representar os múltiplos atributos?

QP5 De que forma a topologia do grafo é representada?

QP6 Quais tarefas de exploração são suportadas pelas técnicas?

QP7 Quais mecanismos de interação são disponibilizados pelas técnicas?

O objetivo da questão QP1 é identificar quais técnicas são independentes da aplicação para as quais foram projetadas, podendo ser utilizadas de forma genérica. Em relação à representação de múltiplos atributos de arestas (QP2), as técnicas foram divididas em duas categorias: as que representam até três atributos e as que representam mais de três atributos. Para responder a questão QP3, foi usada a classificação proposta por Jusufi et. (2013), que distribui as técnicas em cinco abordagens considerando a forma utilizada para integrar a representação da topologia e seus atributos: Múltiplas visões coordenadas, abordagem

integrada, substrato semântico, *layout* dirigido pelos atributos e abordagem híbrida. Na questão QP4 foi empregado o critério apresentado por Ward. et al (2010), que classifica as técnicas segundo os tipos de primitivas gráficas usadas: baseadas em pontos, baseadas em linhas, baseadas em regiões e constituídas de combinações dessas. Em relação à investigação do tipo de representação adotada para exibir a topologia do grafo (QP5), as técnicas foram classificadas em três categorias: as que utilizam o diagrama de nodos e arestas, as que usam a matriz de adjacências e as que empregam outras formas de representação. Para investigar quais tarefas de exploração são admitidas pelas técnicas (ou ferramentas) (QP6), foi adotada uma taxonomia baseada no trabalho de Pretorius et al; (2014). Por fim, para responder a questão QP7 foi utilizada a abordagem proposta por (WYBROW et al. 2014), que utiliza o modelo de referência (CARD et al. 1999a) para dividir os mecanismos de interação suportados pelas técnicas conforme o nível do *pipeline* que eles afetam.

2.1.2 Processo de busca

A segunda etapa do protocolo tem por objetivo efetuar a escolha das publicações que farão parte dessa revisão. O primeiro passo na seleção das publicações foi realizar consultas em mecanismos de buscas específicos para a área acadêmica ou bibliotecas digitais (tabela 2.1) utilizando combinações de palavras chaves, como, por exemplo, “*multivariate network visualization*”, “*multidimensional graph visualization*”, “*multivariate graph exploration*”, entre outras. Os resultados foram classificados por ordem de relevância e, a partir da análise do título, foi executada uma primeira filtragem. Cada relação de publicações retornada foi avaliada página por página, sequencialmente até que os resultados não se relacionassem mais com os objetivos da busca. As publicações retornadas por mais de um mecanismo de busca foram marcadas como candidatas para análise.

Tabela 2.1 – Mecanismos de busca utilizados

Mecanismos de busca
<i>Google Scholar</i>
<i>Microsoft Academic Search</i>
<i>IEEEExplore Digital Library</i>
<i>ACM Digital Library</i>

2.1.3 Critérios de exclusão

Após o registro das publicações candidatas, foi realizada uma leitura dos resumos com o objetivo de descartar aquelas que se encaixam nos critérios de exclusão descritos abaixo:

- 1) Publicações que não se referem à visualização de grafos multivariados.
- 2) Publicações que não descrevem técnicas ou ferramentas de visualização.

- 3) Publicações anteriores a 2005. Foram incluídos apenas os trabalhos publicados até fevereiro de 2016.
- 4) Publicações que apresentam *surveys*. Embora esse tipo de publicação não tenha sido incluído na revisão, foram utilizadas na busca de novas técnicas ou ferramentas que não tenham sido selecionadas.
- 5) Publicações que tenham sido publicadas em dois locais diferentes ou que descrevam aperfeiçoamentos de uma mesma técnica. Nesse caso somente a publicação mais antiga foi citada com o objetivo de marcar o ponto inicial da investigação. Entretanto todas as versões foram usadas na avaliação.

Em um segundo momento, foi realizada uma leitura completa de cada publicação com o objetivo de verificar se realmente ela descreve uma técnica ou sistema de visualização de grafos multivariados. Uma atenção especial foi dada à seção relativa a trabalhos relacionados (e suas referências), seção cujo teor foi usado como um critério de determinação da importância relativa de cada artigo, podendo acarretar a inclusão de novas publicações à lista daquelas a serem analisadas.

2.1.4 Realização da análise

Tabela 2.2 – Técnicas classificadas

Técnicas	Referências	Ano	Implementação	Vídeos
<i>RelVis</i>	Pinzger et al. 2005	2005	---	---
<i>PivotGraph</i>	(Wattenberg, 2006)	2006	---	---
<i>NYSS</i>	Shneiderman et al. 2006	2006	www.cs.umd.edu/hcil/nvss	Mesmo
<i>NoodleView</i>	(Pretorius et al. 2006)	2006	www.comp.leeds.ac.uk/applications/noodleview	---
<i>GraphSpace</i>	(Xu et al. 2007)	2007	---	https://vimeo.com/5106969
<i>AttriGraph</i>	(Pretorius et al. 2008)	2008	www.comp.leeds.ac.uk/scsajp/applications/attrigraph	---
<i>Shannon et al.</i>	(Shannon et al. 2008)	2008	---	---
<i>GraphDice</i>	(Bezerianos et al. 2010)	2010	www.aviz.fr/graphdice	Mesmo
<i>PEX-Graph</i>	(Andery, 2010)	2010	http://infoserver.lcad.icmc.usp.br/infovis2/PEXGraph	Mesmo
<i>P-SPLOM</i>	(Viau et al. 2010)	2010	---	https://www.youtube.com/watch?v=f9Z0mPOnG_M
<i>NetworkLens</i>	(Jusufi et al. 2010)	2010	---	https://vimeo.com/128241200
<i>Cluma-GO</i>	(Kerren et al. 2011)	2011	sourceforge.net/projects/go-cluster	---
<i>GraphTrail</i>	(Dunne et al. 2012)	2012	---	https://www.youtube.com/watch?v=Wic49KHXf88
<i>ViNCent</i>	(Kerren et al. 2012)	2012	---	https://vimeo.com/128268456
<i>enRoute</i>	(Partl et al. 2012)	2012	http://www.caleydo.org/tools/pathways/	Mesmo
<i>SocialNetSense</i>	(Gou et al. 2012)	2012	---	http://zhang.ist.psu.edu/demo/SocialNetSense/SocialNetSense.mov
<i>JauntyNets</i>	(Jusufi et al. 2013)	2013	---	---
<i>PNLB</i>	(Ghani et al. 2013)	2013	---	---
<i>PIWI</i>	(Yang et al. 2013)	2013	---	https://www.youtube.com/watch?v=KC3cpBM5G3Q
<i>Jusufi et al.</i>	(Jusufi et al. 2014)	2014	---	---
<i>OnionGraph</i>	(Shi et al. 2014)	2014	---	http://lcs.ios.ac.cn/~shil/video/OnionGraph_2014.wmv
<i>Elzen e Wijk</i>	(Elzen et al. 2014)	2014	---	https://www.youtube.com/watch?v=E1PVTij7h0
<i>Ko et al.</i>	(Ko et al. 2014)	2014	---	https://vimeo.com/121038387
<i>AmbiguityVis</i>	(Wang et al. 2015)	2015	---	vimeo.com/136205887
<i>g-Miner</i>	(Cao et al. 2015)	2015	---	nancao.org/videos/gminer.mp4
<i>Wu et al.</i>	(Wu et al. 2015)	2015	---	vimeo.com/129251877
<i>egoSlider</i>	(Wu et al. 2016)	2016	vis.cse.ust.hk/egoslider	vimeo.com/131969585

E, finalmente, na última etapa, foi realizada a análise propriamente dita das 27 técnicas selecionadas (Tabela 2.2) a partir das descrições contidas nos artigos que as apresentam. Para responder as questões relacionadas aos aspectos interativos (QP6 e QP7), consultas aos sites dos autores foram realizadas com objetivo de localizar implementações ou vídeos de demonstração.

2.2 Descrição das técnicas selecionadas para análise

Esta seção descreve de forma resumida as 27 técnicas (ou sistemas) selecionadas para a análise. A tabela 2.2 mostra a relação de técnicas por ordem de ano em que os artigos que as descrevem foram publicados. O significado de cada coluna é apresentado a seguir:

- a) **Técnica:** Termo que identifica a técnica (ou sistema). Caso não haja um identificador, é usado o sobrenome do primeiro autor do artigo.
- b) **Referência:** Referência para o artigo que contém a descrição da técnica.
- c) **Ano:** Ano que o artigo que descreve a técnica foi publicado.
- d) **Implementação:** Endereço do site que contém a ferramenta que implementa a técnica, quando disponível.
- e) **Vídeo:** Endereço do site no qual o vídeo de demonstração da ferramenta foi publicado (se disponível).

RelVis (PINZGER et al. 2005) foi proposta para a visualização da evolução das métricas de entidades e relacionamentos existentes em várias versões do código fonte de um software (Figura 2.1).

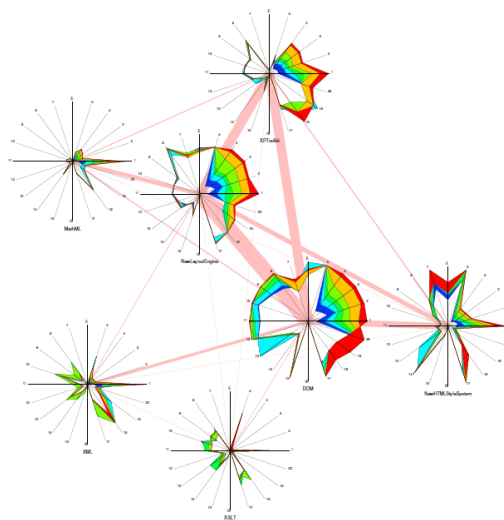


Figura 2.1 – *RelVis*
Fonte: Pinzger et al. (2005)

Os relacionamentos são representados por meio de um diagrama de nodos e arestas, com gráficos *Kiviat* exibidos na posição dos nodos para representar como as medidas das

várias métricas do software evoluem em cada versão. As arestas que representam os relacionamentos são exibidas na forma de um retângulo cujas dimensões e cor possibilitam mapear até três atributos. Embora essa abordagem permita mostrar os valores dos atributos e relacionamentos simultaneamente, possui problemas de escalabilidade em função das dimensões necessárias para exibir de forma legível os diagramas *Kiviat*.

A técnica *PivotGraph* (WATTENBERG, 2006) foi proposta com o objetivo de destacar a correlação entre as dimensões de um grafo multivariado (Figura 2.2).

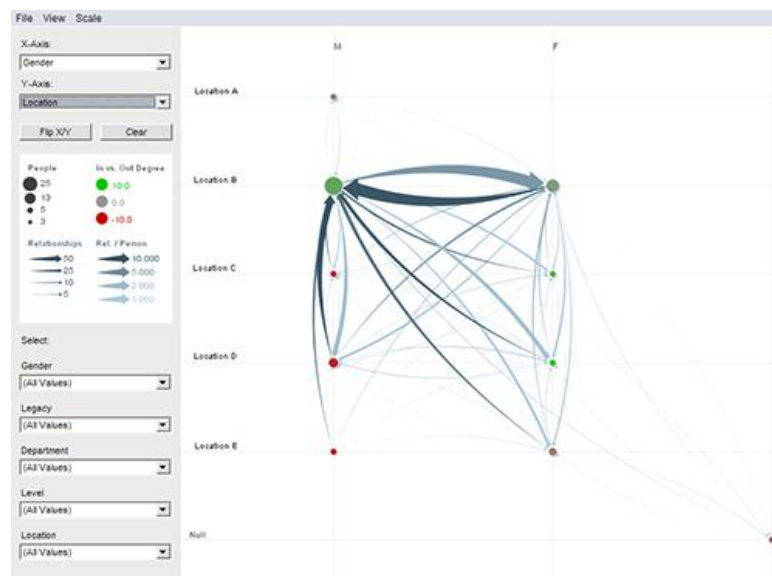


Figura 2.2 – *PivotGraph*
Fonte: Wattenberg (2006)

Nela, os nodos que possuem o mesmo valor para determinados atributos, definidos estes pelo usuário, são agregados e dispostos em uma grade bidimensional semelhante a uma planilha. Essa estratégia possibilita aos usuários reduzir substancialmente a complexidade da estrutura a ser analisada. Embora possa ser utilizada para avaliar a relação de vários atributos, no máximo dois podem ser representados simultaneamente, já que cada um dos eixos X, Y da grade representa uma dimensão. Cada nodo é desenhado na forma de um círculo, com área proporcional a quantidade de nodos agregados que ele representa. As arestas são desenhadas no formato de curvas quadráticas, para evitar sobreposição, já que a disposição dos nodos na forma de grade pode gerar várias arestas com a mesma direção conectando nodos em linha. A espessura de cada aresta é proporcional ao peso que elas representam e são desenhadas segundo uma ordem pré-estabelecida. *PivotGraph* possui dois mecanismos principais para a interação: uma operação denominada *roll-up*, que agrega os nodos com o mesmo valor de atributo, como mencionado anteriormente, e a seleção, que permite obter um subgrafo que possui apenas os nodos cujos valores dos atributos foram determinados pelo usuário.

Na ferramenta NVSS (*Network Visualization by Semantic Substrates*) (Figura 2.3), proposta por Shneiderman e Aris (2006), os nodos são agrupados de acordo com o valor de um de seus atributos em regiões não sobrepostas. Cada grupo de nodos é exibido em uma das regiões utilizando um método de posicionamento escolhido pelo usuário dentro de um conjunto pré-estabelecido que define a forma como um ou mais atributos serão mapeados para os eixos x e y. As dimensões de cada região são proporcionais à quantidade de nodos que elas contêm. A visibilidade das arestas também é controlada pelo usuário permitindo a sua exibição de acordo com as regiões de origem e destino de cada uma delas, minimizando a desordem visual.

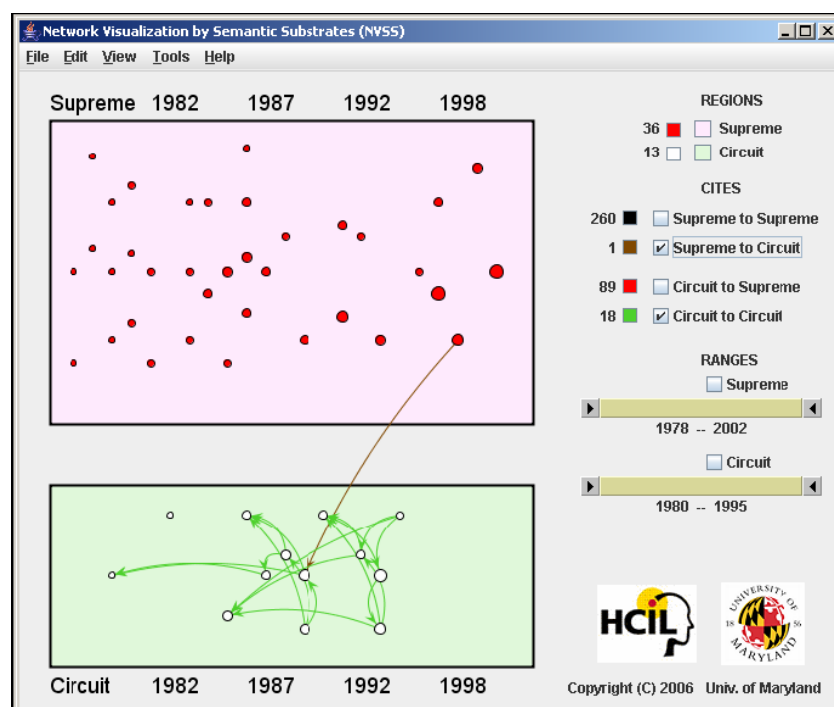


Figura 2.3 – NVSS: *Network Visualization by Semantic Substrates*
 Fonte: Shneiderman e Aris (2006)

NoodleView (PRETORIUS et al. 2006) é um protótipo que implementa uma técnica para visualização de grafos multivariados que representam transição de estados (Figura 2.4). A estratégia adotada pela técnica integra a representação de três diferentes estruturas (hierárquica, relacional e métrica) em uma única área de visualização. A hierarquia, representada por uma árvore, é gerada a partir de um agrupamento dos nodos levando em conta um conjunto de atributos selecionados. Os valores quantitativos dos nodos são visualizados por meio de uma técnica denominada *bar tree*, que se assemelha a um gráfico de barras posicionado abaixo da hierarquia na metade inferior da área de visualização, enfatizando a relação existente entre elas. As arestas que definem a transição entre os estados

do sistema representado também são agrupadas e conectam somente os nodos folhas da hierarquia. Cada folha representa uma coleção de nodos interconectados por arestas dirigidas. As que possuem o mesmo sentido são empacotadas em um mesmo arco de círculo (WATTENBERG, 2002), com espessura proporcional à quantidade de arestas que ele representa, de forma a ligar seu nodo de origem e destino.

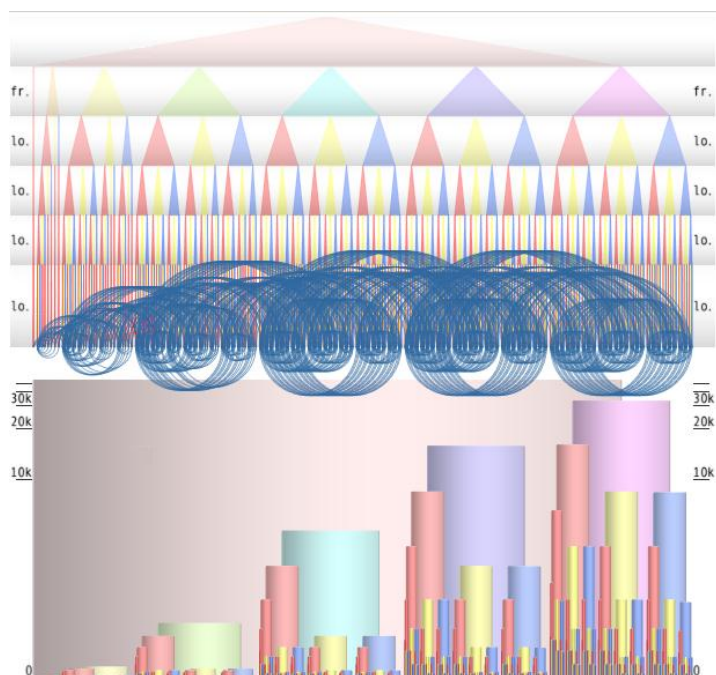


Figura 2.4 – *NoodleView*
Fonte: Pretorius et al. (2006)

GraphScape (XU et al. 2007) utiliza a metáfora *landscape* para visualizar simultaneamente a topologia do grafo e o valor de seus atributos (Figura 2.5). Nela, o grafo é exibido sobre um plano e o valor de seus atributos representados por uma superfície tridimensional. Os nodos são posicionados por meio de uma versão modificada do algoritmo *force-directed* (EADES, 1984) de forma que aqueles com atributos similares fiquem próximos uns dos outros. A superfície mencionada anteriormente é construída com o uso de uma malha tridimensional onde a altura dos vértices de seus polígonos é definida em função dos valores de atributo de todos os nodos localizados a certa distância. A contribuição de cada nodo no cálculo da altura de um vértice diminui à medida que a sua distância aumenta. Essa estratégia ajuda a revelar agrupamentos, pois nodos próximos entre si com valores semelhantes para um atributo tendem a destacar a superfície gerada diferenciando-se do seu entorno. O valor da constante que define essa distância pode ser modificado pelo usuário, provocando uma significativa alteração na representação. Valores pequenos enfatizam o atributo de nodos individuais, ao contrário de valores maiores, que facilitam a visualização de agrupamentos.

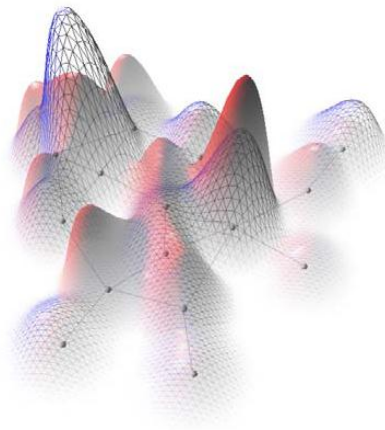


Figura 2.5 – *GraphScape*
Fonte: Xu et al. (2007)

Da mesma forma que a técnica *NoodleView* (PRETORIUS et al. 2006), *AttriGraph* (PRETORIUS et al. 2008) também possui o objetivo de visualizar grafos de transição de estados (Figura 2.6). Entretanto, nessa técnica, Pretorius enfatiza a visualização das arestas. Nela, cada aresta representa uma transição entre dois estados e possui um texto associado que é exibido no interior de um retângulo no centro da área de visualização. As arestas que possuem o mesmo texto compartilham o mesmo retângulo. Elas são dirigidas e visualizadas por meio de linhas que conectam seus nodos de origem, exibidos na área localizada à esquerda, aos de destino, exibidos à direita, representados como um grafo bipartido. De forma semelhante à técnica descrita anteriormente, os nodos são agrupados baseados no valor dos atributos selecionados pelo usuário formando duas hierarquias na qual seus nodos folhas são representados na parte interna da área de visualização. A técnica possui mecanismos de seleção e filtragem que permitem ao usuário realizar consultas na estrutura.

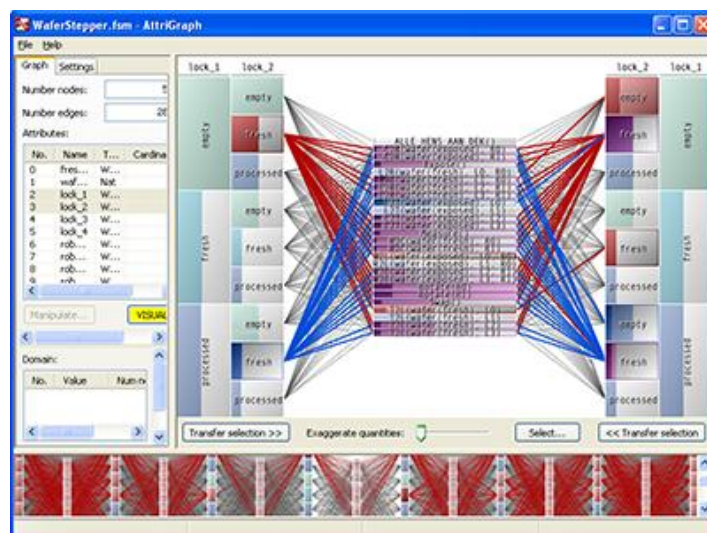


Figura 2.6 – *AttriGraph*
Fonte: Pretorius et al. (2008)

A ferramenta (Figura 2.7) proposta por Shannon et al. (2008) utiliza duas visões coordenadas distintas. Em uma é exibida a topologia do grafo por meio de um diagrama de nodos e arestas cujo *layout* é determinado por meio de um algoritmo dirigido por força. Na outra, os valores dos atributos são correlacionados pelo uso de um diagrama de coordenadas paralelas. Várias técnicas de interação são incorporadas com o objetivo de filtrar e manter a coordenação entre as visões. A seleção é usada para destacar nodos (individuais ou grupos) de interesse simultaneamente nas visões e pode ser executada tanto sobre o diagrama de nodos e arestas como nas coordenadas paralelas.

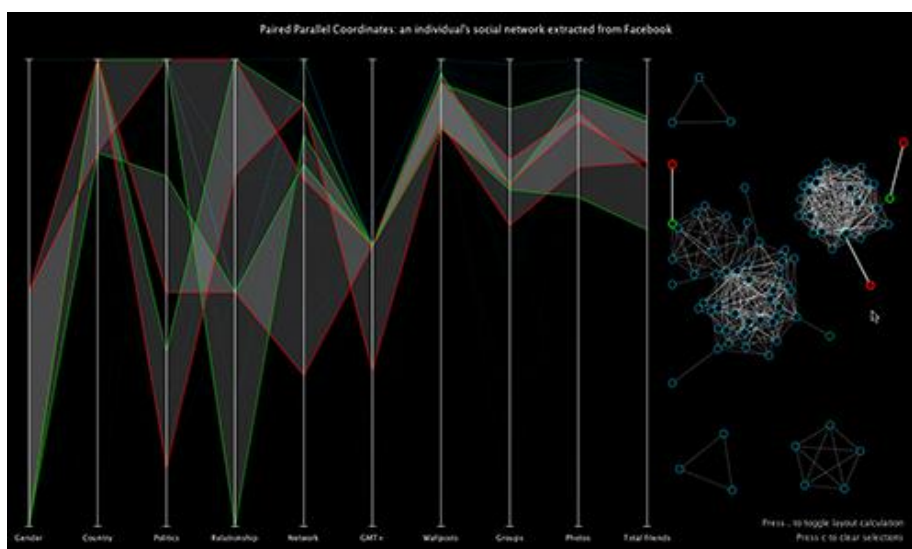


Figura 2.7 – *Paired Parallel Coordinates*
Fonte: Shannon et al. (2008)

GraphDice (BEZERIANOS et al. 2010) é um sistema de visualização projetado para a exploração de redes sociais multivariadas (Figura 2.8). É baseado na *ScatterDice* (ELMQVIST et al. 2008), uma técnica usada para navegação em tabelas de dados multidimensionais e adaptada para visualização de grafos. *GraphDice* exibe o grafo em uma espécie de *scatterplot* onde a posição de cada nodo é determinada pelo valor de dois de seus atributos selecionados para cada um dos dois eixos. Linhas conectando seus pontos são acrescentadas ao gráfico com o objetivo de exibir os relacionamentos, gerando um diagrama de nodos e arestas. A seleção dos dois atributos que serão visualizados é realizada com o auxílio de uma matriz de miniaturas de *scatterplots* que apresenta todas as combinações possíveis de atributos. A visão geral fornecida por esse mecanismo define um espaço de navegação e possibilita ao usuário uma forma rápida de exploração de hipóteses. Um mecanismo denominado *excentric labels lens* (FEKETE; PLAISANT, 1999), formado por um quadrado que pode ser movido sobre o grafo, está disponível para exibir detalhes de nodos

específicos. *GraphDice* também fornece um mecanismo de consultas visuais que permitem que o usuário possa filtrar interativamente o conjunto de dados. É mantido um histórico das consultas, exibindo miniaturas de seu respectivo *scatterplots*, o que permite ao usuário refazer os passos do processo de exploração quando necessário.

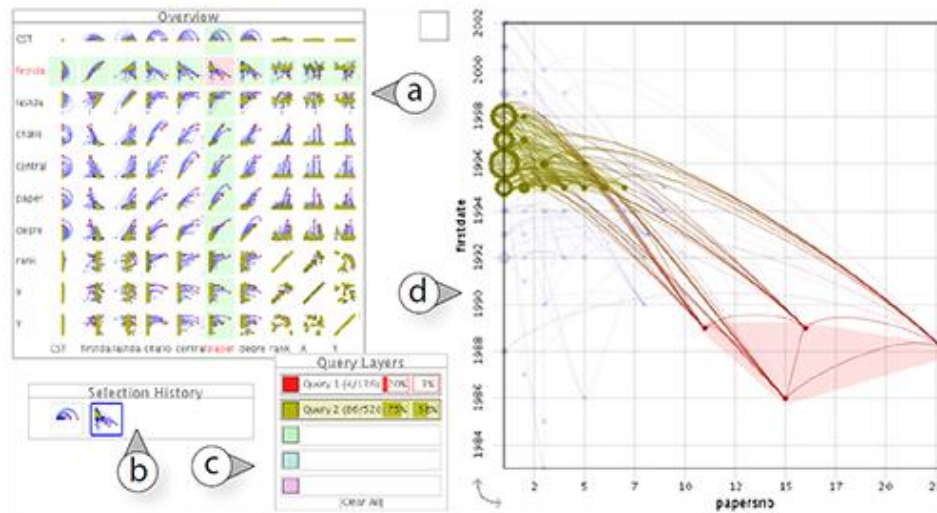


Figura 2.8 – *GraphDice*
Fonte: Bezerianos et al. (2010)

PEX-Graph (ANDERY 2010) foi proposta para visualização de redes sociais (Figura 2.9). Possibilita a exploração de redes heterogêneas que representam grafos com diferentes tipos de vértices e utiliza técnicas de projeções multidimensionais para reduzir a quantidade de dimensões do conjunto de dados, de forma a mapeá-los no espaço bidimensional. O critério utilizado nesse mapeamento é o grau de dissimilaridade entre os vértices, levando-se em conta o valor de seus atributos ou o peso de seus relacionamentos. Dessa forma, torna-se possível identificar grupos de indivíduos que possuem valores de atributos similares por meio de sua proximidade no plano. *PEX-Graph* suporta dois tipos de coordenação: coordenação por identidade e coordenação relacional. A coordenação por identidade cria um relacionamento direto entre os mesmos vértices exibidos em diferentes visualizações de forma que um vértice (ou conjunto) selecionado em uma representação visual seja destacado nas outras. Na coordenação relacional é estabelecido um relacionamento de muitos para muitos entre os vértices nas diferentes representações visuais. A ferramenta *PEX-Graph* possui ainda vários mecanismos de interação que permitem a seleção, filtragem e busca textual de seus elementos, assim como alteração de aspectos visuais.

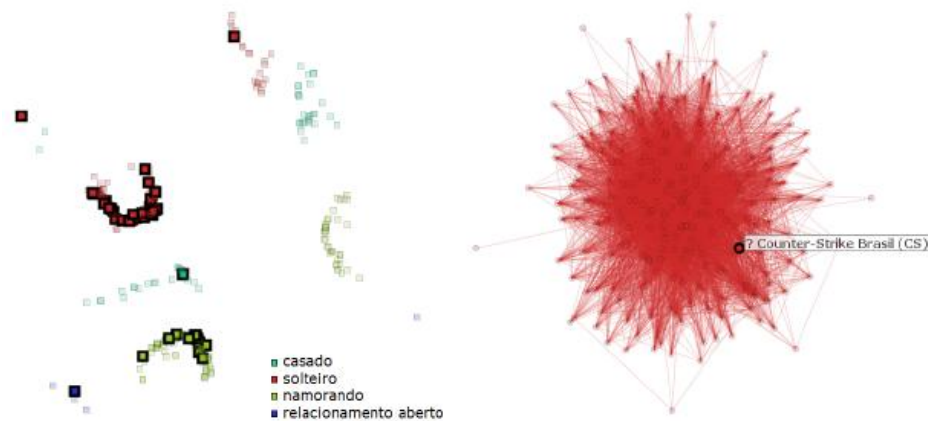


Figura 2.9 – *PEx-Graph*
Fonte: Andery (2010)

P-SPLOM (*Parallel Scatterplot Matrix*) (VIAU et al. 2010) é uma técnica que combina uma matriz de miniaturas de *scatterplots* e coordenadas paralelas com o objetivo de permitir a visualização e exploração de redes multivariadas (Figura 2.10). A matriz exibe uma visão geral dos dados com todas as combinações de atributos, facilitando a comparação entre eles. Um *scatterplot* ampliado mostra uma visão detalhada da miniatura selecionada, na qual a relação existente entre seus elementos é representada por meio de um diagrama de nodos e arestas. O *layout* padrão utilizado para o diagrama de nodos e arestas é determinado pelo valor dos atributos. Nodos podem ser fixados na área de visualização em posições determinadas pelo usuário (*layout* manual) ou selecionados para serem posicionados segundo um algoritmo dirigido por força. Dessa forma, novas seleções de atributos visualizados no *scatterplot* não terão influência no seu posicionamento. A técnica também utiliza visões coordenadas para conectar os nodos selecionados no *scatterplot* com um gráfico de coordenadas paralelas. Esse gráfico é obtido a partir de uma linha da matriz de miniaturas de *scatterplots* e pode ser rotacionado no espaço tridimensional em torno dos eixos vertical ou horizontal produzindo um efeito 3D no gráfico de coordenadas paralelas.

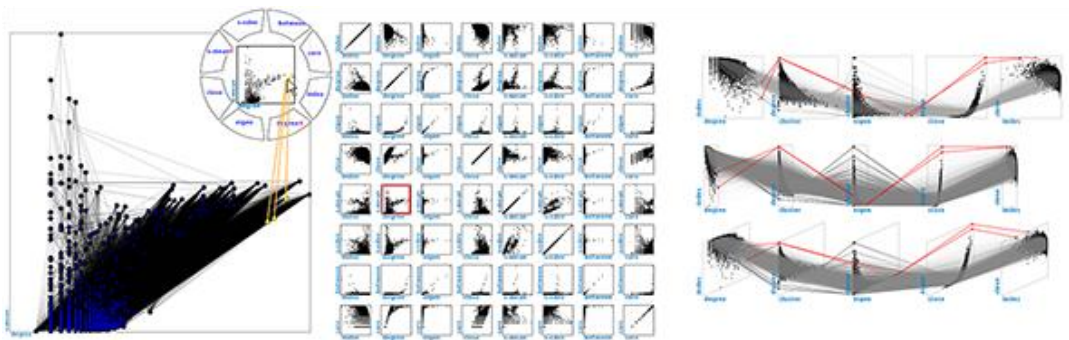


Figura 2.10 – *P-SPLOM: Parallel Scatterplot Matrix*
Fonte: Viau et al. (2010)

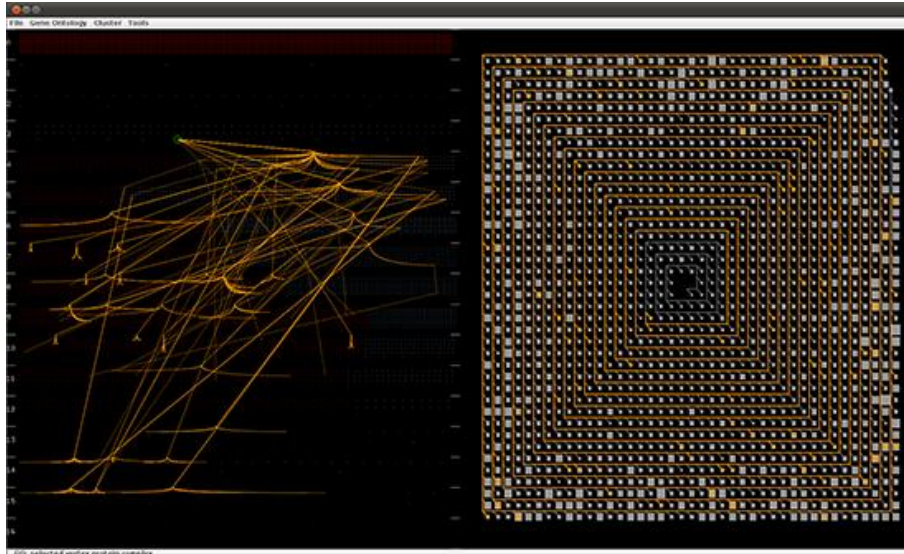


Figura 2.12 – *CluMa-GO*
Fonte: Kerren et al. (2011)

GraphTrail (DUNNE et al. 2012) é uma ferramenta criada para a exploração de grafos multivariados com vários atributos associados tanto a nodos quanto a arestas (Figura 2.13).

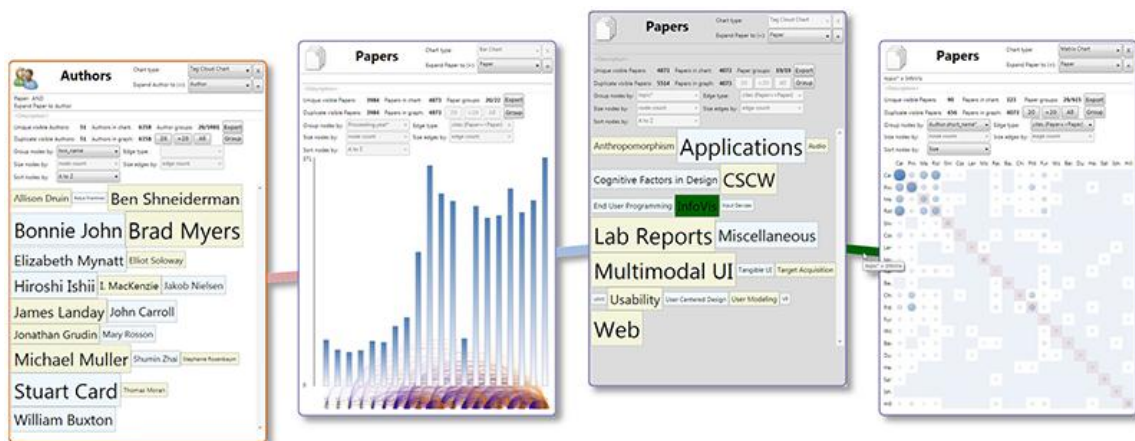


Figura 2.13 – *GraphTrail*
Fonte: Dunne et al. (2012)

Ela permite que o usuário execute uma agregação nos elementos do grafo baseada em atributos de nodos e arestas. Dessa forma, os dados podem ser interativamente explorados por meio de gráficos familiares como *tag clouds*, gráfico de barras, matrizes, entre outros, possibilitando que a melhor alternativa de visualização possa ser escolhida. A exploração é realizada a partir de um gráfico na qual parte do conjunto de dados exibido pode ser selecionada para a criação de uma nova visualização conectada a anterior, criando uma cadeia que mantém o histórico da exploração. *GraphTrail* gerencia a coordenação entre os vários

gráficos exibidos na área de visualização de forma que os elementos selecionados em uma visualização sejam destacados nas outras.

ViNCent (*Visualization of Network Centralities*) (KERREN et al. 2012) é uma ferramenta projetada para a análise de diversas medidas de centralidade dos nodos de um grafo (Figura 2.14). Nela o grafo representado por um diagrama de nodos e arestas com *layout* radial. Os valores de centralidade de cada nodo são exibidos por meio de *stacked bars* de forma a posicioná-los, sobre cada nodo, na direção dos raios do círculo. Múltiplas visões coordenadas são usadas para apoiar as análises de centralidade. Em uma delas, vários histogramas exibem dados estatísticos sobre cada medida possibilitando ao usuário uma melhor compreensão da distribuição dos valores de centralidade sobre todos os nodos. Em outra, um histograma mostra os valores de todas as medidas para um nodo selecionado do grafo. Mecanismos de filtragem podem ser acionados a partir dos histogramas para reduzir a quantidade de nodos exibidos.

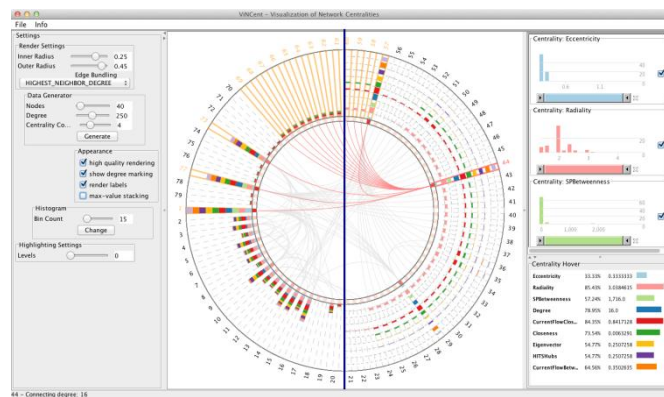


Figura 2.14 – *ViNCent: Visualization of Network Centralities*
Fonte: Jusufi et al. (2013)

enRoute (PARTL et al. 2012) é uma ferramenta projetada para a exploração de dados experimentais em processos biomoleculares extraídos de *pathways* metabólicas (Figura 2.15). Possui duas visões coordenadas denominadas de *pathway view* e *enRoute view*. A *pathway view* exibe um grafo representando um *pathway map* que fornece informações topológicas completas bem como uma visão geral dos dados experimentais. Possui várias alternativas de seleção que permitem destacar, pelo uso de uma versão modificada da técnica *Bubble Sets* (COLLINS et al. 2009), os caminhos a serem analisados. O caminho escolhido é exibido na *enRoute view* em um *layout* linear vertical na qual os atributos relacionados ao dados experimentais são visualizados ao lado de cada nodo, estando esses atributos representados por uma matriz de gráficos de barras. Vários mecanismos de interação são fornecidos para a exploração dos dados.

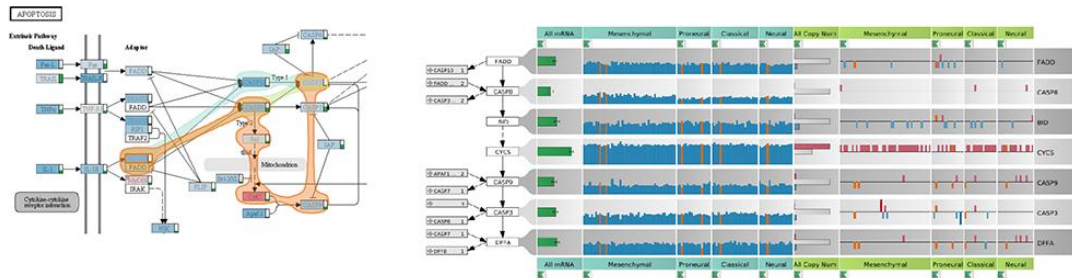


Figura 2.15 – *enRoute*
Fonte: Partl et al. (2012)

SocialNetSense (GOU et al. 2012), ferramenta construída para análise de redes sociais (Figura 2.16), é constituída de dois módulos denominados *Network Exploring Space (NES)* e *Representation Building Space (RBS)*.

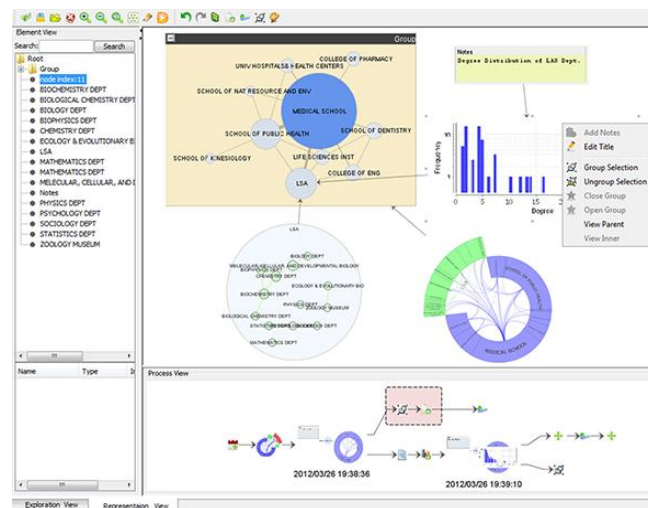


Figura 2.16 – *SocialNetSense*
Fonte: Gou et al. (2012)

NES faz uso de múltiplas visões coordenadas que incluem um diagrama de nodos e arestas tradicional para exibir a topologia da rede, uma visualização de árvore que mostra a estrutura social dos atores e um painel que exibe uma visão geral da rede de forma a facilitar a navegação. *NES* também faz uso da técnica *TreeNetViz* (GOU et al. 2011), que permite exibir uma estrutura híbrida formada pela rede e por uma hierarquia na mesma área de visualização. *NES* oferece ainda um conjunto de ferramentas analíticas, tais como diferentes métricas de agregação de conectividade e centralidade assim como funções de busca. O módulo *RBS* é uma espécie de editor que permite que o usuário copie evidências oriundas dos diagramas construídos com a *NES* (sub-grafos, métricas, nodos, *clusters*, entre outras) organizando-as em uma representação visual própria à medida que a exploração vai sendo executada. Notas com informações adicionais podem ser acrescentadas a ela e relações entre os elementos

podem ser estabelecidas por meio de conexões. Um diagrama do processo de criação dessa representação é exibido em outra área de visualização.

JauntyNets (JUSUFI et al. 2013) combina um *layout* dirigido por atributos com múltiplas visões coordenadas (Figura 2.17). Nela o grafo é representado por um diagrama de nodos e arestas localizado no interior de uma área circular que é delimitada por retângulos dispostos em um *layout* radial representando atributos. Estes são conectados aos respectivos nodos por meio de linhas. A técnica utiliza um algoritmo baseado em força que leva em conta o valor dos atributos no cálculo do posicionamento dos nodos. São disponibilizados vários mecanismos de interação, como filtros para controlar a visibilidade das arestas e seleção de nodos baseados nos atributos ou estrutura do grafo como, por exemplo, o destaque dos nodos. Em uma de suas visões coordenadas um gráfico de barras é usado para exibir a quantidade de nodos que possui um determinado valor de atributo. Outra é utilizada para exibir o grafo por meio de técnicas de projeção multidimensional.

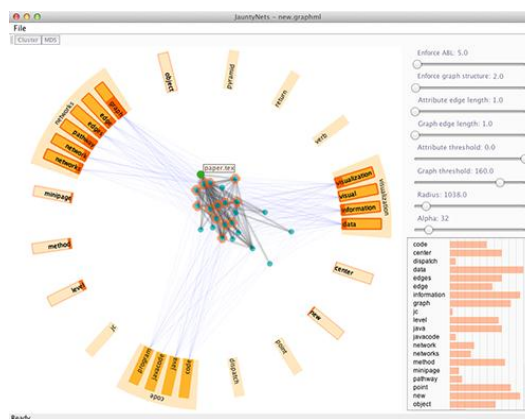


Figura 2.17 – *JauntyNets*
Fonte: Jusufi et al. (2013)

PNLB (*Parallel Node-Link Bands*) (GHANI et al. 2013) é uma técnica projetada para a visualização e análise de redes sociais multimodais, as quais são constituídas de vários tipos de nodos e arestas (Figura 2.18). Ela usa uma representação que divide a área de visualização em várias listas verticais onde os nodos separados por modos são exibidos. Linhas são usadas para representar os relacionamentos e apenas aquelas que relacionam nodos pertencentes a bandas adjacentes são exibidas. As listas podem ser reordenadas com o objetivo de permitir que os relacionamentos omitidos passem a ser exibidos. Além de mecanismos tradicionais como *panning* e *zoom*, a *PNLB* utiliza uma função *fisheye* que permite a expansão e compressão dos retângulos que representam os nodos quando o cursor do mouse é movido sobre eles. Também permite a exibição de um diagrama de coordenadas paralelas integrado a uma das listas, escolhida pelo usuário, para mostrar a correlação entre seus atributos.

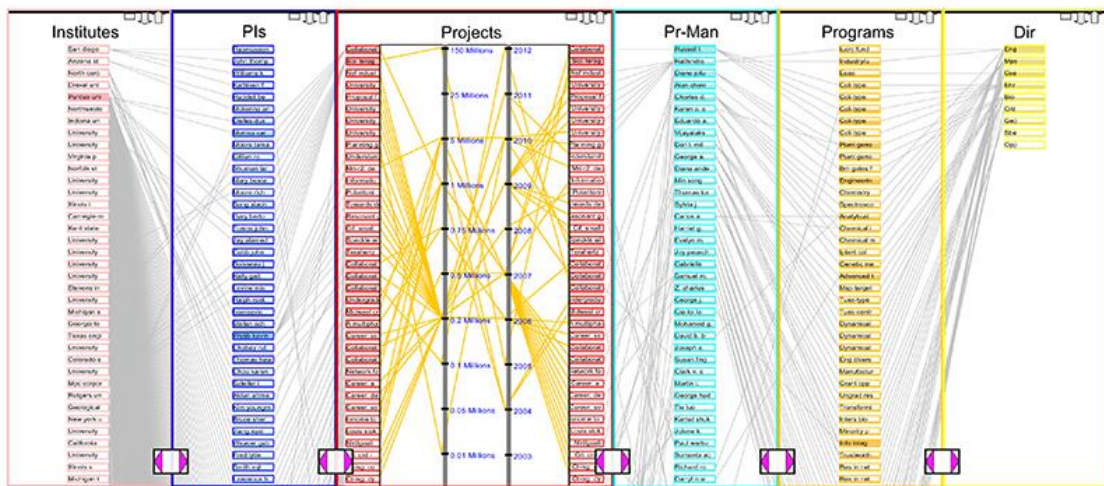


Figura 2.18 – PNLB: *Parallel Node-Link Bands*
Fonte: Ghani et al. (2013)

PIWI (YANG et al. 2013) tem por objetivo a exploração visual de comunidades em grafos que podem ser detectadas automaticamente por diversos algoritmos integrados ao sistema (Figura 2.19). *PIWI* utiliza *tags clouds* e técnicas orientadas a *pixel*, para exibir respectivamente os membros de cada comunidade e seus relacionamentos. As informações são apresentadas em uma grande tabela na qual cada linha representa uma comunidade. Na primeira coluna de cada linha, *labels* dos nodos pertencentes à comunidade são exibidos na forma de uma *tag cloud* na qual atributos topológicos, escolhidos pelo usuário, como grau ou valor de centralidade, são mapeados para o tamanho da fonte. A coluna seguinte apresenta uma sequência de gráficos formada por pontos (*vertex plots*), sendo que os do primeiro gráfico representam os nodos contidos na respectiva comunidade. Já no segundo, os pontos representam os vértices adjacentes ao primeiro grupo e, no terceiro, aqueles localizados a uma distância menor ou igual a dois. Não há linhas representando as conexões entre os nodos. Gráficos de pontos também são usados para exibir apenas os nodos que possuem o valor 1 para um determinado atributo binário. Mecanismos de seleção permitem que nodos sejam selecionados baseados em características estruturais ou em seus atributos.

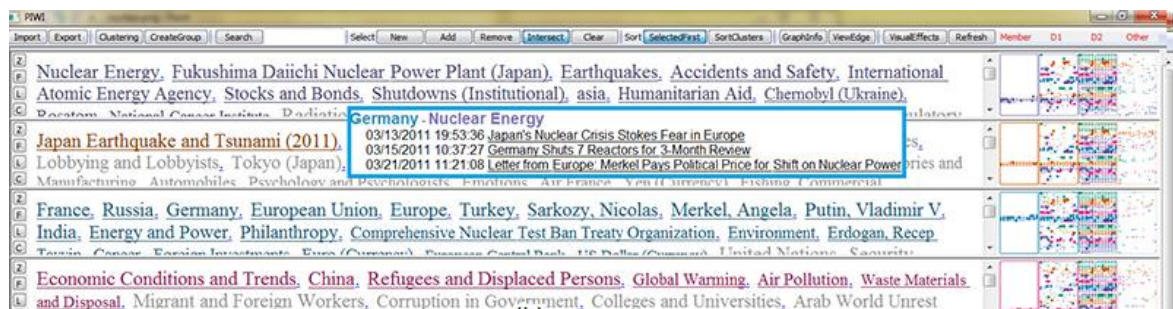


Figura 2.19 – *PIWI*
Fonte: Yang et al. (2013)

Jusufi et al. (2014) desenvolveram uma técnica para a visualização de documentos (Figura 2.20) e seus relacionamentos de coautoria, na qual são aplicados aos nodos algoritmos de clusterização que os agrupam baseado em determinados valores de seus atributos. Cada *cluster* é representado por um círculo e seus respectivos elementos são exibidos na forma de pequenos setores distribuídos radialmente formando um anel em sua parte externa. Uma *tag cloud*, representando a frequência com a qual aparecem os termos conceituais dos documentos contidos no *cluster*, é exibida no interior do círculo. Os relacionamentos entre os *clusters*, bem como os relacionamentos entre seus elementos, são representados por linhas exibidas sob demanda, conforme a seleção executada pelo usuário. Cores são usadas para mapear atributos dos nodos e arestas e vários mecanismos de interação são disponibilizados para facilitar a exploração.

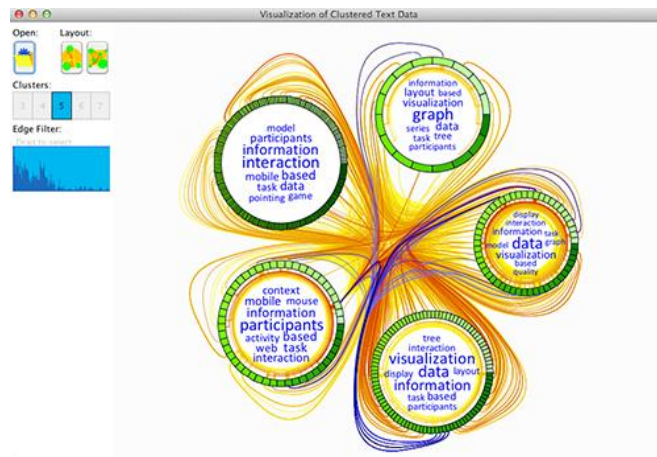


Figura 2.20 – Visualização de *cluster* de documentos
Fonte: Jusufi et al. (2014)

OnionGraph (SHI et al. 2014) é um sistema projetado para a visualização de grandes grafos heterogêneos, possibilitando ao usuário a exploração dos padrões topológicos e atributos atrelados a nodos e arestas (Figura 2.21). Ele permite criar vários níveis de abstração do grafo a partir da aplicação de diversos algoritmos de agregação que levam em conta o valor dos atributos, a estrutura topológica ou uma mistura de ambos. O sistema oferece mecanismos para que o usuário selecione uma porção de interesse do grafo, especifique um nível de abstração e acesse uma representação que exiba informações mais detalhadas da estrutura. *OnionGraph* utiliza uma representação baseada em nodos e arestas. Os nodos que representam agregações são exibidos com múltiplos círculos concêntricos indicando quantos níveis de abstração eles contêm. Assim como o diâmetro dos círculos codifica a quantidade de elementos que ele representa, a espessura das arestas também codifica a quantidade de ligações existentes entre os grupos. Durante a exploração, histogramas associados aos nodos podem ser ativados para exibir os valores de seus atributos.

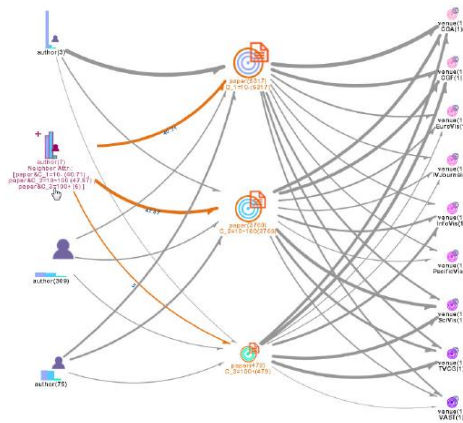


Figura 2.21 – *OnionGraph*

Fonte: Shi et al. (2014)

O sistema proposto por Elzen e Wijk (2014) tem por objetivo explorar e analisar a estrutura do grafo e os atributos associados a seus nodos e arestas de forma simultânea (Figura 2.22). Ele utiliza uma estratégia denominada *DOSA* (*Detail to Overview via Selections and Aggregations*) na qual fornece aos usuários mecanismos para que grupos de dados, selecionados em uma área que contém uma visão geral do grafo, possam ser explorados em seus detalhes. Nessa área cada nodo da estrutura é exibido utilizando-se um *layout* definido por uma projeção bidimensional baseada nos atributos dos nodos. Somente os nodos são exibidos. Durante o processo de seleção apenas as arestas conectadas aos nodos selecionados são exibidas. Cada grupo selecionado é visualizado em uma área justaposta por meio de um diagrama de nodos e arestas, na qual os múltiplos atributos dos elementos que o compõe podem ser explorados por meio de vários tipos de visualização para dados multidimensionais como histogramas, *scatterplots* e diagramas de coordenadas paralelas. O sistema calcula vários atributos relacionados à topologia do grafo que podem ser usados na filtragem dos elementos dos grupos selecionados.

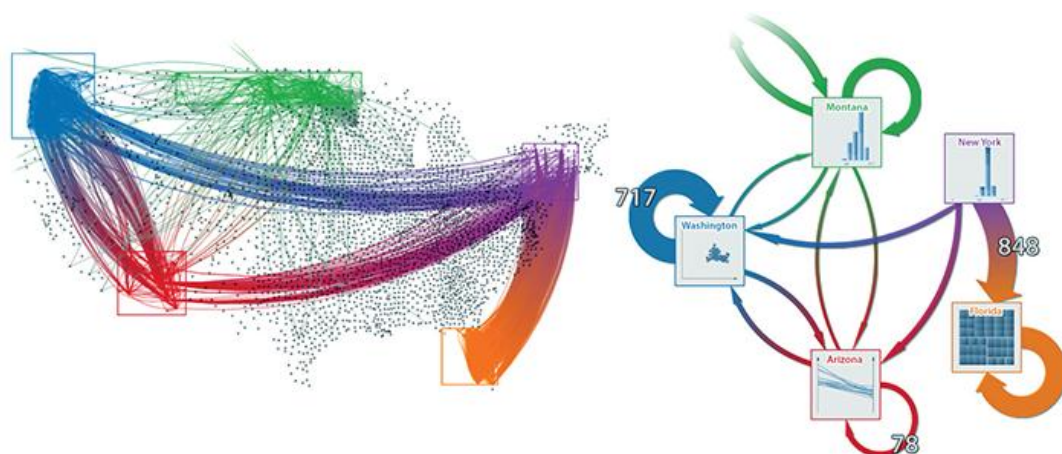


Figura 2.22 – Sistema proposto por Elzen e Wijk (2014)

Fonte: Elzen e Wijk (2014)

KO et al. (2014) propõem um sistema para explorar e analisar grafos multivariados que usa uma família de técnicas de visualização composta por histogramas, gráficos de linhas, *calendar view* entre outras. Embora ele utilize múltiplas visões coordenadas para gerenciá-las todas, neste trabalho serão descritas apenas as características das técnicas *Petal* (Figura 2.23a) e *Thread* (Figura 2.23b) projetadas para a análise dos relacionamentos existentes entre os nodos. *Petal* é exibida em cada nodo selecionado pelo usuário e tem por objetivo agregar em uma única representação informações sobre as arestas que ligam um nodo a seus nodos adjacentes. Ela utiliza um círculo, dividido em setores (pétalas) nos quais cada um representa uma aresta. Cada setor é subdividido de forma similar a um *stacked bar* para mapear os múltiplos atributos. A técnica *Thread* é usada para representar um par origem-destino de um relacionamento. Ele é representado por várias linhas, nas quais a espessura e cor são usadas para mapear os múltiplos atributos. As *Threads* que mostram os relacionamentos de um nodo com seus nodos adjacentes são exibidas quando um nodo é selecionado.

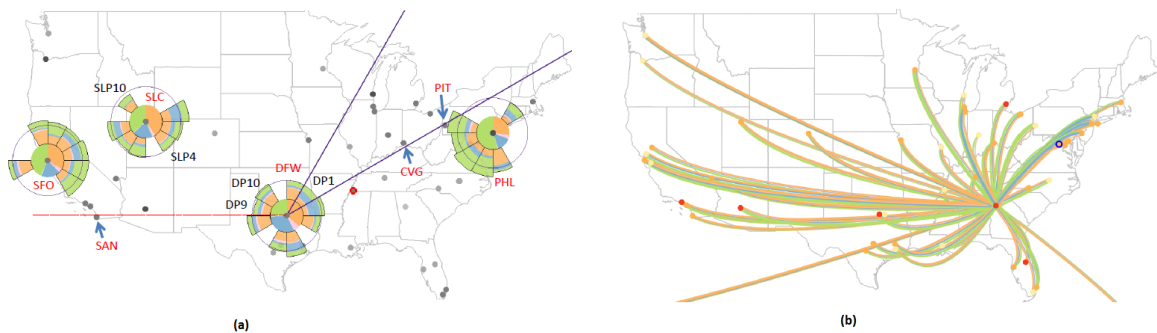


Figura 2.23 – (a) *Petal* (b) *Threads*

Fonte: Ko et al. (2014)

AmbiguityVis (Wang et al. 2015) é uma técnica projetada para revelar três classes de ambiguidades que podem ocorrer em diagramas de nodos e arestas na relação espacial entre nodos e arestas (occlusão de nodos, cruzamentos de arestas, etc.), na sobreposição visual entre estruturas de comunidades e ambiguidades visuais resultantes de agregação de nodos/arestas. Um conjunto de métricas quantificam essas ambiguidades, sendo usadas visualizações baseadas em *heatmaps* para realçar as áreas onde se localizam (Figura 2.24b). Os valores das métricas associadas tanto a nodos como a arestas definem a cor a ser usada no mapeamento. Outras três visualizações mostram o grafo original (Figura 2.24a), o grafo com *edges bundles* (Figura 2.24c) e com nodos agregados (Figura 2.24d). Além dessas visualizações, *AmbiguityVis* utiliza um gráfico de barras para visualizar atributos de supernodos resultante de agregações (quantidade de nodos e arestas dentro dos *clusters*, de arestas entre *clusters*

distintos e ainda a densidade do grafo) e um *scatterplot* gerado por projeções MDS que mostram a consistência das *edges bundled*.

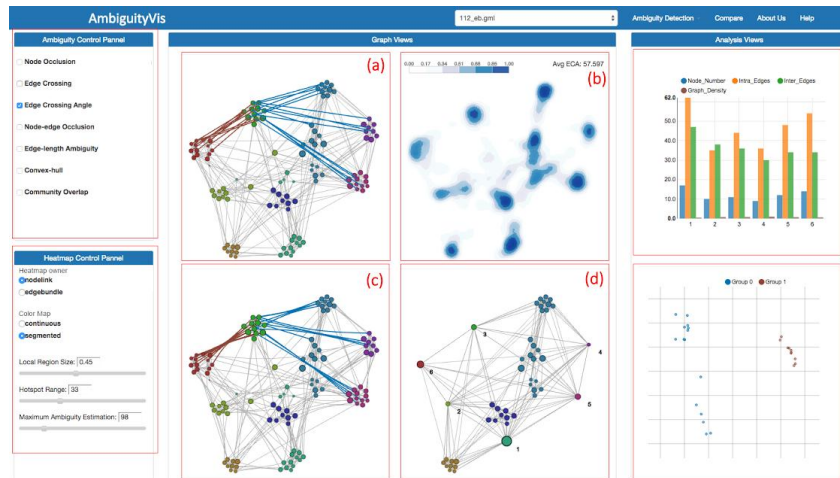


Figura 2.24 – (a) Grafo original (b) *Heatmaps* (c) *Edges Bundled* (d) Com nodos agregados
Fonte: Wang et al. (2015)

G-Miner (Cao et al. 2015) é um sistema interativo projetado com o objetivo de explorar visualmente agrupamentos em grafos multivariados. Ele é constituído de um conjunto de ferramentas que permite localizar e comparar grupos. A primeira, denominada *Hierarchy Explorer* (Figura 2.25a), mostra uma visão geral da hierarquia de agrupamentos através de uma *icicle tree*. Essa estratégia possibilita uma navegação rápida na estrutura de forma que um determinado agrupamento de interesse seja escolhido para ser visualizado através de um diagrama de nodos e arestas no *Group Explorer* (Figura 2.25b). Uma terceira ferramenta (Figura 2.25c) pode ser utilizada para permitir que o usuário altere os critérios usados na formação do grupo selecionado por meio da comparação dele com vários grupos candidatos determinados pelo sistema. *G-Miner* propõe ainda três tipos de glifos (Figura 2.25d) integrados ao diagrama de nodos e arestas para representar diferentes aspectos dos grupos, como, por exemplo, a estrutura topológica e a distribuição dos atributos.

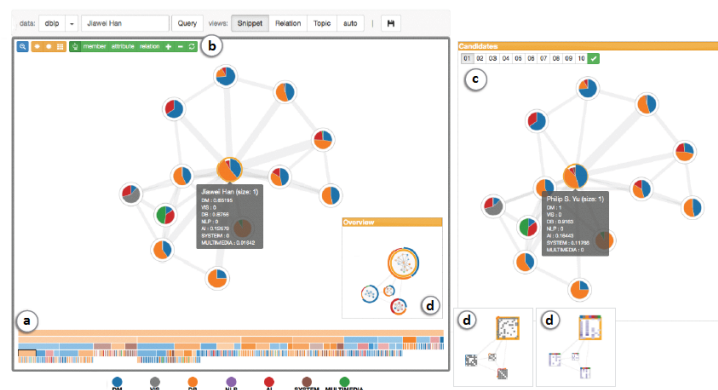


Figura 2.25 – (a) *Hierarchy Explorer* (b) *Group Explorer* (c) *Candidate Explorer* (d) Glifos
Fonte: Cao et al. (2015)

O objetivo da ferramenta proposta por Wu et al. (2015) é oportunizar a exploração e comparação das principais comunidades encontradas em um grafo. Ela utiliza uma *Voronoi Treemap* (BALZER e DEUSEEN, 2007) para mapear as comunidades em polígonos inscritos em cada célula *Voronoi* (Figura 2.26a).

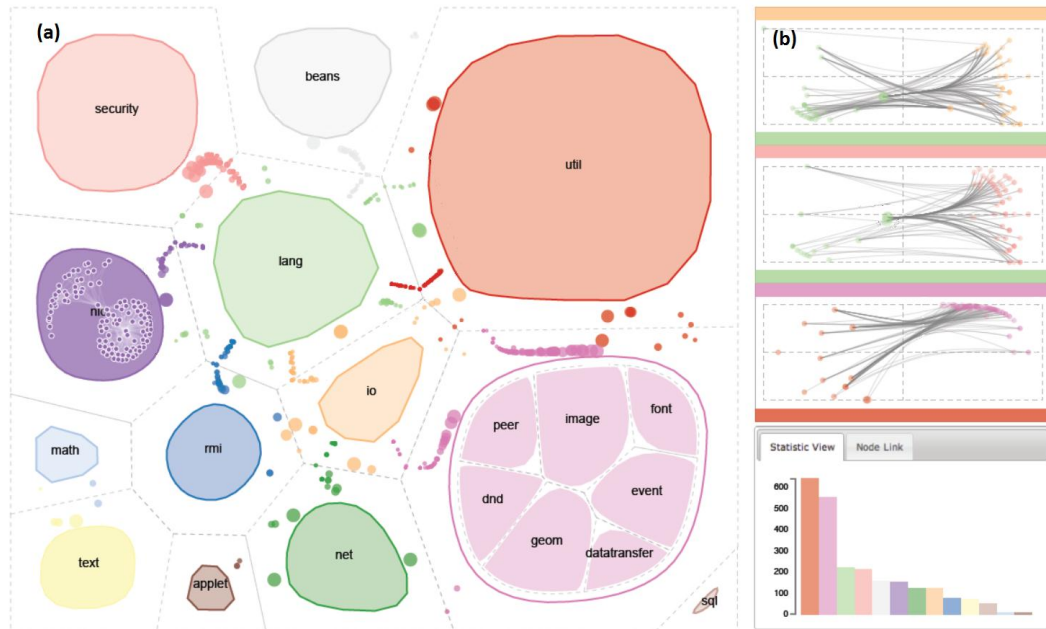


Figura 2.26 – (a) Comunidades (b) Nodos de fronteiras
Fonte: Wu et al. (2015)

Os relacionamentos entre as comunidades são representados pela posição relativa dos polígonos. Nodos que as conectam (nodos de fronteira) são distribuídos nos espaços vazios entre os polígonos. Visões que exibem de forma detalhada a conexão dos nodos de fronteira entre duas comunidades são disponibilizadas (Figura 2.26b). As comunidades com relacionamentos fortes são localizadas próximas entre si e seus diferentes atributos são mapeados através da forma, tamanho e cor dos polígonos. Gráficos de barras são fornecidos para exibir dados estatísticos de propriedades das comunidades. Mecanismos de interação para executar a fusão ou divisão de nodos são disponibilizados.

EgoSlider é um sistema de análise visual para explorar e comparar redes egocêntricas dinâmicas. Uma rede egocêntrica é aquela que representa relacionamentos entre um item específico (ego) e os elementos conectados a ele (alter). O sistema utiliza múltiplas visões coordenadas revelando padrões evolucionários em três escalas distintas: macro, média e micro. No nível macro, é exibida uma sumarização de toda a rede por meio de um conjunto de pontos localizados na área de exibição (*Data Overview*), utilizando um *layout Multidimensional Scaling* (Figura 2.27a). No nível médio, é exibida na *Timeline View* (Figura 2.27b) uma sumarização do processo evolucionário da rede egocêntrica selecionada. Dois

tipos de glifos representam respectivamente propriedades estruturais das redes (*snapshot glyph*) em um ponto no tempo e trocas entre dois pontos consecutivos (*transition glyph*). E, por fim, no nível micro é explorado por meio da *Alter Timeline View* (Figura 2.27c) que exhibe informações temporais detalhadas dos egos e seus alters.

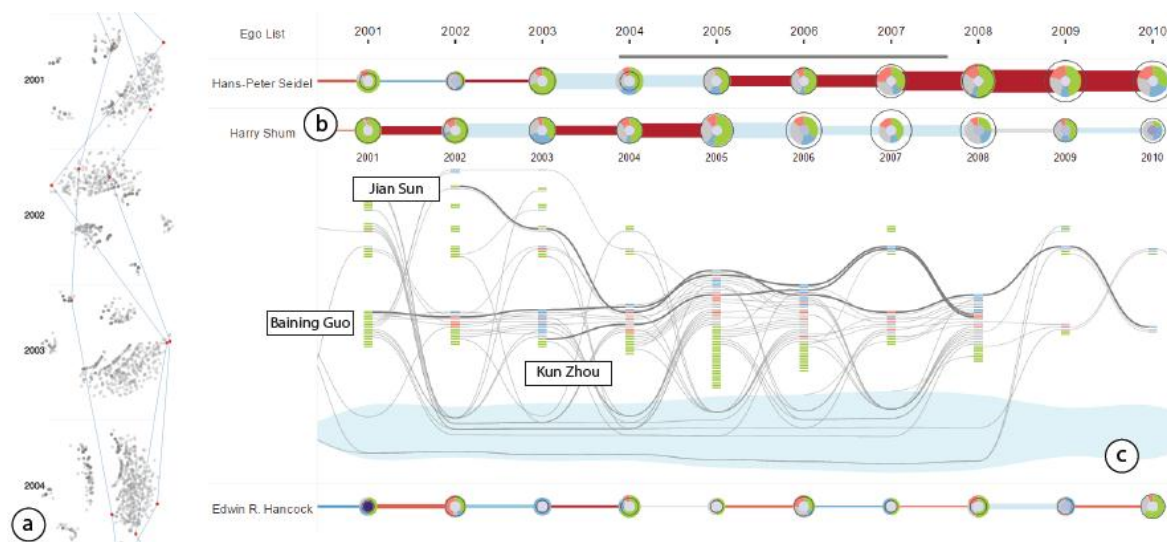


Figura 2.27 – (a) *Data Overview* (b) *Timeline View* (c) *Alter Timeline View*
Fonte: Wu et al. (2015)

2.3 Resultados da análise

Esta seção apresenta os resultados da análise das 27 técnicas (ou sistemas) selecionadas. Essa análise foi realizada a partir das descrições contidas nos artigos que as apresentam e da observação dos respectivos vídeos de demonstração (quando disponíveis). A experimentação dos softwares que as implementam também foi levada em conta, quando disponibilizados para o usuário. Marcas (●) foram usadas para identificar as características suportadas pelas técnicas. No entanto, como nem todas as características são facilmente identificáveis, pontos de interrogação foram usados para anotar aquelas em que a identificação não foi possível.

2.3.1 Área de aplicação e atributos de arestas

A tabela 2.3 apresenta o resultado da análise para as questões QP1 e QP2. A coluna “**Aplicação**” (QP1) mostra a área de aplicação para a qual a técnica (ou sistema) foi proposta. Nos casos em que as técnicas não apresentam um propósito específico uma marca é colocada na coluna “**Genérica**” (QP1). Ainda na coluna “**Aplicação**” é mostrada a área que foi adotada nos estudos de casos realizados para a avaliação dos seus protótipos. Na coluna “**Atributos de arestas**” (QP2) as técnicas foram separadas pela quantidade máxima de

atributos de arestas que podem ser representados: (i) De 1 a três atributos (ii) mais de três atributos.

Tabela 2.3 – Resultados da análise das questões QP1 e QP2

Técnicas	QP1		QP2	
	Aplicação	Genérica	Atributos de arestas	
			<=3	>3
<i>RelVis</i>	Evolução das versões de software		●	
<i>PivotGraph</i>	Redes sociais	●	●	
<i>NVSS</i>	Processos judiciais	●		
<i>NoodleView</i>	Transição de estados		●	
<i>GraphSpace</i>	Redes sociais e biológicas	●		
<i>AttriGraph</i>	Transição de estados		●	
<i>Shannon et al.</i>	Redes sociais	●	●	
<i>GraphDice</i>	Redes sociais	●		●
<i>PEX-Graph</i>	Redes sociais	●	●	
<i>P-SPLOM</i>	Redes biológicas	●	?	?
<i>NetworkLens</i>	Similaridade entre documentos	●	●	
<i>Cluma-GO</i>	Ontologia de genes			
<i>GraphTrail</i>	Redes de coautoria	●	●	
<i>ViNCent</i>	Centralidade de grafos			
<i>enRoute</i>	<i>Pathways</i> biológicos		●	
<i>SocialNetSense</i>	Redes sociais	●	●	
<i>JauntyNets</i>	Documentos	●	●	
<i>PNLB</i>	Redes sociais		●	
<i>PIWI</i>	Comunidades		●	
<i>Jusufi et al.</i>	Documentos		●	
<i>OnionGraph</i>	Autoria de artigos	●	●	
<i>Elzen e Wijk</i>	Migração e censo USA	●		●
<i>Ko et al.</i>	Atraso em voos	●		●
<i>AmbiguityVis</i>	Ambiguidade em <i>layout</i> de grafos		●	
<i>g-Miner</i>	Redes de coautoria	●	●	
<i>Wu et al.</i>	Redes de colaboração	●		
<i>egoSlider</i>	Redes de colaboração	●		

Observa-se no gráfico da Figura 2.28, referente aos dados da tabela 2.3, que mais da metade das técnicas (17=63%) possibilitam a exploração de grafos sem que estejam relacionados a uma área específica. Em relação aos atributos de arestas, percebe-se que, embora a maioria represente até três atributos, apenas 3 (11%) permite a exploração de uma maior quantidade.

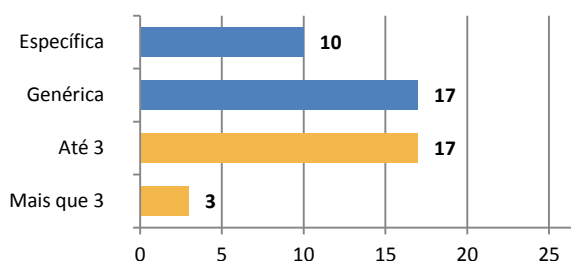


Figura 2.28 – Área de aplicação e existência de atributos nas arestas

2.3.2 Abordagens de integração da topologia e seus atributos

Há várias abordagens de representação para a visualização de grafos multivariados. Algumas exibem melhor a topologia do grafo do que os atributos associados aos nodos e

arestas e vice-versa. Nesta seção será apresentada a análise da questão QP3 (Tabela 2.4) na qual as técnicas são classificadas de acordo com os cinco critérios adotados por Jusufi et al. (2013) que as agrupam segundo a forma como as representações são construídas, levando em conta a maneira usada para integrar a estrutura do grafo a seus respectivos atributos. A seguir é apresentada uma descrição de cada abordagem.

- i) **Múltiplas visões coordenadas:** As técnicas incluídas nessa abordagem combinam várias visões para representar os dados. Essa estratégia possibilita que o usuário possa escolher a técnica mais adequada para visualizar cada conjunto de dados. Geralmente mecanismos de *brushing and linking* são incorporados de forma que alterações em determinados elementos de uma visão possam ser propagadas nas outras. Uma desvantagem dessa abordagem é que, devido à separação visual dos elementos, fica a cargo do processo de cognição do usuário a tarefa de integrá-las.
- ii) **Abordagem integrada:** Combina a representação da estrutura do grafo e seus atributos em uma única área de visualização. Essa estratégia pode afetar a legibilidade do grafo devido às sobreposições que podem ocorrer quando a quantidade de nodos e atributos é alta, fato que pode ser minimizado pelo uso de técnicas de interação. Entretanto, essa abordagem pode decrementar o tempo que os usuários necessitam para encontrar as relações em razão de que as informações são exibidas em um único lugar (GÖRG et al. 2007).
- iii) **Substrato semântico:** Essa abordagem foi inicialmente utilizada na ferramenta NVSS (SHNEIDERMAN e ARIS 2006) com o objetivo de reduzir a desordem visual na visualização de grafos multivariados. Caracteriza-se por agrupar os nodos de acordo com o valor de um de seus atributos em regiões não sobrepostas.
- iv) **Layout dirigido pelos atributos:** Essa abordagem utiliza o valor dos atributos pertencentes aos elementos do grafo para determinar a posição dos nodos de forma a priorizar a exibição da correlação entre eles. Ela possui a desvantagem de sacrificar a exibição da topologia dificultando a percepção visual das diferentes estruturas do grafo e difere da anterior (substrato semântico) em função de que os elementos não são necessariamente apresentados em múltiplas regiões.
- v) **Abordagem híbrida:** Técnicas que utilizam mais de uma abordagem entre as quatro primeiras serão classificadas nessa última.

Tabela 2.4 – Técnicas classificadas pela abordagem

Técnicas	QP3				
	Múltiplas visões coordenadas	Integrada	Substrato semântico	Layout dirigido pelos atributos	Híbrida
<i>RelVis</i>		●			
<i>PivotGraph</i>				●	
<i>NVSS</i>			●		
<i>NoodleView</i>			●		
<i>GraphSpace</i>		●		●	●
<i>AttriGraph</i>			●		
<i>Shannon et al.</i>	●				
<i>GraphDice</i>				●	
<i>PEx-Graph</i>	●			●	●
<i>P-SPLOM</i>	●			●	●
<i>NetworkLens</i>		●			
<i>Cluma-GO</i>	●				
<i>GraphTrail</i>	●				
<i>ViNCent</i>	●	●			●
<i>enRoute</i>	●				
<i>SocialNetSense</i>	●				
<i>JauntyNets</i>	●			●	●
<i>PNLB</i>		●	●		●
<i>PIWI</i>	●		●		●
<i>Jusuft et al.</i>			●		
<i>OnionGraph</i>		●			
<i>Elzen e Wijk</i>	●	●		●	●
<i>Ko et al.</i>		●			
<i>AmbiguityVis</i>	●				
<i>g-Miner</i>	●	●			●
<i>Wu et al.</i>	●				
<i>egoSlider</i>	●	●			●

O gráfico da Figura 2.29, referente aos dados da tabela 2.4, apresenta um resumo das técnicas classificadas segundo as abordagens descritas anteriormente.

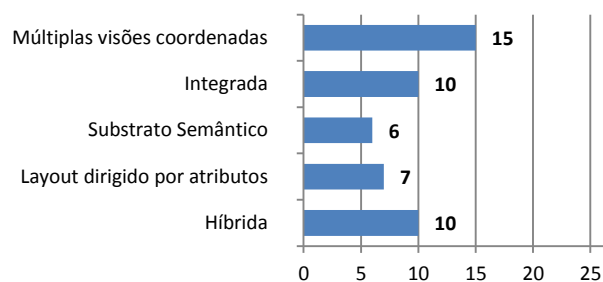


Figura 2.29 – Classificação das técnicas por abordagem

Observa-se que mais da metade das técnicas analisadas (15=56%) faz uso de múltiplas visões coordenadas que permitem o uso de diferentes representações para a visualização da topologia e dos múltiplos atributos. A abordagem integrada é usada em pouco mais de um terço das técnicas (10=37%), fato que talvez se justifique em razão de que a exibição da correlação dos atributos integrada à topologia resulte uma sobrecarga visual que irá reduzir a quantidade de nodos que poderão ser visualizados de forma simultânea. As abordagens

referentes a substrato semântico e *layout* dirigido pelos atributos são as menos utilizadas, entretanto, se considerarmos que são abordagens bastante similares, em conjunto também representam uma parcela significativa.

2.3.3 Representação visual

Na construção da representação visual das técnicas de exploração de grafos multivariados é necessário que tanto a topologia do grafo como a correlação de seus atributos seja percebido. Nessa seção, as técnicas serão classificadas de acordo com as primitivas gráficas (WARD et al. 2010) utilizadas na representação dos múltiplos atributos do grafo (QP4) e na forma como a topologia é representada (QP5). As técnicas foram classificadas segundo as primitivas gráficas descritas a seguir.

- i) **Técnicas baseadas em pontos:** Nesse tipo de representação as técnicas projetam registros de um espaço de dados n-dimensional para um espaço de visualização k-dimensional. Para cada registro, uma representação gráfica, marca ou outra entidade estética é desenhada no ponto k-dimensional. Os pontos de plotagem podem ser definidos para mostrar registros individuais ou sumarizados, e podem ser estruturados por várias técnicas de projeção. Técnicas como *scatterplots* e matrizes de *scatterplots* são exemplos que aderem a essa classificação.
- ii) **Técnicas baseadas em linhas:** São caracterizadas por unir os pontos que representam cada valor ou dimensão por uma linha reta ou curva. Essas linhas, além de reforçar o relacionamento entre os valores de dados, também facilitam a percepção de outros aspectos por meio de características das linhas como, por exemplo, inclinação, curvatura e cruzamentos. Gráficos de linhas e coordenadas paralelas são exemplos dessa classificação.
- iii) **Técnicas baseadas em regiões:** São caracterizadas pelo uso de polígonos preenchidos na representação dos valores. Os múltiplos atributos podem ser mapeados para propriedades visuais, como, por exemplo, tamanho, forma, cor entre outras. Gráficos de barras e *stacked bar chart* são exemplos de técnicas baseadas em regiões.
- iv) **Combinação de técnicas:** Enquadram-se nesse item as técnicas que combinam características das três classes anteriores (pontos, linhas e regiões). Entre as diversas alternativas destacam-se as técnicas baseadas em glifos (WARD, 2002) e orientadas a *pixels* (KEIM, 2000). Um glifo é uma representação visual de uma

informação na qual as características de uma entidade gráfica são determinadas pelo valor dos atributos contidos nessa informação. Cada item do conjunto de dados a ser visualizado é representado por um glifo com suas n dimensões. Já as orientadas a *pixels* são caracterizadas por mapear um atributo de cada instância de um conjunto de dados para um pixel colorido da área de visualização baseado no seu valor.

Em relação à forma como a topologia do grafo é representada, as técnicas foram classificadas segundo três critérios: (i) técnicas que usam o **diagrama de nodos e arestas**, independente do tipo de *layout* adotado; (ii) técnicas que utilizam uma **matriz de adjacências**; (iii) e as que utilizam **outras formas** de representação.

A tabela 2.5 apresenta o resultado da análise segundo os critérios descritos anteriormente. Na coluna referente à combinação de técnicas foi incluída uma letra (G-Glifo, P-Orientada a *pixel*) indicando o tipo de técnica utilizada.

Tabela 2.5 – Representação visual dos atributos e da topologia

Técnicas	QP4				QP5		
	Pontos	Linhas	Regiões	Combinação	Nodos-arestas	Matriz adjac.	Outras formas
<i>RelVis</i>			•		•		
<i>PivotGraph</i>	•				•		
<i>NVSS</i>	•				•		
<i>NoodleView</i>			•				•
<i>GraphSpace</i>			•		•		
<i>AttriGraph</i>			•				•
<i>Shannon et al.</i>		•			•		
<i>GraphDice</i>	•				•		
<i>PEX-Graph</i>					•		
<i>P-SPLOM</i>	•	•			•		
<i>NetworkLens</i>		•	•	G	•		
<i>Cluma-GO</i>				P			•
<i>GraphTrail</i>	•		•			•	•
<i>ViNCent</i>			•		•		
<i>enRoute</i>			•		•		
<i>SocialNetSense</i>			•		•		
<i>JauntyNets</i>			•		•		
<i>PNLB</i>		•					•
<i>PIWI</i>				P			•
<i>Jusufi et al.</i>			•		•		
<i>OnionGraph</i>			•		•		
<i>Elzen e Wijk</i>	•	•	•		•		
<i>Ko et al.</i>			•		•		
<i>AmbiguityVis</i>	•		•	P	•		
<i>g-Miner</i>			•	G	•	•	
<i>Wu et al.</i>			•				•
<i>egoSlider</i>	•		•				•

Observando-se o gráfico da Figura 2.30, percebe-se que as técnicas para visualização de dados multidimensionais baseadas em regiões são as mais utilizadas (18=67%). De fato, a grande parte utiliza histogramas ou suas variações para representar os atributos. Em relação à visualização da topologia do grafo, o diagrama de nodos e arestas é o mais usado (19=70%), provavelmente por ser o que melhor favoreça a percepção dos relacionamentos existentes entre os seus elementos.

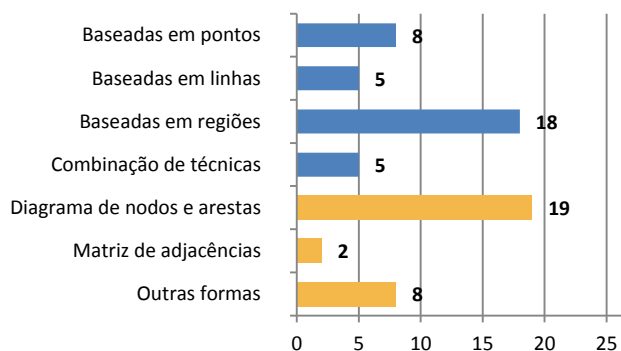


Figura 2.30 – Representação visual dos atributos e da topologia

2.3.4 Tarefas para a exploração de grafos multivariados

As tarefas descritas nesta seção são definidas como as atividades que o usuário deseja realizar por meio de interação com a representação visual de um grafo multivariado (PRETORIUS et al. 2014). Isso implica que há uma intenção do usuário (YI et al. 2007) e que o grafo foi apresentado visualmente. Vários autores têm proposto taxonomias de tarefas para a visualização de informações em geral (SHNEIDERMAN 2006, AMAR et al 2005, BREHMER; MUNZNER 2013). Algumas são focadas na avaliação de visualização multidimensional (VALIATI et al. 2006) e outras específicas para a visualização de grafos (SHNEIDERMAN;ARIS 2006, LEE et al. 2006, AHN et al. 2014). Para realizar a análise das técnicas (QP6) utilizamos como base o trabalho de Pretorius et al. (2014) que, embora seja fortemente baseado na taxonomia definida por Lee et al. (2006), é dirigido especificamente para a visualização de grafos multivariados. Com o objetivo de simplificar a análise das tarefas suportadas pelas técnicas de visualização, foram utilizados apenas dois níveis da classificação proposta, cuja descrição será apresentada a seguir.

a) **Tarefas baseadas na estrutura:** São tarefas que levam em conta a topologia do grafo para a sua realização. Foram subdivididas em três subgrupos como descrito a seguir:

- i) **Tarefas de adjacências:** São tarefas que permitem inferir conhecimento sobre a adjacência existente entre os nodos do grafo, como, por exemplo, encontrar a

quantidade de pessoas adjacentes ou os tipos de relacionamentos adjacentes a uma pessoa que possui um nome determinado.

ii) **Tarefas de acessibilidade:** Permitem inferir conhecimento sobre o fato de um nodo ser ou não acessível a partir de outro, como por exemplo, encontrar os nomes de todos os amigos que são amigos de uma determinada pessoa em uma rede social.

iii) **Tarefas de conectividade:** Permite inferir conhecimento sobre subgrafos, como por exemplo, encontrar *clusters* ou componentes conectados.

b) **Tarefas baseadas nos atributos:** Foram subdivididas em tarefas que permitem inferir conhecimento sobre atributos de nodos e áreas.

i) **Nodos:** Tarefas, como, por exemplo, encontrar os nodos que possuem um valor específico de atributo.

ii) **Arestas:** Dado um nodo, encontrar todos os nodos conectados a ele por arestas com um valor específico de atributo.

c) **Tarefas de navegação:** São aquelas que permitem ao usuário inferir conhecimento a partir da navegação nos caminhos definidos pela estrutura de um grafo multivariado.

d) **Tarefas de overview:** São tarefas relacionadas a fornecer uma compreensão da estrutura geral do grafo, como por exemplo, identificar um conjunto de nodos como potencialmente pertencentes a grupos altamente conectados.

A tabela 2.6 apresenta o resultado da análise segundo as tarefas de exploração suportadas pelas técnicas (QP6). Analisando-se o gráfico da Figura 2.31, referente aos dados da tabela 2.6, percebe-se que, das tarefas baseadas na estrutura, apenas as que permitem inferir conhecimento sobre a adjacência existente entre nodos são executadas pela maioria das técnicas (21=78%). Nove (33%) executam tarefas de conectividade e apenas três (11%) disponibilizam mecanismos para a execução de tarefas de acessibilidade. Em relação a tarefas baseadas no valor dos atributos, a totalidade está relacionada a nodos e apenas duas executam tarefas sobre as arestas: *AttriGraph* (PRETORIUS; WIJK, 2008), que enfatiza a visualização das arestas, e *GraphDice* (BEZERIANOS et al. 2010), que explora nodos e arestas de uma forma unificada. Tarefas que permitem ao usuário inferir conhecimento a partir da navegação nos caminhos definidos pela estrutura também não são suportadas pela grande maioria. Das técnicas avaliadas apenas a *enRoute* (PARTL et al. 2012) e a *egoSlider* (WU et al. 2016) as suportam. Para as tarefas de *overview*, quase a totalidade (25=93%) das técnicas as suportam.

Tabela 2.6 – Tarefas para exploração de grafos multivariados

Técnicas	QP6							Quantidade de tarefas
	Estrutura			Atributos		Navegação	Overview	
	Adjacências	Acessibilidade	Conectividade	Nodos	Arestas			
<i>enRoute</i>	●	●	●	●		●	●	6
<i>GraphDice</i>	●		●	●	●		●	5
<i>Elzen e Wijk</i>	●	●	●	●			●	5
<i>PEx-Graph</i>	●		●	●			●	4
<i>JauntyNets</i>	●		●	●			●	4
<i>PIWI</i>	●	●	●	●				4
<i>g-Miner</i>	●		●	●			●	4
<i>Wu et al.</i>	●		●	●			●	4
<i>egoSlider</i>	●			●		●	●	4
<i>NYSS</i>	●			●			●	3
<i>NoodleView</i>	●			●			●	3
<i>AttriGraph</i>	●			●	●			3
<i>Shannon et al.</i>	●			●			●	3
<i>Cluma-GO</i>	●			●			●	3
<i>GraphTrail</i>	●			●			●	3
<i>ViNCent</i>	●			●			●	3
<i>SocialNetSense</i>	●			●			●	3
<i>PNLB</i>	●			●			●	3
<i>Jusufi et al.</i>	●			●			●	3
<i>OnionGraph</i>	●			●			●	3
<i>Ko et al.</i>	●			●			●	3
<i>AmbiguityVis</i>			●	●			●	3
<i>RelVis</i>				●			●	2
<i>PivotGraph</i>				●			●	2
<i>GraphSpace</i>				●			●	2
<i>P-SPLOM</i>				●			●	2
<i>NetworkLens</i>				●			●	2

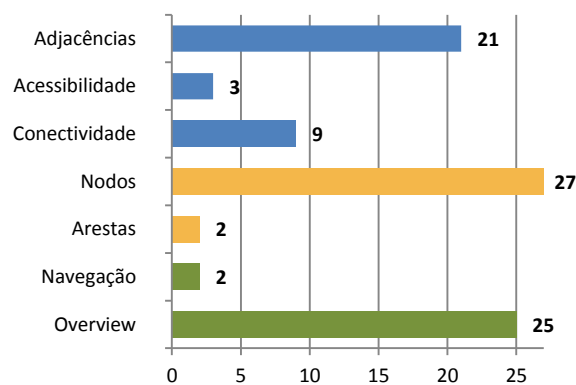


Figura 2.31 – Tarefas para exploração de grafos multivariados

2.3.5 Técnicas de interação para exploração de grafos multivariados

Para que o usuário possa executar as tarefas de exploração, é necessário que as técnicas de visualização ofereçam mecanismos de interação que ocasionem alterações na representação visual de modo que novos aspectos do conjunto de dados possam ser

observados. Nesta seção as técnicas são analisadas (QP7) de acordo com a abordagem proposta em (WYBROW et al. 2014), que utiliza o modelo de referência (CARD et al, 1999a) para dividir os mecanismos de interação suportados pela visualização de grafos multivariados em três grupos, conforme o nível do pipeline que eles afetam: visões, estruturas visuais e dados.

a) **Interação ao nível das visões:** Neste grupo os mecanismos estão relacionados ao destaque dos elementos de interesse e à navegação através do conjunto de dados.

- i) ***Hovering***: Usado para destacar elementos relacionados ou exibir informações adicionais quando o mouse passa sobre eles.
- ii) ***Brushing and linking (B&L)***: Usado com múltiplas visões coordenadas, permitindo que um determinado elemento apontado em uma área de visualização seja destacado em outra.
- iii) ***Magic lenses***: Funcionam como uma espécie de filtro que modifica a apresentação dos elementos, em uma região de foco, de forma a revelar novas informações ou suprimir elementos de distração.
- iv) ***Panning and zooming (P&Z)***: Permitem o reposicionamento do observador com o objetivo de alterar a parte do conjunto de dados a ser exibida na área de visualização ou proporcionar uma visão mais detalhada da mesma.
- v) ***View distortion***: São mecanismos que permitem a exibição mais detalhada dos elementos que estão posicionados no foco. Podem ser aplicados tanto aos nodos como às arestas.

b) **Interação ao nível das estruturas visuais:** Neste grupo os mecanismos estão relacionados à alteração das propriedades visuais e à localização dos elementos gráficos que compõem a representação utilizada nas técnicas de visualização.

- i) **Seleção**: Modifica as propriedades visuais dos elementos com o objetivo de destacar determinadas áreas do grafo. Pode ser aplicada por meio de um apontamento direto aos elementos a serem destacados, pela seleção de nodos baseadas na estrutura do grafo ou pelo valor de seus atributos.
- ii) **Mapeamento de atributos**: Mecanismos que permitem a alteração do mapeamento de determinados atributos para as propriedades visuais com o objetivo de explorar as várias dimensões dos dados.

- iii) **Layout:** Mecanismos que permitem a alteração da posição de nodos e arestas baseada nas propriedades do grafo. Exemplos desse tipo de interação incluem o posicionamento de nodos para enfatizar similaridade, (como, por exemplo, por meio do uso de técnicas de projeções de dados multidimensionais), pela aplicação de algoritmos de *layout* de grafos automáticos, ou pela interação direta (arrasto) com os elementos do grafo.
 - iv) **Representação:** Mecanismos que permitem a alteração da representação utilizada, como, por exemplo, substituir um diagrama de nodos e arestas por uma matriz de adjacências.
- c) **Interação ao nível de dados:** Está relacionada à seleção de quais dados deverão ser mostrados.
- i) **Filtragem:** Interação que permite definir qual subconjunto de dados deve ser exibido conforme o interesse do usuário. Pode ser executada diretamente sobre a representação visual ou pelo uso de um componente de interface, como, por exemplo, *sliders*.
 - ii) **Busca:** Permitem que um subconjunto de interesse seja destacado do conjunto inteiro de dados. Pode ser usada para buscar nodos/arestas com determinados valores de atributos ou com certas propriedades topológicas. Sua execução pode ser realizada a partir de consultas textuais.
 - iii) **Edição:** Mecanismo para permitir que elementos da estrutura visualizada sejam alterados, como por exemplo, a exclusão/adição de nodos ou arestas.
 - iv) **Agregação:** Permite que vários nodos/arestas sejam agrupados dando origem a supernodos/superarestas com o objetivo de simplificar o grafo. A agregação pode ser realizada a partir da escolha de estratégias pré-definidas ou definidas interativamente pelo usuário.
 - v) **Anotação:** Mecanismo para permitir que o usuário possa acrescentar informações adicionais aos elementos para marcar pontos de interesse.
 - vi) **Histórico:** Mecanismo usado para registrar o histórico do processo de exploração.

A tabela 2.7 apresenta o resultado da análise segundo as técnicas de interação disponíveis nas ferramentas de exploração (QP7).

Tabela 2.7 - Técnicas de interação para exploração de grafos multivariados

Técnicas	QP7														Qde. Técnicas de interação	
	Visões					Estruturas visuais				Dados						
	<i>Hovering</i>	<i>B & L</i>	<i>Magic Lenses</i>	<i>P & Z</i>	<i>View Distortion</i>	<i>Seleção</i>	<i>Map. atributos</i>	<i>Layout</i>	<i>Representação</i>	<i>Filtragem</i>	<i>Busca</i>	<i>Edição</i>	<i>Agregação</i>	<i>Anotação</i>		<i>Histórico</i>
<i>GraphDice</i>	●	●	●	●		●	●	●		●	●				●	10
<i>SocialNetSense</i>	●	●		●		●	●	●		●	●		●	●	●	10
<i>NetworkLens</i>	●		●	●	●	●	●	●	●	●	●					9
<i>Elzen e Wijk</i>	●	●		●		●	●	●	●	●	●		●			9
<i>PEx-Graph</i>	●	●		●		●	●	●		●	●					8
<i>Cluma-GO</i>	●	●		●		●		●		●	●		●			8
<i>GraphTrail</i>	●	●		●		●		●	●	●	●		●		●	8
<i>g-Miner</i>	●			●		●	●	●	●	●	●		●			8
<i>enRoute</i>	●	●		●		●			●	●	●		●			7
<i>JauntyNets</i>	●	●		●		●		●	●	●	●					7
<i>PNLB</i>	●	●		●	●	●			●	●	●					7
<i>OnionGraph</i>	●	●		●		●	●	●	●	●	●		●			7
<i>egoSlider</i>	●	●		●		●	●	●	●	●	●					7
<i>NVSS</i>	●	●				●	●	●	●	●	●		●			6
<i>NoodleView</i>	●			●	●	●			●	●	●		●			6
<i>AttriGraph</i>	●					●			●	●	●		●		●	6
<i>Jusufi et al.</i>	●	●				●		●	●	●	●		●			6
<i>GraphSpace</i>			●	●		●	●	●								5
<i>P-SPLOM</i>	●	●		●		●		●								5
<i>ViNCent</i>	●	●				●		●		●						5
<i>PIWI</i>		●		●		●			●	●	●					5
<i>Ko et al.</i>	●	●		●		●			●	●	●					5
<i>Shannon et al.</i>	●	●				●			●	●	●					4
<i>PivotGraph</i>	●					●			●	●	●		●			4
<i>AmbiguityVis</i>		●		●		●			●	●	●					4
<i>Wu et al.</i>						●		●					●			3
<i>RelVis</i>									●							1

Analisando-se o gráfico da Figura 2.32 baseada nos dados da tabela 2.7, observa-se que no grupo referente ao nível de visão (Figura 2.32), técnicas como *hovering*, *panning and zooming* e *brushing and linking* são bastante utilizadas. Esta última, principalmente pelo fato de que inúmeras técnicas fazem uso de visões coordenadas como mencionado na seção 2.3.2. Técnicas como *magic lenses* e *view distortion* são menos utilizadas.

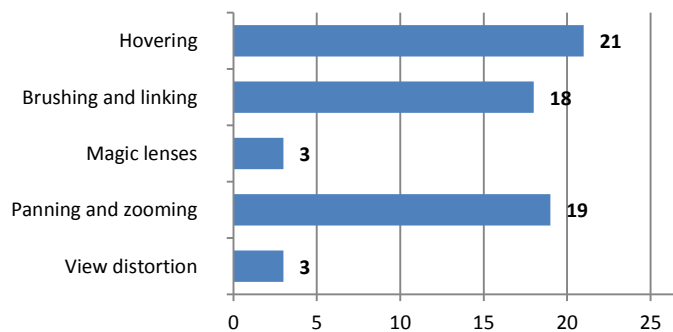


Figura 2.32 – Técnicas de interação - Visões

Em nível das estruturas visuais (Figura 2.33), mecanismos de seleção são utilizados em quase a totalidade das técnicas (26=96%). Isso se justifica pelo fato de que é necessário o destaque de um grupo de elementos para executar a sua exploração mais detalhada. Mecanismos para alteração de *layout* são disponibilizados pela metade das técnicas (14=52%) e o recurso de alteração do mapeamento de atributos e representação é pouco utilizado.

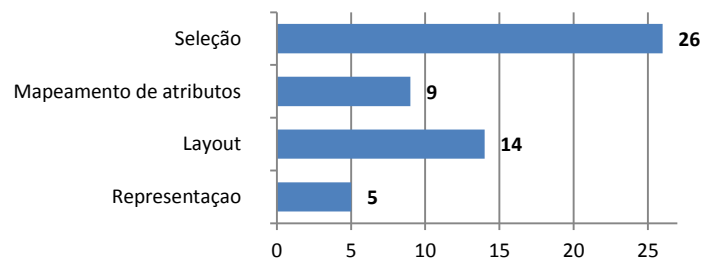


Figura 2.33 – Técnicas de interação – Estruturas visuais

Em relação ao último grupo (Figura 2.34), que permite estabelecer quais dados serão exibidos, observa-se que as técnicas de filtragem são as mais usadas (24=89%), seguidas pela busca e agregação. Esse fato justifica-se em função de serem eficientes para reduzir a quantidade de elementos explorados em grandes estruturas. Mecanismos que permitem a edição e inclusão de informações adicionais (anotações) aos elementos representados são pouco utilizados, assim como aqueles que permitem registrar o histórico da exploração.

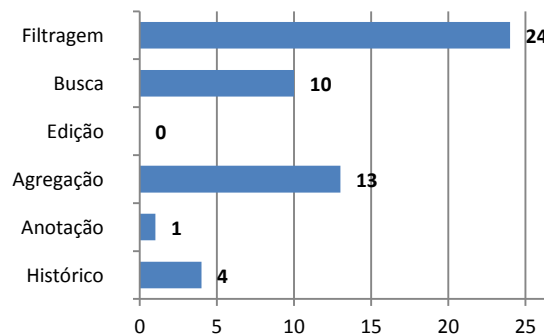


Figura 2.34 – Técnicas de interação – Dados

2.4 Comentários finais

Nesta seção foram descritas técnicas de visualização para grafos multivariados desenvolvidas nos últimos doze anos, apresentando um resumo de suas principais características relacionadas à representação visual, tarefas de exploração suportadas e técnicas de interação. Observando-se o gráfico da Figura 2.35 percebe-se que das técnicas selecionadas para análise 7 foram desenvolvidas nos primeiros cinco anos desse período. Já

no restante do período houve um incremento de publicações, o que demonstra o crescente interesse da comunidade científica pela investigação desse tipo de visualização.

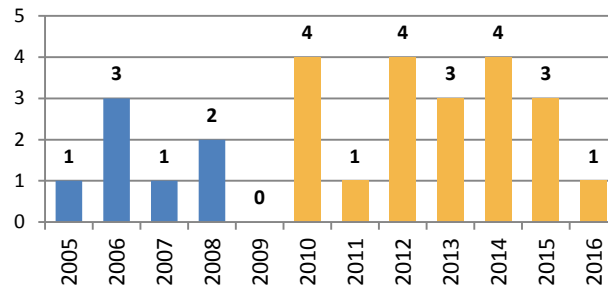


Figura 2.35 – Desenvolvimento das técnicas no período de 2005 a 2016

Para identificar a importância de cada técnica, considerando apenas as selecionadas para compor esse trabalho, foi construído um diagrama (Figura 2.36) que representa a relação existente entre elas, baseada nas citações de suas referências bibliográficas. Nele observa-se que há um destaque no ano de 2006 no qual surgiram as técnicas *PivotGraph* (WATTENBERG, 2006) e *NVSS* (SHNEIDERMAN; ARIS, 2006) já que foram citadas pela maioria das técnicas analisadas. De fato elas são pioneiras no tema e propõem a exibição de um grafo integrado a uma espécie de *scatterplot*, um gráfico até então utilizado tipicamente em visualização de dados multidimensionais. Percebe-se também um destaque da técnica *GraphDice* (Bezerianos et al. 2010) que, embora recente, recebeu 12 citações das 19 publicações no período compreendido entre 2010 e 2016. Ela é uma ferramenta com inúmeros recursos e tem um enfoque bastante acentuado na visualização da relação entre os atributos do grafo. O mecanismo de consultas visuais incorporado à ferramenta permite que o usuário filtre interativamente o conjunto de dados e a manutenção de um histórico dessas consultas possibilita ao usuário refazer os passos do processo de exploração quando necessário.

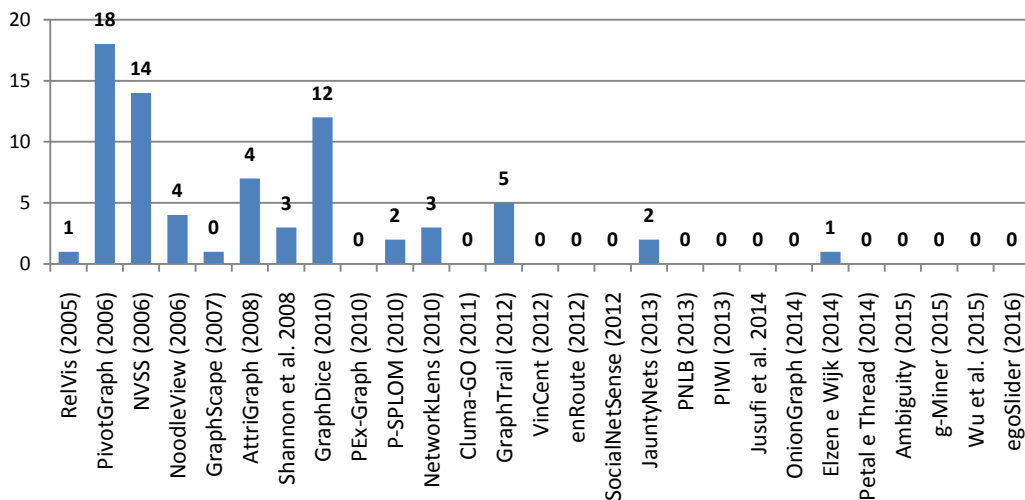


Figura 2.36 – Quantidade de citações entre as técnicas selecionadas

Considerando-se as tarefas específicas para a exploração de grafos multivariados, percebe-se que a maioria das técnicas não as suportam. Apenas tarefas que permitem inferir conhecimento sobre os atributos dos nodos, suas adjacências e visão geral do grafo são contempladas. Entretanto, cabe destacar que as técnicas *enRoute* (PARTL et al. 2012), *GraphDice* (BEZERIANOS et al. 2010) e a proposta por Elzen e Wijk (2014) são as que executam a maior parte das tarefas selecionadas para análise. Assim, a criação de mecanismos específicos para a exploração em grafos multivariados, tema ainda pouco explorado, pode ser um interessante problema a ser investigado.

Os mecanismos de navegação e exploração das estruturas são elementos fundamentais em todas as técnicas apresentadas. Seu grau de importância cresce à medida que a quantidade de dados a serem visualizados aumenta. Em função da quantidade de nodos (e seus atributos) exibidos, pode não ser possível fazer uma interpretação adequada da estrutura, exigindo que haja alguma forma de acesso aos seus detalhes. Dentre as técnicas apresentadas *GraphDice* (BEZERIANOS et al. 2010), *SocialNetSense* (GOU et al. 2012), *Network Lens* (JUSUFI et al. 2010) e a proposta por Elzen e Wijk (2014), destacam-se nesse quesito com uma quantidade significativa de mecanismos que facilitam a exploração visual.

3 EXPLORAÇÃO INTERATIVA DE GRAFOS MULTIVARIADOS

Encontrar uma forma adequada para facilitar a compreensão de um grafo multivariado torna-se um grande desafio em razão da sua complexidade. A exploração das informações representadas nesse tipo de estrutura exige uma análise integrada de dois aspectos distintos: a topologia, que define as relações existentes entre seus elementos, e o conjunto de atributos associados a eles. Assim, é necessário alcançar um equilíbrio entre mostrar o máximo possível de informação, para minimizar a necessidade de navegação, e o custo de mostrar uma quantidade inadequada de uma única vez, na qual o usuário seria sobrecarregado pela desordem visual.

Na última década, várias técnicas/ferramentas para visualização de grafos multivariados foram desenvolvidas. Entretanto, a análise apresentada no Capítulo 2 mostra que existem ainda aspectos que poderiam ser investigados. Nesse sentido, devido à complexidade dos grafos multivariados presume-se ser improvável encontrar uma única combinação de técnicas que apoie plenamente a exploração de tais conjuntos de dados. Assim, a solução sugerida neste trabalho fornece ao usuário diversas técnicas para serem usadas de forma combinada na exploração de grafos multivariados. Essa abordagem coloca em prática o *framework* “*what-why-how*”, proposto por Munzner (2014), que se baseia nas seguintes questões:

- i) Quais (*what*) dados os usuários veem?
- ii) Por que (*why*) o usuário pretende usar a ferramenta de visualização?
- iii) Como (*how*) o mapeamento e interação são construídos em termos de opções de projeto?

Cada combinação dessas três questões define uma instância de visualização, estando elas associadas respectivamente a: (i) Quais dados serão visualizados? (ii) Quais tarefas que serão executadas? (iii) Qual a representação visual e técnicas de interação serão adotadas? Embora ferramentas simples possam ser concebidas com uma única instância, sistemas complexos podem exigir uma sequência de instâncias expressando dependências entre elas. Pela complexidade inerente aos grafos multivariados, a abordagem proposta adota o uso das múltiplas instâncias, mas sem definir uma ordem específica, já que o usuário é o responsável pela escolha de cada uma delas. Além disso, o *framework* é usado em um sentido mais amplo por manter o registro de quais ações de exploração o usuário executa em busca da compreensão dos dados, e quais técnicas de visualização foram empregadas na exploração.

Para organizar o processo de *design* e validação das técnicas propostas neste trabalho, foi utilizado um modelo proposto por Munzner (2009, 2014), denominado *Nested Model*.

Nele, o processo de *design* é dividido em quatro níveis aninhados de forma que a saída de um torna-se entrada do outro. A Figura 3.1 apresenta um esquema do modelo, composto pelos seguintes níveis: (i) caracterização do domínio do problema, (ii) *design* da abstração dos dados e tarefas, (iii) *design* das técnicas de interação e da codificação visual e (iv) *design* dos algoritmos.

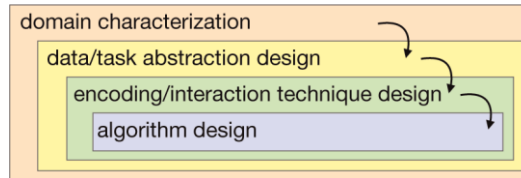


Figura 3.1 – Modelo para *design* e validação de ferramentas de visualização

Fonte: Munzner (2009)

O primeiro nível refere-se à caracterização do domínio da aplicação de forma a determinar quem são os usuários alvo. Embora este trabalho apresente um estudo de caso utilizando um domínio específico (redes de coautorias), espera-se que as ideias apresentadas possam ser adotadas para o desenvolvimento de qualquer aplicação que tenha como objetivo a exploração de grafos multivariados.

No segundo nível, os problemas e dados devem ser mapeados do domínio específico para uma descrição mais abstrata e genérica dentro do vocabulário da visualização de informações. Essa abstração de dados está relacionada à questão “Quais (*what*) dados os usuários veem?” mencionada anteriormente. Assim, foram adotadas na representação dos dados estruturas como grafos e tabelas. Neste mesmo nível, deve ser considerado também o conjunto de tarefas que serão utilizadas no próximo nível para orientar a definição das técnicas de interação. A determinação dessas tarefas responde à questão: “Por que (*why*) o usuário pretende usar a ferramenta de visualização?”

O terceiro nível do processo de *design* é responsável pela determinação da forma a ser utilizada pela ferramenta para criar e manipular dinamicamente a representação visual gerada. Essa determinação está relacionada à questão “Como (*how*) o mapeamento e interação são construídos em termos de opções de projeto?”. Nesse nível, é definida uma codificação visual que controla exatamente como os dados serão percebidos assim como é estabelecido um conjunto de técnicas de interação que permitam a execução do conjunto de tarefas determinadas no nível anterior do processo de *design*. Apresentadas no capítulo 4, três técnicas foram propostas neste trabalho: *ClusterVis*, *GlyphMatrix* e *Iris*.

Por fim, o nível mais interno está relacionado ao projeto e implementação dos algoritmos que permitam à máquina executar eficientemente as interações projetadas para o

sistema. Neste trabalho não houve uma preocupação em encontrar algoritmos eficientes já que o objetivo da implementação foi apenas viabilizar a prova de conceito e não gerar uma ferramenta plenamente funcional que possa ser utilizadas pelos usuários finais.

3.1 Premissas para a exploração interativa de grafos multivariados

A abordagem proposta tem por objetivo proporcionar uma integração de diferentes técnicas de visualização para apoiar a exploração de grafos multivariados. Mais especificamente, ela foi concebida para permitir:

- i) Integração de múltiplas instâncias de visualização;
- ii) Exibição de visões simultâneas de grafos multivariados;
- iii) Representação do encadeamento entre as visões utilizadas na exploração.

A partir da investigação apresentada no capítulo anterior, foram identificadas três premissas empregadas na visualização de grafos multivariados:

- 1) Durante o processo de exploração de um grafo multivariado pode ser necessário que os usuários façam uso de diferentes instâncias de visualização.

Considerando as diversas necessidades do usuário, é razoável esperar que possa acontecer que uma única instância de visualização não atenda a todas as situações. Pode ser preciso que múltiplas instâncias sejam usadas durante a exploração, permitindo ao usuário diversas combinações entre os dados, tarefas e técnicas de visualização (Figura 3.2).

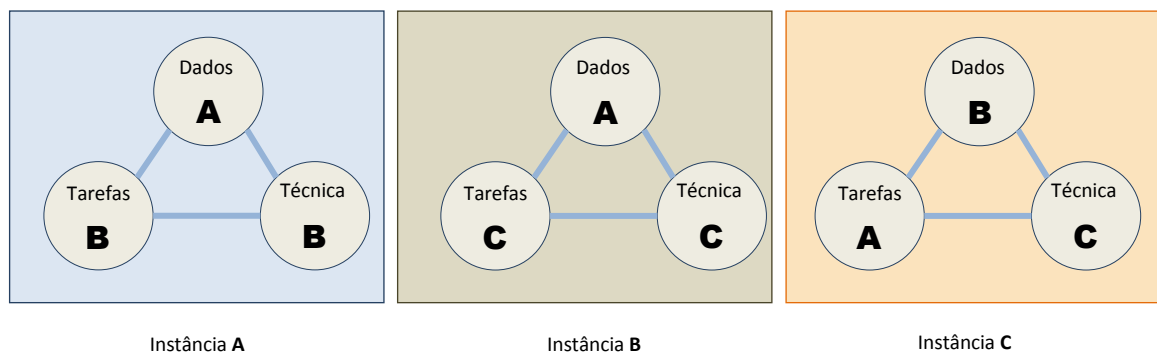


Figura 3.2 – Múltiplas instâncias de visualização

Por exemplo, na análise da topologia do grafo é comum a utilização de um diagrama de nodos e arestas, como na *Network Lens* (JUSUFI et al., 2010). Já na análise de múltiplos atributos dos nodos pode-se usar um diagrama de coordenadas paralelas como Viau et al. (2010). Apesar de existirem alternativas que combinam *scatterplots* (BEZERIANOS et al., 2010; SNHEIDERMAN e ARIS, 2006; VIAU et al., 2010) e diagrama de nodos e arestas na

mesma representação, cabe salientar que essa estratégia prioriza a visualização da correlação dos atributos em detrimento da topologia. Outras ferramentas, como *SocialNetSense* (GOU et al. 2012) e *enRoute* (PARTL et al., 2012), disponibilizam diferentes técnicas de visualização. *SocialNetSense* utiliza um diagrama de nodos e arestas para exibir a topologia de redes sociais, um diagrama de árvore para representar a estrutura social dos atores da rede, e a técnica *TreeNetVis* (GOU; ZHANG, 2011) para mostrar a estrutura híbrida formada pelo grafo e a hierarquia na mesma área de visualização. *enRoute* é uma ferramenta projetada para a exploração de dados experimentais em processos biomoleculares extraídos de rotas (*pathways*) metabólicas e que também utiliza diagrama de nodos e arestas para representar informações topológicas. No entanto, emprega uma matriz de gráfico de barras para visualizar os atributos relacionados aos dados experimentais.

- 2) Os itens de informação armazenados em grafos multivariados não podem ser facilmente exibidos todos de uma única vez. Portanto, um processo exploratório deve permitir que os usuários inspecionem e comparem subconjuntos de dados por meio de visões interligadas.

A exploração de grafos multivariados pode exigir a análise de vários subconjuntos de dados que são melhor inspecionados pelo uso de diferentes instâncias de visualização. Nos casos em que essa exploração gera resultados parciais, isto é, um subconjunto de dados obtido a partir de outra instância, é recomendável que haja uma representação da origem de cada etapa do processo (Figura 3.3).

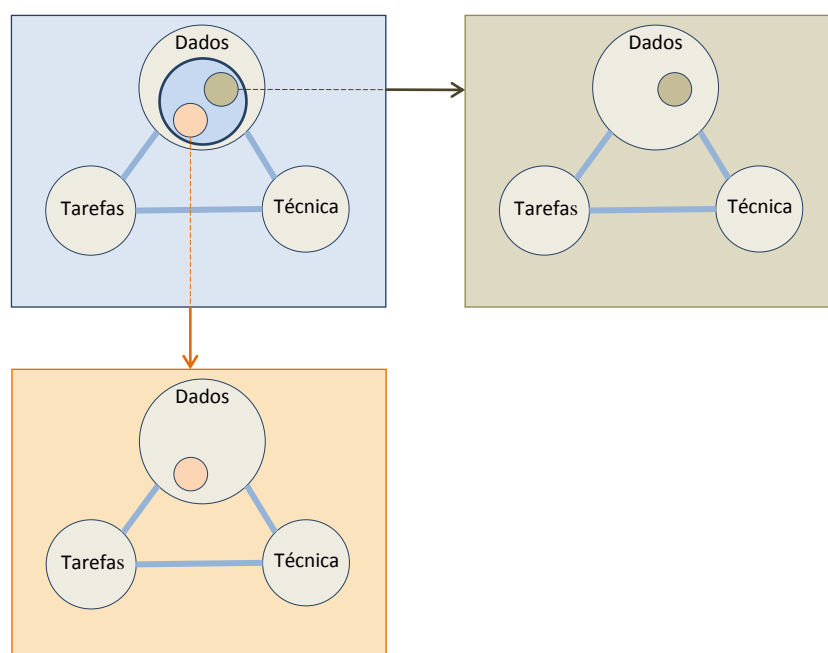


Figura 3.3 – Instâncias de visualização interligadas

Dessa forma, a criação de múltiplas visões dos subconjuntos pode implicar que elas de alguma forma estarão inter-relacionadas. Embora a representação dessa dependência nem sempre seja obrigatória, o uso de visões interligadas ajuda a compreender o processo de exploração.

O princípio de visões interligadas é implementado pela ferramenta *SocialNetSense* (GOU et al., 2012), que possui um editor gráfico no qual estruturas visuais obtidas durante o processo de exploração podem ser registradas como *snapshots* que podem ser conectados manualmente pelo usuário através de linhas, permitindo lembrar os passos seguidos durante o processo de exploração. A criação de várias visões durante o processo de exploração é também suportada pela ferramenta *GraphTrail* (DUNNE et al., 2012), que desenha automaticamente as ligações que mostram as dependências entre elas.

- 3) A exploração de grafos multivariados não exige apenas a exploração de nodos, mas também pode exigir a inspeção de arestas, subgrafos e relacionamentos criados e função do valor dos atributos de nodos e arestas.

A maioria das técnicas disponíveis para a visualização de grafos multivariados enfatiza a exploração de atributos de nodos. Embora algumas disponibilizem alternativas para a visualização de atributos de arestas, geralmente limitam os seus mapeamentos visuais ao uso da cor e espessura da linha para representar atributos. Com isso, essas representações não são suficientes para explorar as relações de um conjunto maior de atributos já que restringem a visualização a apenas dois atributos. Portanto, torna-se necessário que técnicas para exploração de múltiplos atributos de arestas sejam desenvolvidas.

Outro aspecto a ser considerado é a exploração de subgrafos. Uma forma de minimizar o problema gerado pela sobreposição de nodos e pelos cruzamentos de arestas em grafos com uma grande quantidade de nodos é o uso de estratégias com filtragem e agregação, gerando uma representação simplificada. Essas operações geram subconjuntos de dados que poderiam ser melhor explorados com técnicas de visualização específicas, diferentes. Além disso, outros mecanismos podem gerar novas estruturas levando em conta os valores dos atributos dos nodos ou arestas como será ilustrado no exemplo descrito a seguir.

Dado um conjunto de três tabelas (Figura 3.4) que armazenam respectivamente autores, artigos e a relação de coautoria, pode-se gerar três grafos distintos, levando-se em conta os valores dos atributos. Em todos eles, os nodos representam autores, sendo que, no primeiro (Figura 3.5a) as arestas representam a existência de artigos em coautoria (com atributo igual à quantidade de artigos), no segundo (Figura 3.5b), a existência de artigos

publicados no mesmo evento (com atributo igual à quantidade de eventos) e no terceiro (Figura 3.5c), artigos em coautoria publicados no VAST. Percebe-se então que no segundo duas novas arestas surgem (**Bia-Ana** e **Eva-Lara**), mesmo que não tenham publicações em coautoria. No terceiro as arestas **Eva-Ana** e **Ana-Lara** são removidas.

Autores		Coautorias		Artigos	
Autor		Autor	Artigo	Artigo	Evento
Ana		Ana	A5	A1	VAST
Bia		Ana	A6	A2	VAST
Eva		Bia	A1	A3	VAST
Lara		Bia	A2	A4	AVI
		Bia	A3	A5	AVI
		Bia	A4		
		Eva	A1		
		Eva	A2		
		Eva	A6		
		Lara	A3		
		Lara	A4		
		Lara	A5		

Figura 3.4 – Tabelas de autores, artigos e relação de coautoria

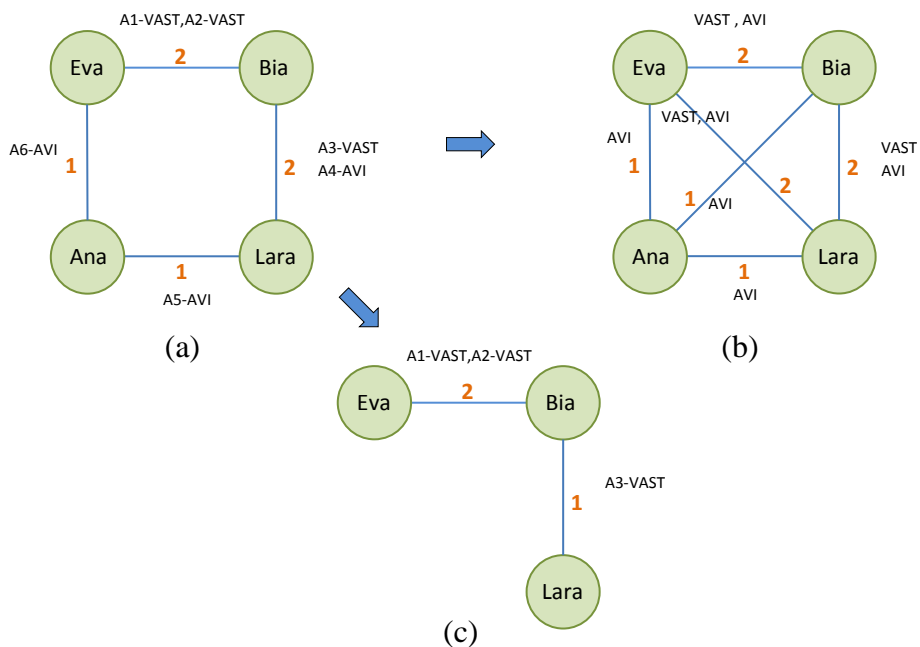


Figura 3.5 – Grafos: (a) Coautoria (b) Eventos em comum (c) coautoria no VAST

3.2 Requisitos para aderência à abordagem

Nesta seção serão descritos os requisitos mínimos (R) que deveriam ser atendidos por uma ferramenta construída levando em conta a abordagem proposta. Tais requisitos foram obtidos a partir das premissas apresentadas no início desse capítulo e estão distribuídos da seguinte forma:

Premissa 1

- R1 Fornecer **múltiplas técnicas** de visualização, de forma que a melhor alternativa possa ser escolhida, considerando as diversas necessidades do usuário.
- R2 Permitir que **múltiplas instâncias** de uma mesma técnica possam ser utilizadas.

Premissa 2

- R3 Possibilitar o uso de **múltiplas visões** (ou áreas de visualização) simultâneas, coordenadas ou não, de forma que um ou vários subconjuntos de dados possam ser explorados sob vários aspectos.
- R4 Permitir a exibição simultânea de uma **quantidade variável de visões**.
- R5 Autorizar o **reposicionamento e/ou redimensionamento** das visões no espaço de visualização de forma que o usuário possa escolher as dimensões e posições mais adequadas durante a execução da exploração.
- R6 Representar as **ligações explícitas** entre as visões através de um elemento visual, de tal modo que a sequência de exploração utilizada possa ser identificada.
- R7 Facilitar a **ocultação/exibição** de visões de forma a proporcionar uma melhor organização do espaço de visualização.
- R8 Representar o **histórico** do processo de exploração.

Premissa 3

- R9 Proporcionar a exploração de múltiplos **atributos de nodos**.
- R10 Proporcionar a exploração de múltiplos **atributos de arestas**.
- R11 Permitir que **subgrafos** possam ser selecionados para a exploração de seus atributos.

Com o objetivo de verificar se as ferramentas de visualização para grafos multivariados que usam as técnicas descritas no Capítulo 2 atendem (ou não) cada um dos

requisitos foi realizada uma análise cujos resultados serão apresentados a seguir. A Tabela 3.1 mostra o conjunto de ferramentas classificadas em ordem decrescente de quantidade de requisitos atendidos.

Tabela 3.1 – Requisitos atendidos pelas ferramentas

Ferramentas	Premissa 1		Premissa 2						Premissa 3			Qde. Requisitos atendidos
	Qde. Técnicas	Múltiplas Instâncias	Múltiplas visões	Quantidade variável	Reposicionamento	Ligações explícitas	Ocultação de visões	Histórico	Atributos nodos	Atributos arestas	Exploração subgrafos	
	R1	R2	R3	R4	R5	R6	R7	R8	R9	R10	R11	
2012 <i>SocialNetSense</i>	● (5)	●	●	●	●	●		●	●	●	●	9
2012 <i>GraphTrail</i>	● (3)	●	●	●	●	●			●		●	8
2013 <i>PIWI</i>	● (2)	●	●	●			●		●		●	7
2010 <i>GraphDice</i>	● (3)		●		●			●	●	●	●	6
2015 <i>Ambiguity</i>	● (4)	●	●						●	●	●	6
2015 <i>g-Miner</i>	● (4)		●		●			●	●		●	6
2012 <i>enRoute</i>	● (2)		●		●			●	●		●	6
2010 <i>PEx-Graph</i>	(1)	●	●	●	●			●	●			5
2008 <i>AttriGraph</i>	(1)	●	●	●				●	●	●		5
2016 <i>egoSlider</i>	● (5)	●	●				●		●		●	5
2014 <i>Ko et al.</i>	● (7)	●	●						●	●		5
2014 <i>Elzen e Wijk</i>	● (6)		●						●	●	●	5
2015 <i>Wu et al.</i>	● (5)		●						●		●	4
2013 <i>MMGraph (PNLB)</i>	● (3)		●					●	●			4
2012 <i>ViNCent</i>	● (2)				●			●	●			4
2006 <i>NVSS</i>	(1)	●	●						●		●	3
2014 <i>OnionGraph</i>	● (3)								●		●	3
2014 <i>Jusufti et al.</i>	● (2)								●		●	3
2013 <i>JauntyNets</i>	● (2)		●						●			3
2011 <i>Cluma-GO</i>	● (2)		●						●			3
2010 <i>P-SPLOM</i>	● (3)		●						●			3
2008 <i>Shannon et al.</i>	● (2)		●						●			3
2010 <i>NetworkLens</i>	● (4)								●			2
2006 <i>NoodleView</i>	● (2)								●			2
2007 <i>GraphSpace</i>	(1)								●			1
2006 <i>PivotGraph</i>	(1)								●			1
2005 <i>RelVis</i>	(1)								●			1

Os gráficos contidos no restante desta seção ilustram as principais características observadas. Como se pode constatar no gráfico da Figura 3.6 não há uma única ferramenta dentre as avaliadas que atende todos os requisitos propostos. Apenas duas contemplam mais de 70% (8 e 9) deles e a maior parte (18) se localiza no intervalo de 3 a 6 requisitos atendidos.

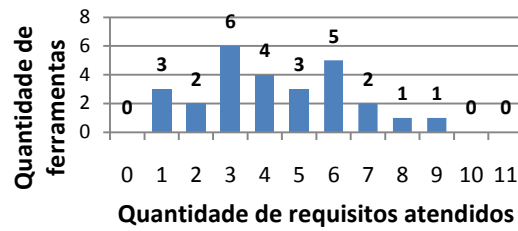


Figura 3.6 – Quantidade de ferramentas x quantidade de requisitos atendidos

O gráfico da Figura 3.7 mostra a quantidade de requisitos atendidos por ferramenta. *GraphTrail* (DUNNE et al., 2012) e *SocialNetSense* (GOU et al., 2012), com barras destacadas na cor verde, são as duas ferramentas que contemplam a maior parte dos requisitos (8 e 9 respectivamente). No entanto, *GraphTrail*, além de bloquear alterações nas visões intermediárias e impossibilitar a sua ocultação, está mais direcionada à visualização de dados agregados. Ela não inclui um diagrama de nodos e arestas tradicional, o que afeta o suporte a tarefas que exigem a compreensão da topologia do grafo. Já na *SocialNetSense*, as diversas visões que podem ser instanciadas comportam-se como *snapshots*, nos quais não há possibilidade de interação com seu conteúdo. Uma vez instanciadas, elas apenas podem ser reposicionadas na área de visualização. Para alterá-las é necessário acessar outro módulo da ferramenta destinado à execução das tarefas de exploração.

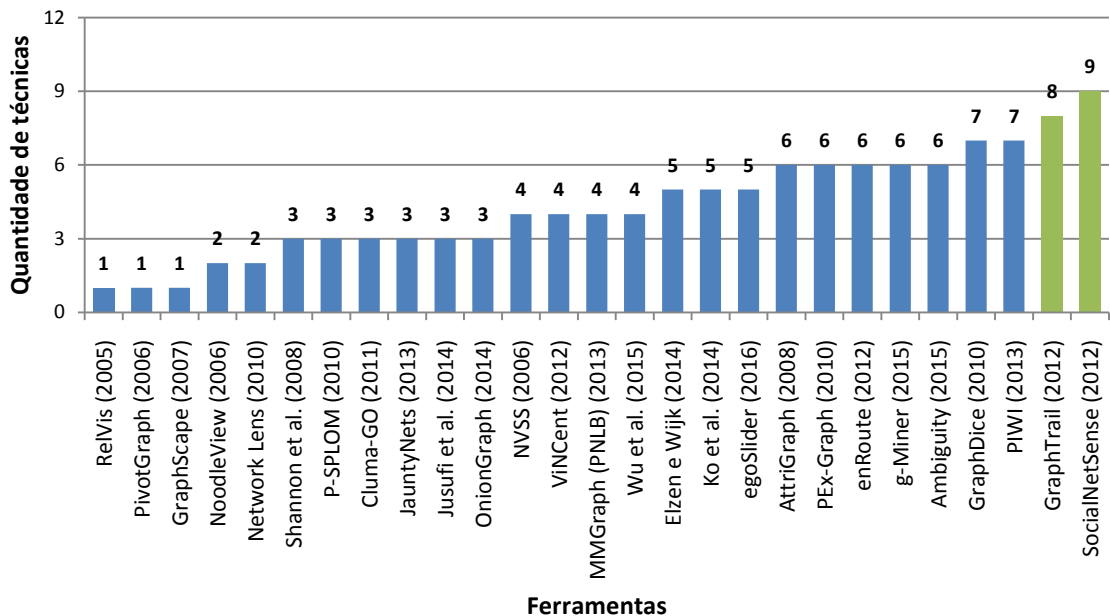


Figura 3.7 – Quantidade de requisitos atendidos em cada ferramenta

Os gráficos da Figura 3.8 ilustram a aderência das ferramentas a cada uma das premissas. Os requisitos são agrupados nesses gráficos, levando-se em conta as premissas às

quais eles pertencem. Observa-se que 24 ferramentas (85%) satisfazem, pelo menos de forma parcial, a premissa 1 (barras azuis), enquanto que apenas 6 (22%) delas atendem em sua totalidade. A premissa 2 (barras verdes) é a que possui menor quantidade de requisitos atendidos. Embora ela seja composta de 6 requisitos, dos quais 4 são frequentemente implementados em sistemas que utilizam um gerenciador de janelas, 7 técnicas (26%) não contemplam nenhum deles. Também não foi registrada uma única técnica que atenda a premissa em sua totalidade. E, por fim, a premissa 3 (barras marrons) aparece contemplada plenamente por apenas 3 (11%) técnicas. No entanto, pelo menos um dos requisitos é atendido por todas as técnicas.

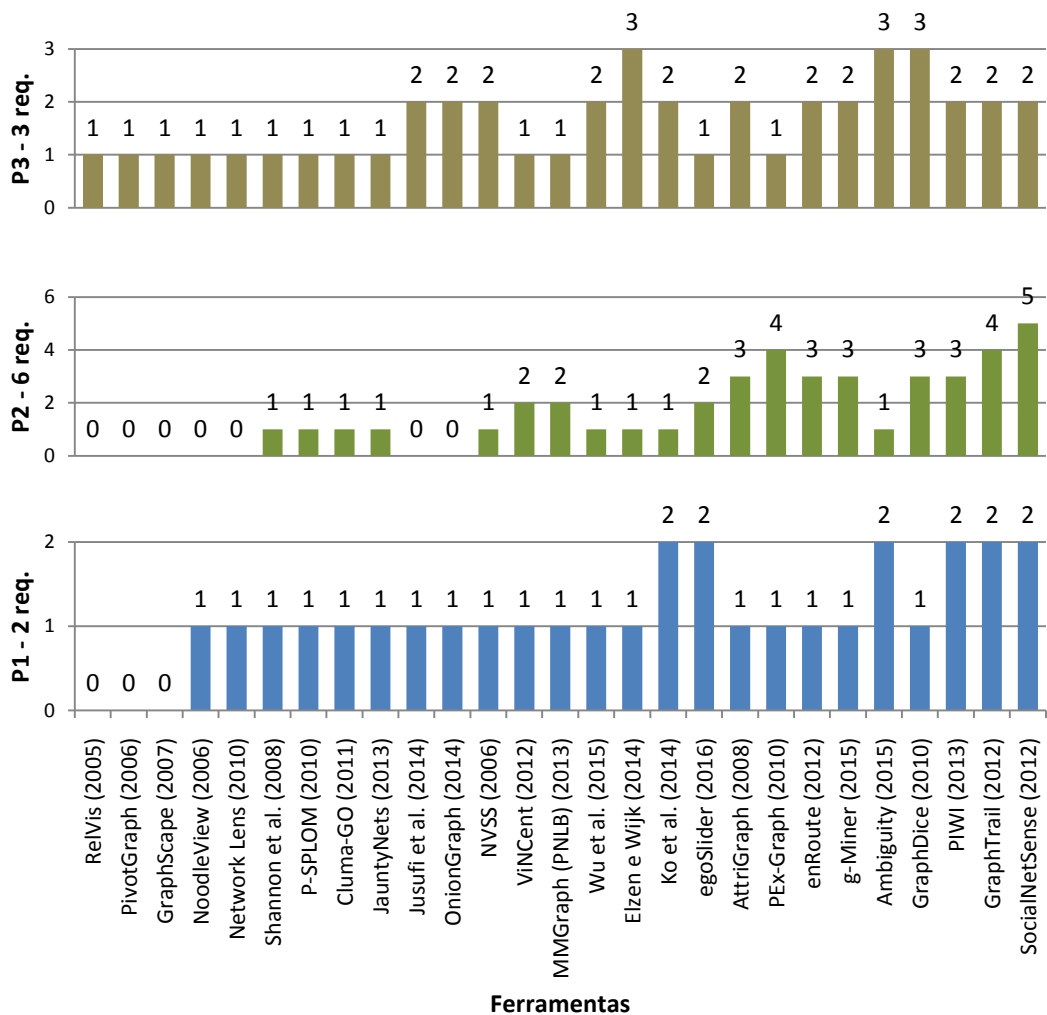


Figura 3.8 – Premissas atendidas em cada ferramenta

O gráfico da Figura 3.9 mostra a aderência das técnicas a cada um dos 11 requisitos. Pode-se observar que todas elas implementam a exploração de múltiplos atributos de nodos (R9), fato já esperado. Outro ponto observado é que a maioria das ferramentas (21=78%)

utiliza diferentes técnicas de visualização (R1), empregando múltiplas visões (R3) (20=74%). No entanto, menos da metade (10=37%) permitem o uso de múltiplas instâncias (R2) e apenas 5 (18%) possibilitam que uma quantidade qualquer seja instanciada (R4). Essas constatações vem do fato de que a maioria das ferramentas não flexibiliza o uso das diversas visões, mantendo as visualizações em áreas restritas que ocupam posições fixas na área de trabalho.

Em relação aos requisitos que se referem à representação do histórico do processo de exploração (R6, R8), percebe-se a aderência de uma quantidade pouco significativa de ferramentas (2=7% e 3=11% respectivamente), fato também observado em relação à exploração de múltiplos atributos de arestas (5=18%). No entanto, no que se refere a atributos de subgrafos, uma quantidade um pouco maior é contemplada (12=44%).

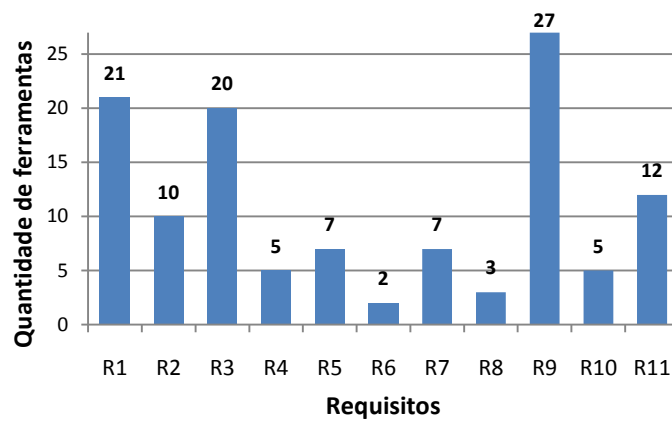


Figura 3.9 – Quantidade de ferramentas que atendem os requisitos

Ainda em relação à quantidade de técnicas utilizadas pelas ferramentas, percebe-se no gráfico da Figura 3.10 que a maior parte destas (13=48%) que atendem ao requisito R1 se utilizam de 2 ou 3 técnicas, o que pode se tornar insuficiente para a exploração de grafos mais complexos. Cerca de um terço (8=30%) das ferramentas utilizam mais de três técnicas e 6 (22%) utilizam apenas 1, não atendendo ao requisito R1.

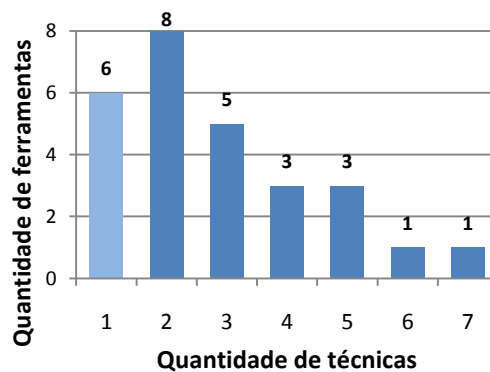


Figura 3.10 – Quantidade de ferramentas x quantidade de técnicas distintas

A Figura 3.11 representa um detalhamento do gráfico anterior, mostrando a relação das ferramentas classificadas por quantidade de técnicas disponibilizadas para a exploração do grafo visualizado. As ferramentas propostas por Ko et al. (2014) e Elzen et al. (2014) são as que oferecem ao usuário uma maior quantidade de técnicas (7 e 6, respectivamente) para a execução da exploração.

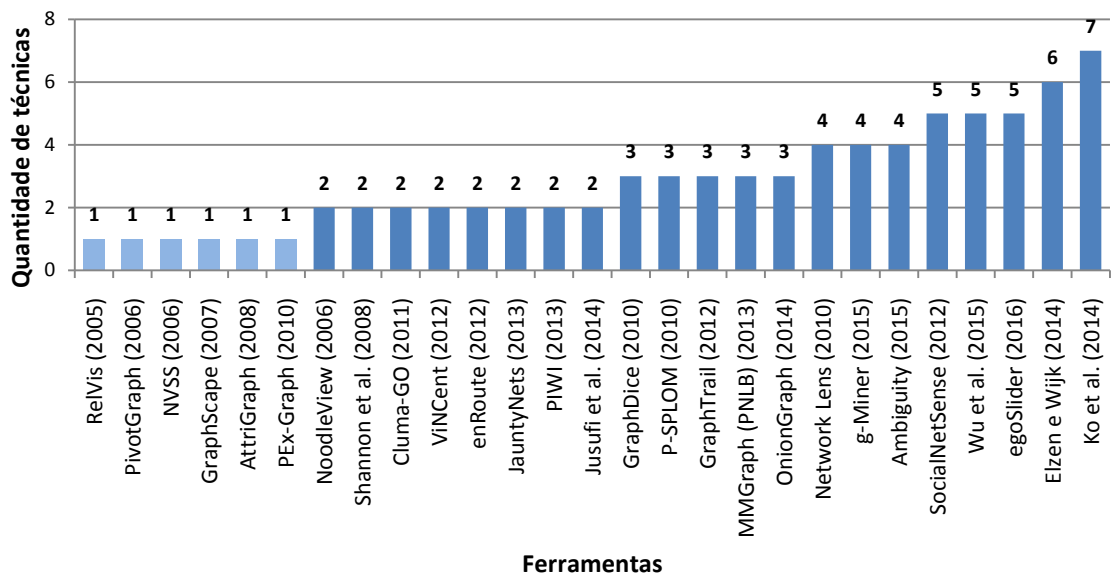


Figura 3.11 – Quantidade de técnicas distintas utilizadas pelas ferramentas

E, finalmente, para avaliar a tendência da aderência aos requisitos com o passar do tempo, foi construído o *scatterplot* da Figura 3.12, que mostra a correlação entre a quantidade de requisitos atendidos e o ano em que cada ferramenta foi divulgada. Percebe-se que, com poucas exceções, as ferramentas construídas no início (2005 a 2008) do período selecionado atendem a um menor número de requisitos, enquanto as que foram publicadas nos últimos 7 anos (2010 a 2016) tendem a contemplar uma quantidade maior, o que era esperado.

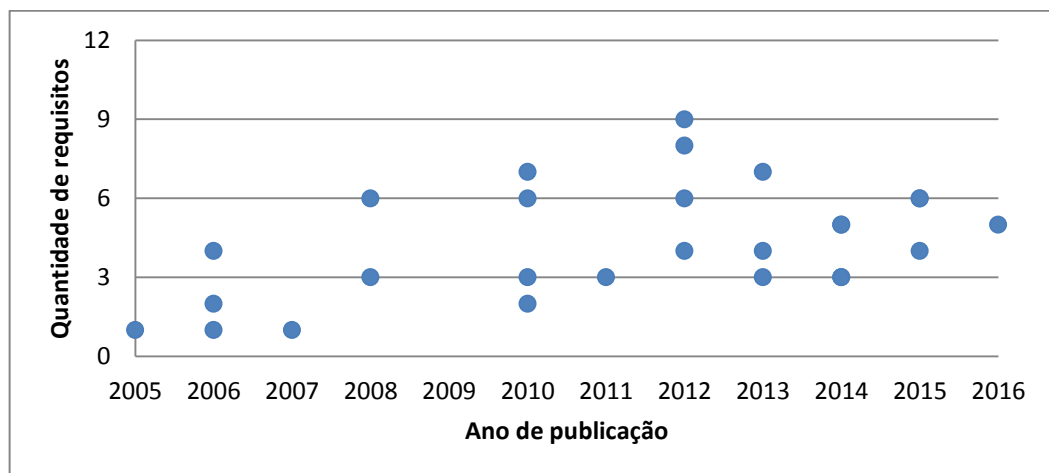


Figura 3.12– Quantidade de técnicas distintas utilizadas pelas ferramentas

3.3 Caminhos possíveis para a exploração de grafos multivariados

Nesta seção será apresentado como as múltiplas instâncias de visualização podem ser usadas na exploração de grafos multivariados. Será definido um conjunto de etapas que descrevem, de uma forma genérica, como executar esse processo de exploração. Seguindo o mantra *Overview first, zoom and filter, then details-on-demand*, apresentado por Shneiderman (1996), o ponto de partida é estabelecido com uma primeira instância de visualização que mostra uma visão geral dos dados a serem explorados. Na sequência, itens de dados podem ser selecionados para serem explorados em outra visualização. Esses itens irão constituir um subgrafo que pode ser obtido a partir das seguintes alternativas de seleção:

- a) **Seleção de nodos baseada na conectividade:** Nesse tipo de seleção os elementos do subgrafo são obtidos a partir dos nodos levando-se em conta aspectos topológicos do grafo:
- **Seleção por nível:** O subgrafo resultante será formado por um nodo inicial (nodo selecionado) e todos os outros que estão acessíveis a partir dele e que se encontram localizados até um determinado nível fornecido pelo usuário. Apenas as arestas que conectam os nodos que atendem ao critério mencionado anteriormente são incluídas. A Figura 3.13 ilustra as duas situações nas quais o nodo inicial está identificado com uma borda mais espessa.

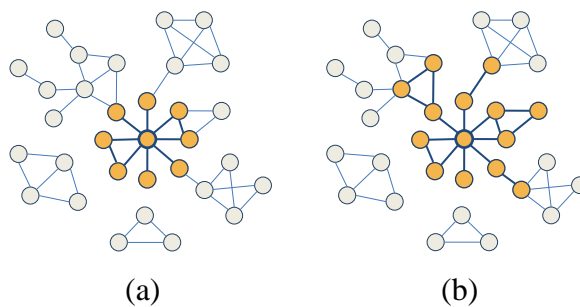


Figura 3.13 – (a) Seleção de 1 nível (b) Seleção de 2 níveis

Exemplo 1 - Para um grafo com nodos representando autores e arestas representando uma relação de coautoria, seria possível a seguinte exploração por nível: Selecionar um autor e seus coautores (1 nível).

Exemplo 2 - Para um grafo com nodos representando professores e arestas uma relação de orientação (ou coorientação), seria possível a seguinte exploração

baseadas em níveis: Selecionar um orientador, seus orientados e respectivos orientandos (2 níveis).

- **Seleção por caminho:** Inclui no grafo gerado os nodos e arestas pertencentes ao caminho existente entre dois nodos selecionados. Vários critérios podem ser usados na escolha desses elementos, como, por exemplo: o caminho mais curto (Figura 3.14a), todos os caminhos simples (sem nodos repetidos) (Figura 3.14b), apenas caminhos que possuem arestas com um determinado valor (ou valores) de atributos.

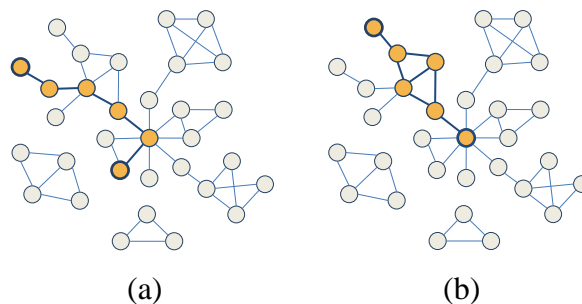


Figura 3.14 – (a) Caminho mais curto (b) Todos os caminhos simples

Exemplo - Para um grafo com nodos representando localidades e arestas estradas, seria possível a seguinte exploração, usando seleção por caminho: Selecionar duas localidades e analisar atributos como tipo de pavimentação, distância, quantidade de pedágios existentes na estrada mais curta que as ligam.

- b) **Seleção manual de nodos e arestas:** Nesse tipo de seleção os elementos são obtidos a partir da escolha pelo usuário de quais nodos e/ou arestas farão parte do subgrafo resultante.
 - **Seleção aleatória de elementos:** Inclui no grafo gerado todos os nodos e arestas selecionados pelo usuário. A seleção pode ocorrer pela escolha individual de elementos, bem como pelo conjunto que estiver contido em uma área de seleção (Figura 3.15).

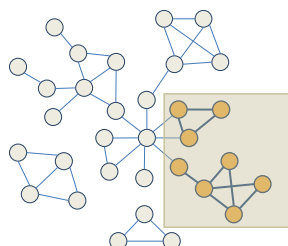


Figura 3.15 –Utilização de área de seleção

Exemplo - Para um grafo com nodos representando funcionários de uma empresa e arestas a troca de e-mails entre eles, uma exploração possível por seleção aleatória seria: Selecionar os nodos pertencentes a uma área visualmente mais densa para análise da comunicação entre esses funcionários.

- **Escolha de caminho:** A partir de um nodo inicial o usuário seleciona recursivamente um dos nodos adjacentes definindo um caminho (Figura 3.16), cujos elementos constituirão o grafo resultante.

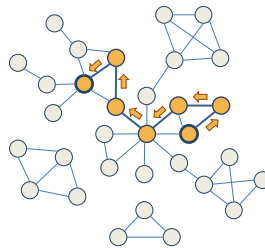


Figura 3.16 –Seleção recursiva de um caminho

Exemplo - Para um grafo com arestas representando ruas e avenidas de uma cidade com atributos, como quantidade de escolas, comércios, bancos, hospitais, etc, seria possível a seguinte exploração por escolha de caminho: Selecionar múltiplos entroncamentos de uma avenida para análise do tipo de uso urbano.

- c) **Seleção por *cluster*:** Nessa seleção os subgrafos são gerados automaticamente por meio de algoritmos computacionais (Figura 3.17a). Estes usam a estratégia de agrupar elementos do grafo original, ficando o mesmo subdividido em supernodos, que representam *clusters*, e superarestas, que consistem nas ligações existentes (Figura 3.17b) entre os nodos de *clusters* diferentes. Após a geração dos *clusters* cabe ao usuário escolher aquele que será explorado. Foram identificados dois modos para executar essa clusterização:

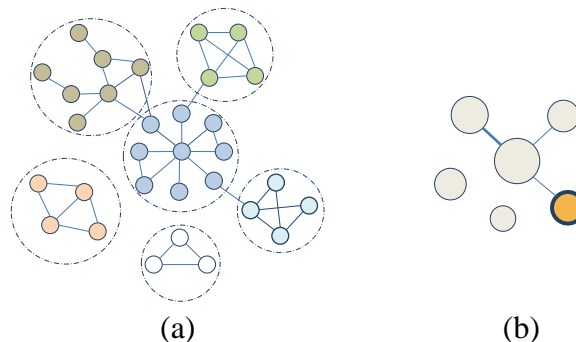


Figura 3.17 –(a) *Clusters* gerados automaticamente (b) Supernodos e superarestas

- **Por algoritmos de clusterização:** A partir da escolha de um algoritmo e seus parâmetros, o grafo inicial é clusterizado gerando um conjunto de supernodos (subgrafos) que poderão ser selecionados pelo usuário.

Exemplo - Para um grafo com nodos representando autores e arestas representando uma relação de coautoria, um exemplo de aplicação desse modo de selecionar um subgrafo poderia ser: Criar uma clusterização que agrupe automaticamente os nodos com maior quantidade de relacionamentos entre si com o objetivo de analisar atributos de grupos de autores que costumam publicar em coautoria.

- **Por valor de atributos:** Nesse caso, o conjunto de nodos pertencentes a cada *cluster* é gerado levando em consideração os valores de um determinado atributo escolhido pelo usuário.

Exemplo - Para um grafo com nodos representando autores e arestas representando uma relação de coautoria, um exemplo de aplicação desse modo de selecionar subgrafo poderia ser: Criar uma clusterização que agrupe automaticamente os nodos que possuem o mesmo valor para o atributo área de pesquisa principal de cada autor.

Na etapa seguinte à seleção do subgrafo, os atributos e/ou relacionamento associados aos seus componentes podem ser explorados por meio de instâncias de visualização, o que pode demandar:

- O uso de várias visões do mesmo subgrafo, utilizando diferentes técnicas de visualização;
- O uso da mesma técnica de visualização para exibir subgrafos distintos;

3.4 Conexões entre visões e histórico de exploração de grafos multivariados

O processo de exploração de grafos multivariados pode seguir vários caminhos e a escolha do caminho e/ou técnica de visualização a ser utilizada depende da tarefa e do objetivo do usuário. Tal flexibilidade durante o processo de exploração pode ser acompanhado pelas conexões entre as visões utilizadas de forma a mostrar o caminho já explorado. A Figura 3.18 mostra um *storyboard* que descreve um processo de exploração de grafos multivariados que mantém o histórico da exploração e a conexão entre as visões.

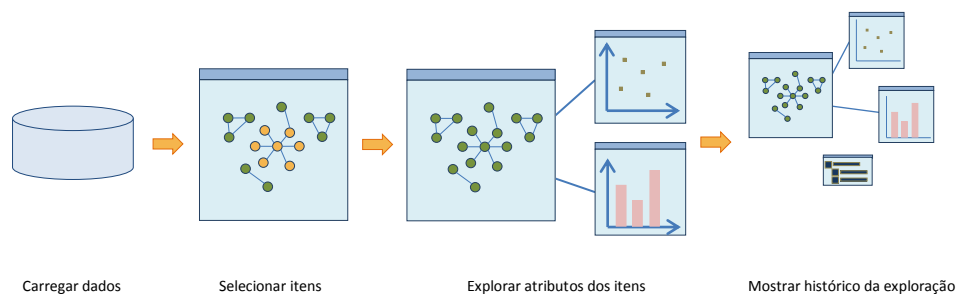


Figura 3.18 – Processo de exploração

A seguir é apresentado um detalhamento de cada etapa do processo de exploração.

- 1) O processo se inicia com uma consulta do usuário que define os dados que serão explorados. O grafo multivariado obtido pela consulta é então carregado na ferramenta de exploração.
- 2) A segunda etapa é a responsável pela seleção dos subconjuntos de dados que serão visualizados nas diversas técnicas. Tal etapa pode ser executada a partir de qualquer visão que já tenha sido instanciada anteriormente.
- 3) Na terceira, é executado o processo de exploração propriamente dito, no qual visões são instanciadas conforme as tarefas de usuário. Ligações explícitas entre as visões são exibidas de forma a mostrar o caminho já explorado. Como mencionado no item 2, novas visões podem ser instanciadas a partir de itens selecionados em uma determinada visão.
- 4) Por fim uma representação adicional deve ser disponibilizada para mostrar todo o histórico da exploração.

Como prova de conceito da abordagem, foi desenvolvida a ferramenta denominada *Multivariate Graph Explorer (MGExplorer)*, descrita no Capítulo 5, no qual exemplos concretos do processo de exploração são apresentados.

4 TÉCNICAS DESENVOLVIDAS

Considerando o *Nested Model* (Munzner 2009, 2014) este capítulo está relacionado ao terceiro nível: “*design* das técnicas de interação e da codificação visual”. Nele é apresentado uma descrição das três técnicas que foram desenvolvidas com o objetivo de atender a premissa 3 reproduzida abaixo.

Premissa 3: A exploração de grafos multivariados não exige apenas a exploração de nodos, mas também pode exigir a inspeção de arestas, subgrafos e relacionamentos criados devido ao valor dos atributos de nodos e arestas.

Embora as técnicas possam ser usadas individualmente, recomenda-se que sejam integradas em um único sistema de forma similar a aplicação apresentada como estudo de caso no capítulo 5. Dessa forma, cada conjunto de dados de interesse pode ser exibido utilizando-se a técnica mais adequada. As três técnicas desenvolvidas neste trabalho são as seguintes: *ClusterVis*, projetada para a exploração de atributos de nodos em subgrafos; *GlyphMatrix*, que explora atributos de arestas em subgrafos; e a *Iris* que visualiza atributos de arestas de nodos adjacentes a um determinado nodo de interesse. As figuras que ilustram a representação visual das técnicas utilizam como exemplo a rede de coautoria entre os docentes, estudantes e participantes externos do Instituto de Informática da UFRGS.

4.1 *ClusterVis*

A técnica *ClusterVis* foi desenvolvida com o objetivo de visualizar as relações entre atributos de nodos pertencentes a agrupamentos em grafos multivariados. Considera-se que o subgrafo que constitui esse agrupamento tenha sido obtido a partir de uma seleção com base em (i) critérios topológicos, (ii) valores de atributos de nodos e/ou arestas, ou (iii) pela aplicação de algoritmos de clusterização. Com a *ClusterVis* pretende-se que vários atributos de nodos sejam exibidos simultaneamente, sem sobreposição, em conjunto com as arestas que representam os relacionamentos existentes entre esses nodos. Assim, a técnica atende aos requisitos R9 e R11 da premissa 3 mencionados na seção 3.1.

4.1.1 Representação visual

Na *ClusterVis* os nodos são dispostos em torno de um círculo em um *layout* radial, com seus atributos numéricos mapeados para retângulos com altura proporcional ao valor que eles representam. Essa estratégia permite a comparação dos valores de um mesmo atributo em nodos diferentes, bem como a busca de possíveis padrões dentro dos atributos de um mesmo nodo. Atributos categóricos são representados por pequenos círculos nos quais a cor é usada

para mapear cada categoria. A Figura 4.1 ilustra essa representação exibindo quatro atributos (3 numéricos e 1 categórico) pertencentes a um subgrafo referente a um conjunto de autores (nodos) conectados entre si por uma relação de coautoria. São eles: quantidade de artigos publicados em *Journals* (retângulos na cor verde), quantidade de capítulos de livros (vermelho), quantidade de artigos publicados em conferências (laranja) e categoria dos autores – Docente, discente e participante externo – (círculos). O subgrafo selecionado contém os autores que possuem a maioria de suas publicações na linha de pesquisa Engenharia de Software.

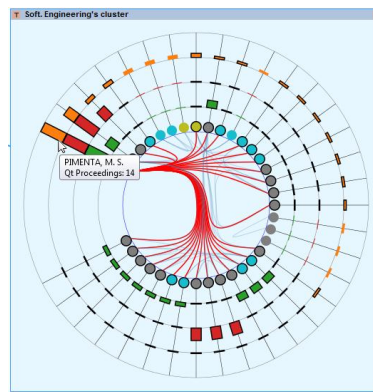


Figura 4.1 – *ClusterVis* – Atributos de nodos em um *cluster*

Cada conjunto de valores de um determinado atributo é exibido em um anel diferente do círculo. Cores são usadas para identificar cada um dos atributos. A parte central do diagrama (círculo) é reservada para exibir as arestas que representam as relações de coautoria. Elas são geradas por meio de uma abordagem semelhante à usada nas *bundles edges* hierárquicas (HOLTEN, 2006) descritas a seguir.

Inicialmente é gerada uma árvore de altura igual a um valor pré-determinado, considerando os nodos distribuídos em torno do círculo como nodos folhas (Figura 4.2a). A seguir, as arestas são desenhadas por meio de *B-splines* cujos pontos de controle são determinados a partir dos nodos dessa árvore (Figura 4.2b).

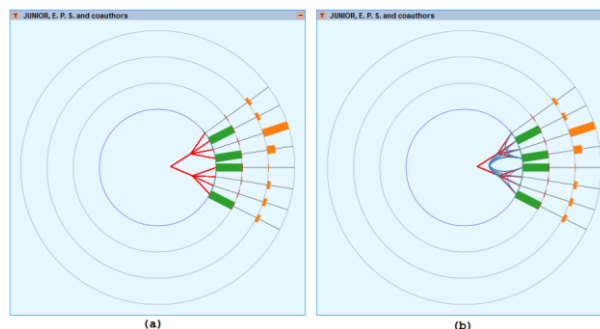


Figura 4.2 – (a) Árvore geradora (b) Árvore geradora e *B-splines*

4.1.2 Interação

Foram previstos os seguintes mecanismos de interação para a técnica *ClusterVis*:

- Seleção dos atributos visualizados.

Os atributos exibidos em cada anel podem ser escolhidos pelo usuário, permitindo que possa explorar várias combinações de atributos.

- Ordenação

Os retângulos que representam os atributos dos nodos podem ser classificados pelo valor de um determinado atributo selecionado pelo usuário, como, por exemplo, a quantidade total de publicações em conferências. Quando o usuário altera o critério usado na classificação, a transição entre os dois estados é executada com animação para minimizar a desorientação do usuário. A seleção do critério é realizada por meio de um *combobox* localizado em um painel associado à área de visualização (Figura 4.3), que pode ser movido de forma independente de acordo com a necessidade do usuário.

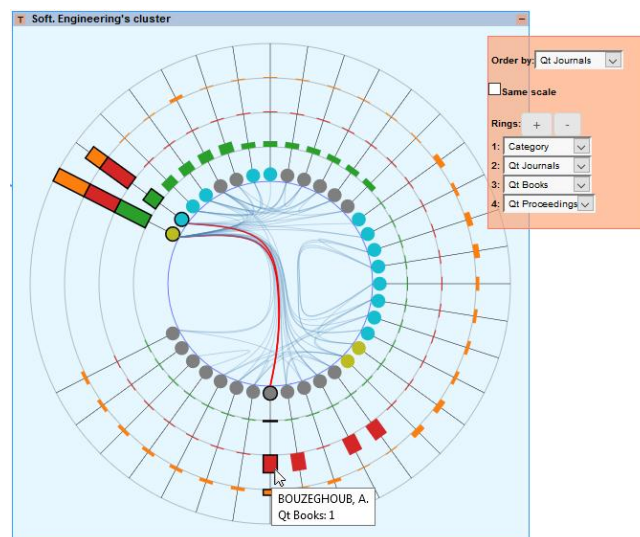


Figura 4.3 – Ordenação dos valores de atributos

- Destaque de arestas

Para facilitar a percepção dos relacionamentos entre os nodos, foi incluído um mecanismo que destaca na cor vermelha as arestas que estão conectadas ao retângulo (nodo) que se encontra sob o cursor do mouse. Além disso, os nodos conectados a outra extremidade de cada aresta destacada (nodos adjacentes) também são realçados com bordas na cor preta.

- *Tooltips*

Tooltips também foram adicionados à representação com o objetivo de exibir os valores numéricos dos atributos sob demanda, bem como outras informações de interesse (ex: nome do autor).

- Redimensionamento dos elementos geométricos

Foram previstos mecanismos para alterar as dimensões dos principais elementos geométricos da representação, como, por exemplo, o raio dos círculos interno e externo, a largura de cada anel, a largura das barras e o espaçamento entre elas. Entretanto, no protótipo desenvolvido neste trabalho, somente o redimensionamento do círculo externo foi implementado.

4.1.3 Limitações

A *ClusterVis* possui algumas limitações no que se refere à escalabilidade. Subgrafos compostos por muitos nodos não podem ser visualizados em razão da limitação do perímetro do círculo que não permite a representação de uma quantidade de nodos acima de certo limite máximo. Alternativas para a minimização desse problema devem ser investigadas. Subgrafos com muitas conexões entre os nodos podem causar uma certa dificuldade na percepção detalhada dos relacionamentos, como mostra a Figura 4.4. Alternativas para os algoritmos de *bundles edges* podem ser investigadas.

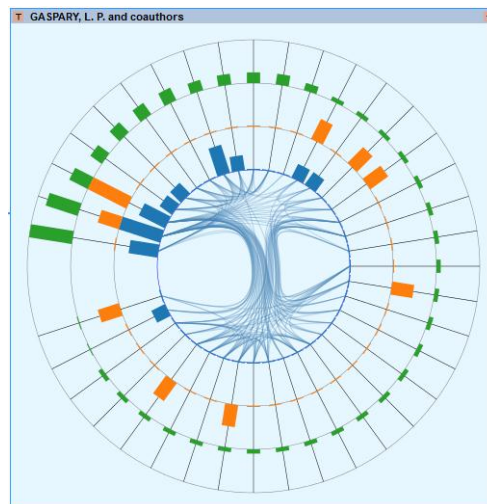


Figura 4.4 – Sobreposição de arestas

4.2 *GlyphMatrix*

A *GlyphMatrix* foi criada como uma alternativa para a visualização dos atributos de arestas pertencentes a agrupamentos em grafos multivariados de forma a atender aos requisitos R10 e R11 da premissa 3, mencionados na seção 3.1. A representação matricial

adotada pela técnica propicia a visualização de grafos densos sem que ocorra o cruzamento de arestas. No entanto, ela possui a desvantagem de dificultar a percepção de caminhos.

4.2.1 Representação visual

Na *GlyphMatrix* os nodos são representados através de linhas e colunas de uma matriz e as arestas, pelas células localizadas nas respectivas intersecções. A representação adota uma abordagem semelhante à usada por Ulrik e Bobo (2011), a qual exibe um glifo em cada célula para mostrar a evolução no tempo dos dados associados às arestas. A área localizada na lateral esquerda e na parte superior da matriz é reservada para exibir um texto escolhido pelo usuário para identificar cada nodo do grafo. O glifo adicionado em cada célula mapeia os atributos da respectiva aresta. A Figura 4.5 ilustra a representação mostrando a relação de coautoria entre pesquisadores das áreas de engenharia de software e modelagem conceitual que possuem publicações em comum. Nesse exemplo, um *starPlot* para cada aresta foi usado para exibir a quantidade de cada tipo de publicação (*Journals*, capítulo de livros e conferências).

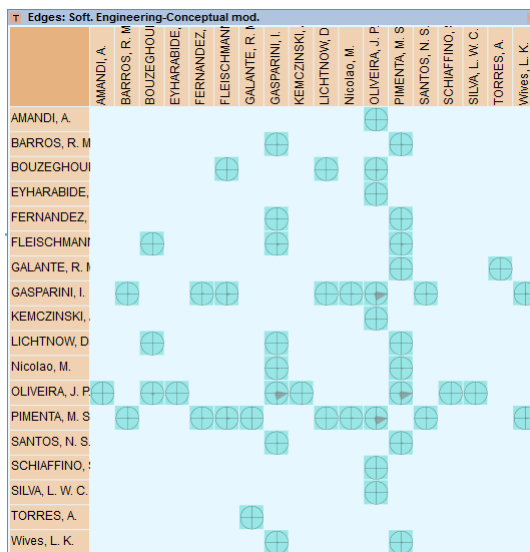


Figura 4.5 – *GlyphMatrix* – atributos de arestas

Outros glifos estão disponíveis como alternativas para a representação dos atributos (Figura 4.6).

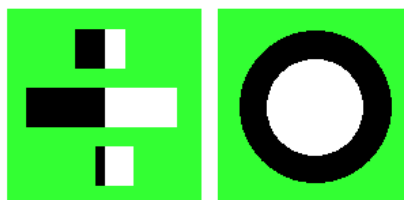


Figura 4.6– Glifos disponíveis na *GlyphMatrix*

A cor de fundo de cada célula também pode ser usada para mapear o valor de um atributo. No exemplo apresentado na Figura 4.5, a cor de fundo não está sendo utilizada para representar atributos, mas apenas para indicar as arestas existentes. Um gráfico de barras pode ser sobreposto nas áreas reservadas a identificação dos nodos com a finalidade de exibir o valor de um de seus atributos.

4.2.2 Interação

Foram previstos os seguintes mecanismos de interação para a técnica *GlyphMatrix*:

- Ordenação

As linhas e colunas da matriz podem ser reordenadas por vários critérios escolhidos livremente pelo usuário. Assim como na *GlyphMatrix*, a transição entre os dois estados da ordenação é executada com animação para minimizar a desorientação do usuário.

- *Tooltips*

Tooltips que mostram os glifos com um tamanho aumentado foram adicionados à representação com o objetivo de mostrar informações dos atributos das arestas de forma mais detalhada (Figura 4.7).

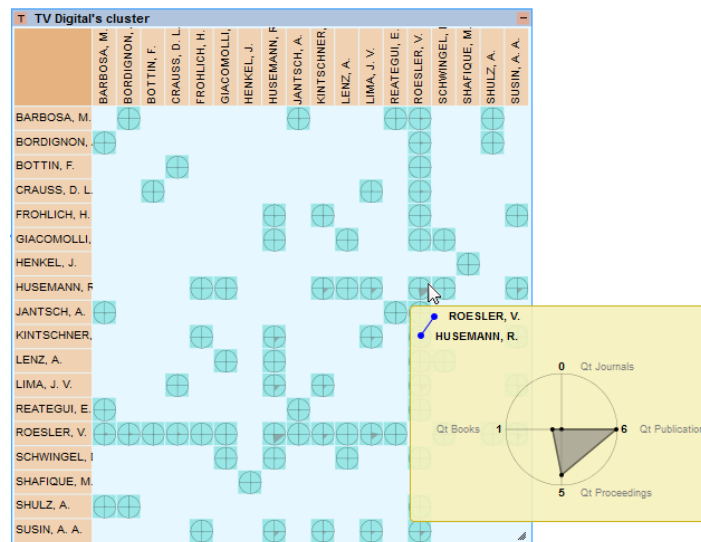


Figura 4.7– *Tooltips* que exibem glifos

- Redimensionamento das células

Dependendo das dimensões da área utilizada para exibir a matriz e da quantidade de nodos contida no subgrafo, o tamanho de cada célula pode se tornar demasiadamente pequeno a ponto de impossibilitar a identificação dos glifos. Para lidar com isso, a representação admite a modificação das dimensões das células através de um mecanismo que permite alterar

a quantidade de linhas (e colunas) disponíveis para exibir os nodos da matriz. Quando essa quantidade é reduzida, a área de visualização passa a exibir apenas uma porção da matriz, fazendo com que cada célula tenha as duas dimensões aumentadas (Figura 4.8). O arrasto do mouse sobre os textos que identificam os nodos ou sobre as células da matriz permite que esta seja movida de modo que o usuário possa controlar quais nodos estarão visíveis. Uma área localizada no canto superior esquerdo mostra uma visão geral da matriz na qual um retângulo marrom indica a porção que está sendo visualizada.

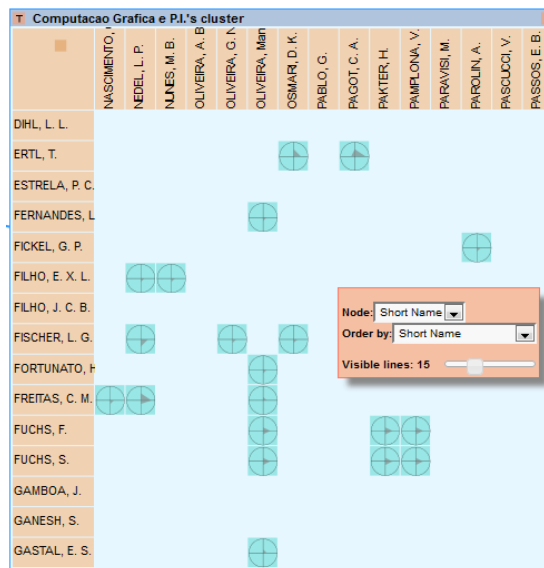


Figura 4.8– Matriz 106 x 106 com 15 linhas visíveis

4.2.3 Limitações

A *GlyphMatrix*, tal como a representação descrita anteriormente, possui limitações no que se refere à escalabilidade. Matrizes com uma maior quantidade de nodos exibem células com dimensões reduzidas o que dificulta a percepção dos glifos. Embora o mecanismo de redimensionamento das células, que permite aumentar as suas dimensões, auxilie na exploração de tais matrizes, a conseqüente redução da porção visível decorrente de seu uso restringe a visão geral da matriz. Outra limitação a ser considerada é quantidade máxima de atributos que pode ser representada em um glifo de forma perceptível em razão de sua dimensão reduzida.

4.3 Iris

De forma similar à *GlyphMatrix*, a *Iris* também foi projetada com o objetivo de visualizar atributos de arestas. Entretanto, ela atende a uma situação mais específica, concentrando-se na visualização dos atributos das arestas (R10 da premissa 3) que conectam

o sentido anti-horário. A rotação é realizada com animação de forma a minimizar a desorientação do usuário (HEER e ROBERTSON, 2007).

4.3.3 Limitações

A solução adotada para o problema da escalabilidade da *Iris* foi usar a distorção *fisheye* para mostrar uma maior quantidade de elementos e permitir a rotação dos elementos quando nem todos estiverem visíveis. Entretanto, embora essa estratégia funcione bem quando o critério usado na ordenação dos nodos for o atributo que está sendo exibido, para outros critérios, como, por exemplo, ordem alfabética, isso pode ocasionar que um valor de destaque se posicione na área não visível impossibilitando a sua percepção. Criar uma forma de indicar essa situação pode ser uma solução para o problema.

Outro ponto a ser considerado é o fato de que, embora o objetivo principal da técnica seja mostrar atributos de arestas entre um determinado nodo e seus adjacentes, a identificação dos relacionamentos existentes entre esses nodos (adjacentes) poderia ser de interesse do usuário. Uma solução possível, ainda que não implementada, seria, ao invés de utilizar linhas para representar as arestas, destacá-los quando o mouse passa sobre eles de forma similar ao que corre na *ClusterVis*.

5 MULTIVARIATE GRAPH EXPLORER

No capítulo 3 foram apresentados os onze requisitos que deveriam ser atendidos por uma ferramenta para que se enquadrasse na abordagem proposta neste trabalho. Este capítulo descreve a ferramenta *MGExplorer* (*Multivariate Graph Explorer*), desenvolvida como prova de conceito, que viabiliza a instanciação de múltiplas técnicas de visualização em janelas redimensionáveis. Além da descrição da ferramenta com os requisitos atendidos (seção 5.1), será apresentado também um estudo de caso (seção 5.2) com dados referentes à rede de coautoria dos pesquisadores do Instituto de Informática/UFRGS, com o objetivo de exemplificar como as múltiplas visões podem ser combinadas para criar visualizações interligadas durante um processo de exploração de grafos multivariados.

5.1 Descrição da *MGExplorer*

A *MGExplorer* é composta de uma área de trabalho na qual uma quantidade variável (R4) de múltiplas técnicas de visualização (R1) pode ser instanciada em visões redimensionáveis (R5) de forma que um ou vários conjuntos de dados possam ser explorados sob vários aspectos (R3). A ferramenta permite ainda que múltiplas instâncias de uma mesma técnica (R2) possam ser usadas de forma a possibilitar a comparação de dois conjuntos de dados. Apesar de ser possível a criação de diversas alternativas de visualização, quatro técnicas foram escolhidas para a prova de conceito: o diagrama de nodos e arestas, ponto de partida da exploração, e as técnicas *ClusterVis* (R9 e R11), *GlyphMatrix* (R10 e R11) e *Iris* (R10), apresentadas no capítulo 4, que foram desenvolvidas para representar características distintas de grafos multivariados. A Figura 5.1 mostra uma visão geral da *MGExplorer*.

A barra superior dessa ferramenta contém dois *combobox* (*Select to load* e *Cluster by*) que permitem a escolha do grafo que será carregado para exploração e a relação de algoritmos que podem ser usados para a execução de uma clusterização. O exemplo da Figura 5.1 mostra um diagrama de nodos e arestas que representa um grafo de coautoria das publicações do ano de 2004 sem clusterização. Cada visão possui uma área associada (cor alaranjada) que pode conter ferramentas de configuração e mecanismos de interação que propiciam alterações na representação.

Dois tipos de linhas (Figura 5.1a) foram usados na representação das ligações explícitas entre as visões (R6) de tal modo que a sequência de exploração utilizada possa ser identificada. A linha cheia conecta duas visões nas quais as técnicas estão sendo exibidas e a linha tracejada indica uma ligação com uma visão oculta (Figura 5.1b) representada por um retângulo cinza. Em uma visão desconectada das restantes, é exibida a árvore de exploração

(Figura 5.1c) com a identificação das visões instanciadas na área de trabalho (R8). Um clique no quadrado localizado no canto superior direito remove a visão, que pode ser novamente exibida (R7) a partir de um clique sobre a sua respectiva representação no histórico.

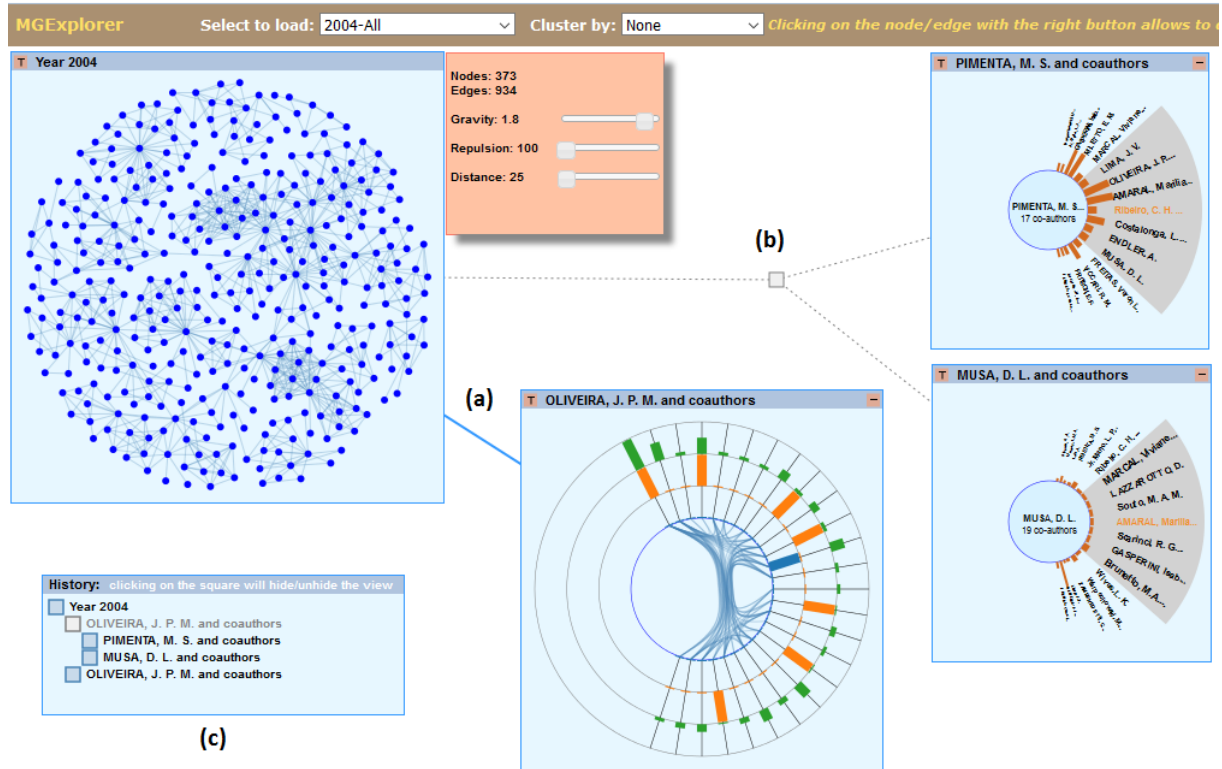


Figura 5.1– *MGExplorer* (a) Ligações entre janelas (b) Janelas ocultas (c) Histórico

A *MGExplorer* foi desenvolvida em *javascript* e fez uso da D3 (*Data-Driven Documents*) (BOSTOCK et al. 2011), uma biblioteca amplamente usada para o desenvolvimento de sistemas de visualização para *web*.

5.2 Estudo de caso

Para exemplificar de que forma a *MGExplorer* pode ser usada para combinar as múltiplas visões em um processo de exploração, será apresentado um estudo de caso com dados da rede de coautoria dos pesquisadores do Instituto de Informática/UFRGS referente a trabalhos publicados no período de 2004 a 2011.

5.2.1 Dados visualizados

Os dados de coautorias estavam disponíveis inicialmente em oito arquivos (um para cada ano) no formato *GraphML*, contendo grafos cujos nodos representavam autores e as arestas, artigos em coautoria. Com o objetivo de reunir todos os dados em uma só estrutura e de flexibilizar a geração dos grafos, foi implementado um software para mesclar os oito arquivos em uma base de dados, cujo E-R é mostrado na Figura 5.2. Dessa forma, grafos que

representam outras relações (ex: nodos representando publicações e arestas, autores que as compartilham) podem ser gerados a partir de consultas realizadas sobre a base. No entanto, nesse trabalho os grafos são obtidos por uma aplicação *ad hoc* que os gera no formato JSON para posteriormente serem carregados na *MGExplorer*.

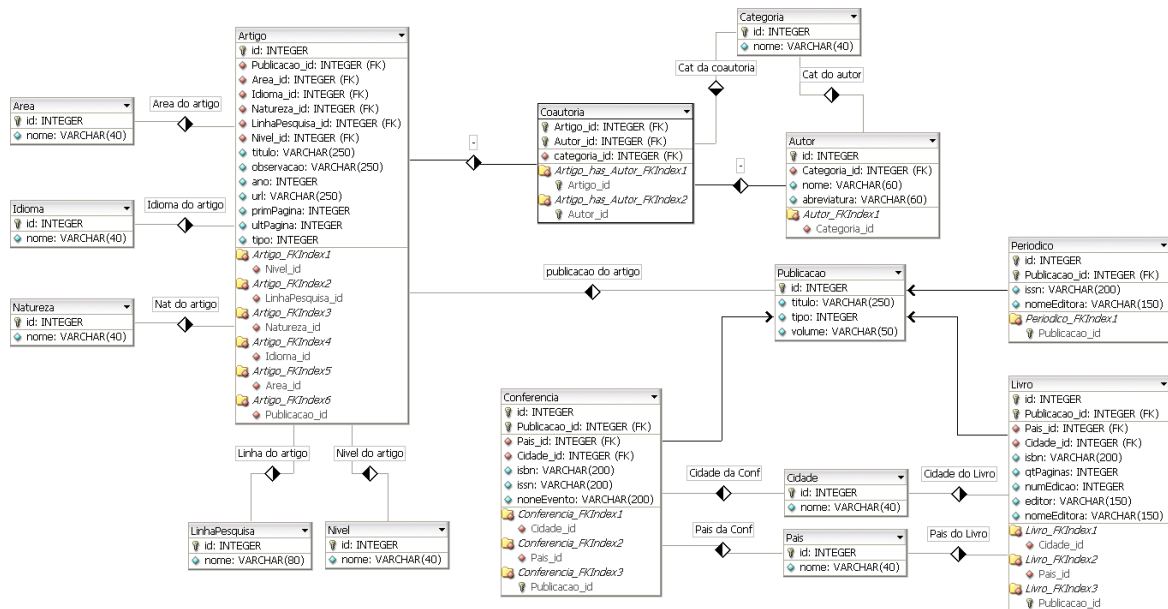


Figura 5.2– E-R da base de dados de coautoria INF-UFRGS

No grafo usado na maioria dos exemplos desse trabalho, os nodos representam autores e as arestas, publicações em coautoria. Considerando o período completo (2004-2011), o grafo contém 1989 nodos e 7250 arestas. No entanto, para os exemplos apresentados, foram gerados grafos filtrados por ano de publicação e linha de pesquisa principal de seus autores. Os nodos possuem atributos nominais (nodo do autor, categoria, linha de pesquisa com maior quantidade de publicações), atributos quantitativos (quantidade total de publicações, independente da coautoria, distribuídas em artigos em *Journals*, capítulos de livros e artigos em conferências) e, finalmente, atributos calculados a partir da topologia (grau, *closeness centrality*, etc). As arestas possuem apenas atributos quantitativos, que representam os totais de publicações em coautoria distribuídas segundo vários critérios: tipos de publicação (quantidade de artigos em *Journals*, capítulos de livros e artigos em conferências), linha de pesquisa e natureza (trabalho completo, resumo, capítulos,...).

5.2.2 Exemplos de exploração

Para ilustrar de que forma a *MGExplorer* pode ser usada para executar uma exploração em um grafo multivariado, serão apresentados nesta seção dois exemplos. O primeiro mostra os passos de um processo de exploração de um grafo de coautoria de todas as publicações de

2011 e o segundo considerando apenas os pesquisadores da linha de pesquisa Computação Gráfica.

Explorando as relações de coautoria de todas as publicações do ano de 2011 do Instituto de Informática/UFRGS considerando as linhas de pesquisa de seus autores.

O grafo da rede de autoria de todas as publicações do ano de 2011 possui 619 nodos e 1986 arestas. O processo inicia com a exibição do grafo em um diagrama de nodos e arestas (Figura 5.3).

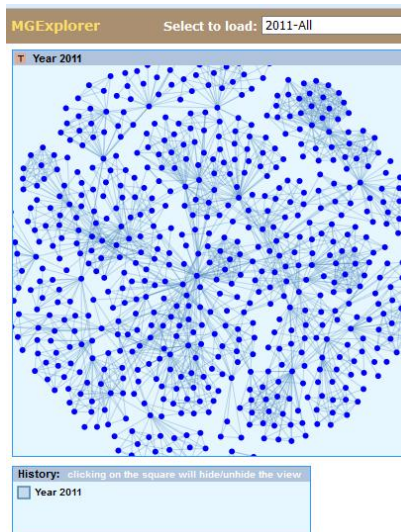


Figura 5.3– Representação inicial do grafo com dados de 2011

Observa-se que, em razão do número de arestas do grafo (1986), essa representação produz uma significativa sobreposição de linhas. Assim, no passo seguinte, um algoritmo de clusterização, baseado no atributo “linha de pesquisa”, é ativado para gerar o grafo da Figura 5.4 no qual cada círculo representa um subgrafo formado pelos autores com maior quantidade de publicações na respectiva área.

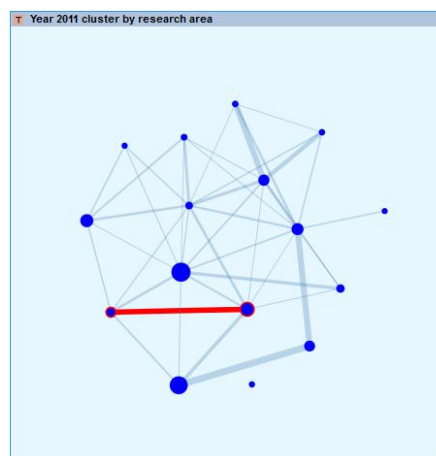


Figura 5.4– Grafo clusterizado por linha de pesquisa

A medida do raio de cada círculo é proporcional à quantidade de nodos do subgrafo que ele representa. As linhas que conectam os círculos representam as publicações cujos coautores pertencem a áreas diferentes e são desenhadas com espessura proporcional à quantidade de publicações. Ao analisar o grafo da Figura 5.4, percebe-se um pequeno círculo que não se conecta com nenhum outro, indicando que os pesquisadores daquela área (Tolerância a Falhas) não publicaram em 2011 em coautoria com os demais.

No passo seguinte, opta-se por investigar as publicações comuns a pesquisadores de duas áreas distintas: Engenharia de Software e Modelagem Conceitual. Dessa forma, a aresta que conecta os dois *clusters* é selecionada para que o subgrafo que ela representa seja exibido na *GlyphMatrix* (Figura 5.5).

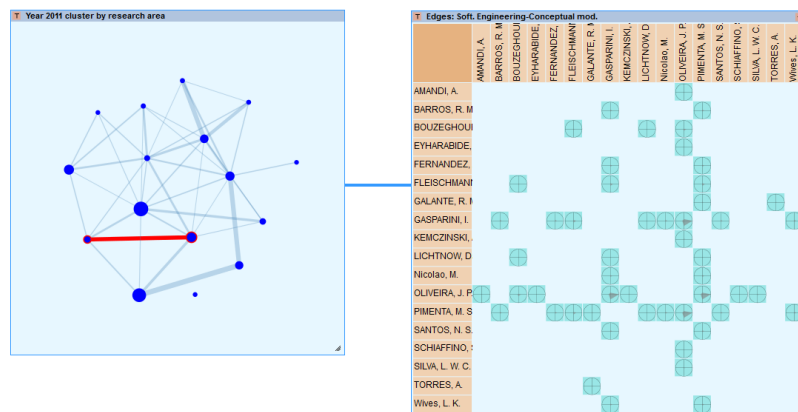


Figura 5.5– Ativação da *GlyphMatrix*

Observando-se a matriz, percebe-se que três linhas (ou colunas) se destacam das outras pela maior quantidade de glifos existentes, indicando que os respectivos pesquisadores (Gasparini, Oliveira e Pimenta) são os que mais publicações possuem em coautoria com pesquisadores da área oposta. Resolve-se então investigar os coautores do pesquisador Pimenta, ativando uma visão com a *Iris* (5.6) a partir da *GlyphMatrix*. Nota-se então que Gasparini é o autor com maior quantidade de publicações em comum.

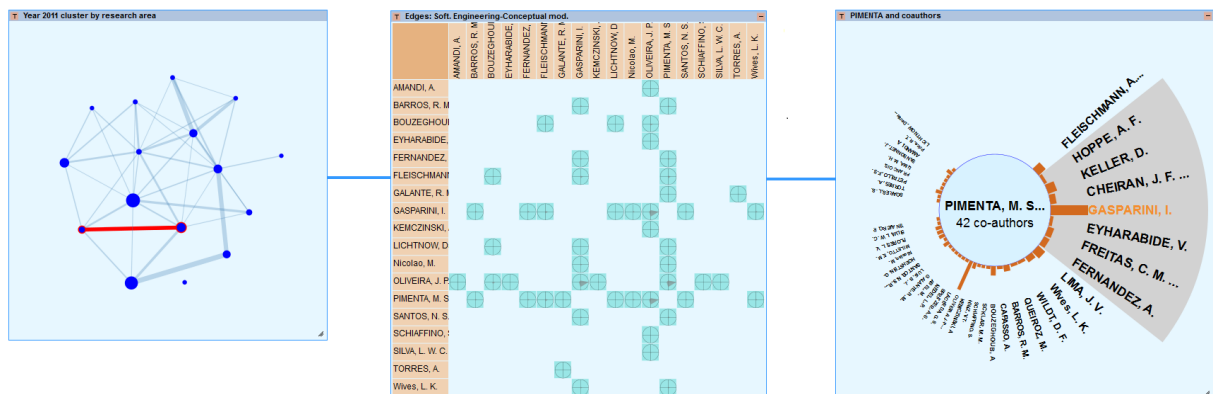


Figura 5.6– Coautores do pesquisador Pimenta

Um segundo ramo de exploração é iniciado a partir do diagrama de nodos e arestas que passa a exibir o subgrafo do *cluster* formado pela coautoria entre os pesquisadores da área de Engenharia de Software em uma *ClusterVis* (Figura 5.7). Observa-se que Pimenta é o autor com maior quantidade de publicações na área e o coautor da maioria dos pesquisadores pertencentes ao grupo. A Figura 5.7 mostra o resultado final do processo de exploração. A hierarquia criada com a instanciação das técnicas é mostrada na visão com o título de *History*.

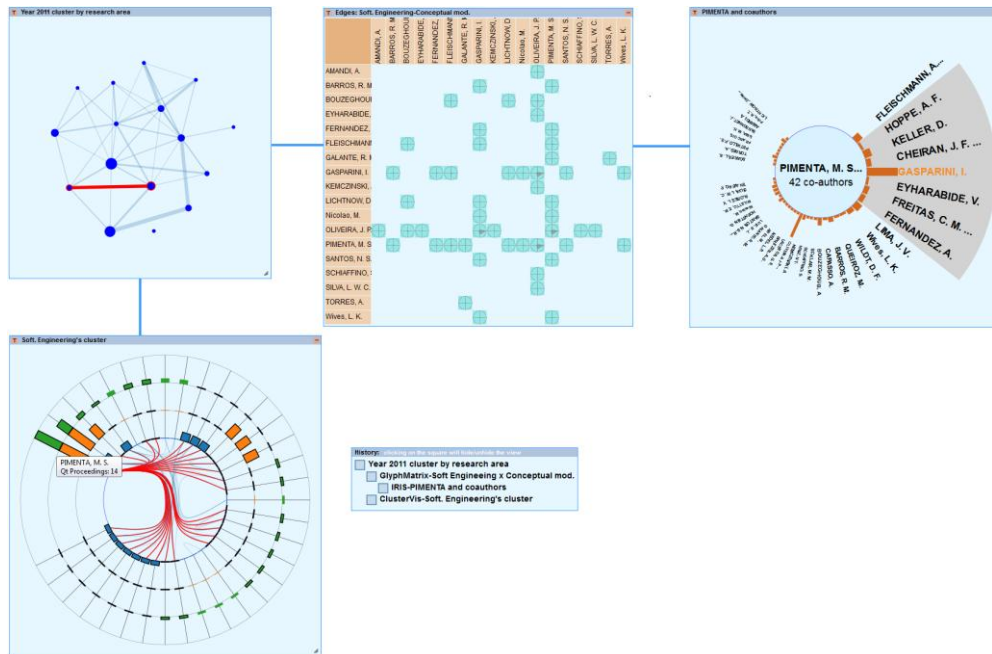


Figura 5.7– Coautorias na linha de pesquisa Engenharia de Software

Explorando as relações de coautoria das publicações do ano de 2011 do Instituto de Informática/UFRGS, dos autores da linha de pesquisa Computação Gráfica.

Neste segundo exemplo, o grafo possui 117 nodos e 366 arestas e, de forma similar ao exemplo anterior, o processo se inicia com a exibição dos dados em um diagrama de nodos e arestas (Figura 5.8). Observa-se então a existência de dois subgrafos separados entre si.

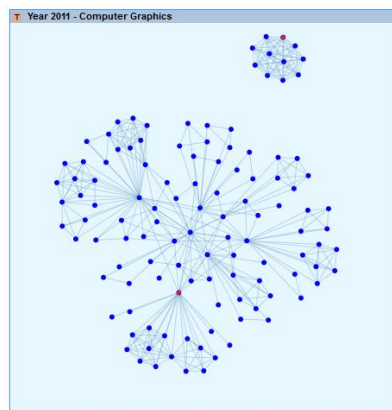


Figura 5.8– Coautorias em Computação Gráfica no ano de 2011

Utilizando-se da *ClusterVis*, uma visão detalhada do subgrafo menor é obtida. O uso da *ClusterVis* para explorar essa rede específica de coautoria permite perceber que todos os 11 nodos do subgrafo estão conectados ao primeiro pesquisador (Navaux) e possuem a mesma quantidade de publicações, sugerindo que todos são coautores de um único artigo.

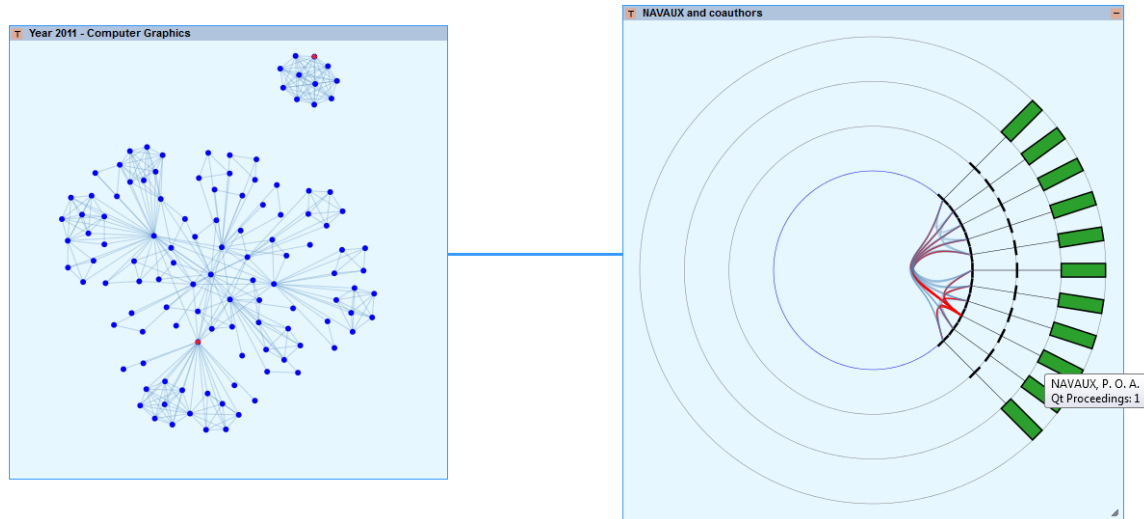


Figura 5.9 – Subgrafo visualizado com a *ClusterVis*

A seguir é ativada uma nova visão usando a *GlyphMatrix* para se poder compará-la com a visualização obtida com a *ClusterVis*. Os glifos confirmam a existência da conexão entre todos os coautores entre si com a mesma quantidade publicações (Figura 5.10).

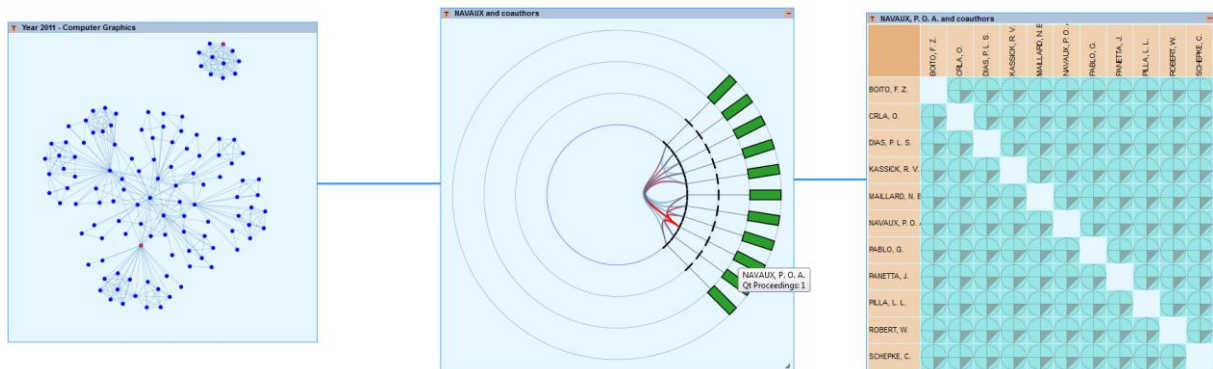


Figura 5.10 – *GlyphMatrix*: – Mostra todas as conexões entre os coautores

A *ClusterVis* que exibe os dados do autor Navaux é oculta da área de trabalho e outro ramo de exploração é iniciado a partir da seleção do pesquisador Oliveira, um dos autores que possui muitos coautores no diagrama de nodos e arestas. O subgrafo formado por Oliveira e seus coautores é exibido com uma *ClusterVis* em uma nova visão (Figura 5.11). As barras azul, laranja e verde correspondem respectivamente a quantidade total de artigos em *Journals*, capítulos de livros e artigos em conferências de autoria de cada pesquisador.



Figura 5.11 – Oliveira e seus coautores

Observando-se os dois retângulos laranjas nessa visão, conclui-se que os respectivos coautores publicaram capítulos de livros naquele ano. Seleciona-se, na sequência, a partir da *ClusterVis* cada um dos coautores (Comba e Maciel) para que sejam exibidos em duas *Iris* distintas. Comparando-se os dois resultados, percebe-se que Comba possui uma maior quantidade de coautores e que Ertl é o coautor com o qual ele publicou mais em 2011. O pequeno quadrado cinza que conecta o diagrama de nodos e aresta à *GlyphMatrix* representa a *ClusterVis* que foi removida da área de trabalho para se obter mais espaço para as novas visões.

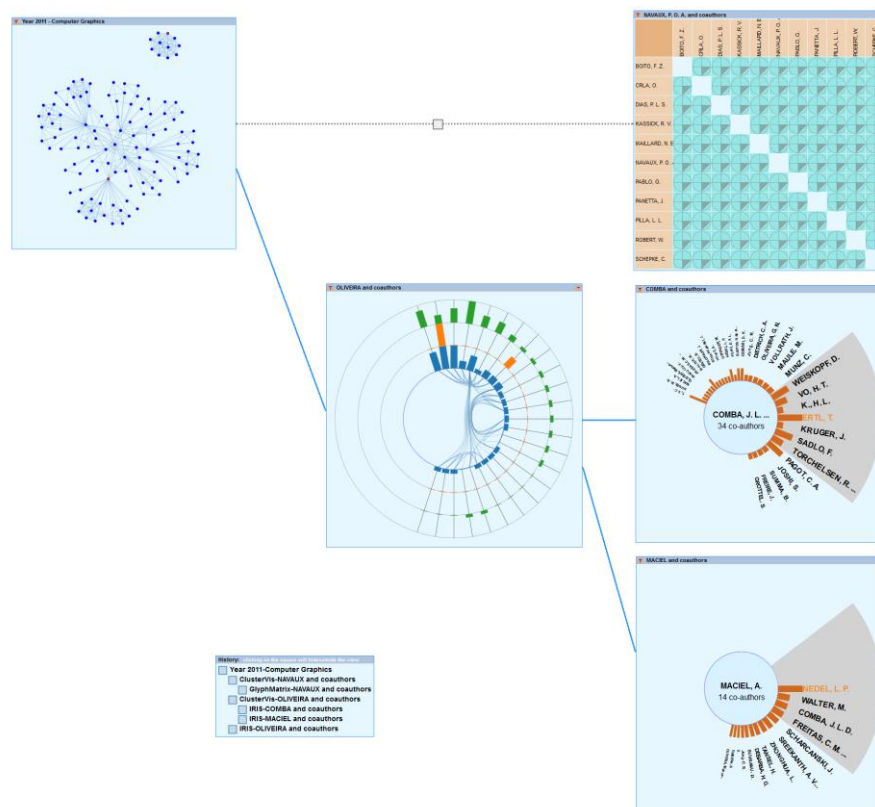


Figura 5.12 – Coautorias dos pesquisadores Comba e Maciel

E, por fim, voltando ao diagrama de nodos e arestas, um terceiro ramo de exploração é iniciado, permitindo a análise da lista de coautores de Oliveira em uma *Iris* (Figura 5.13). Pode-se notar que, embora duas visões tenham sido removidas da área de trabalho, a representação do processo de exploração como um todo é mostrada em uma visão separada (*History*).

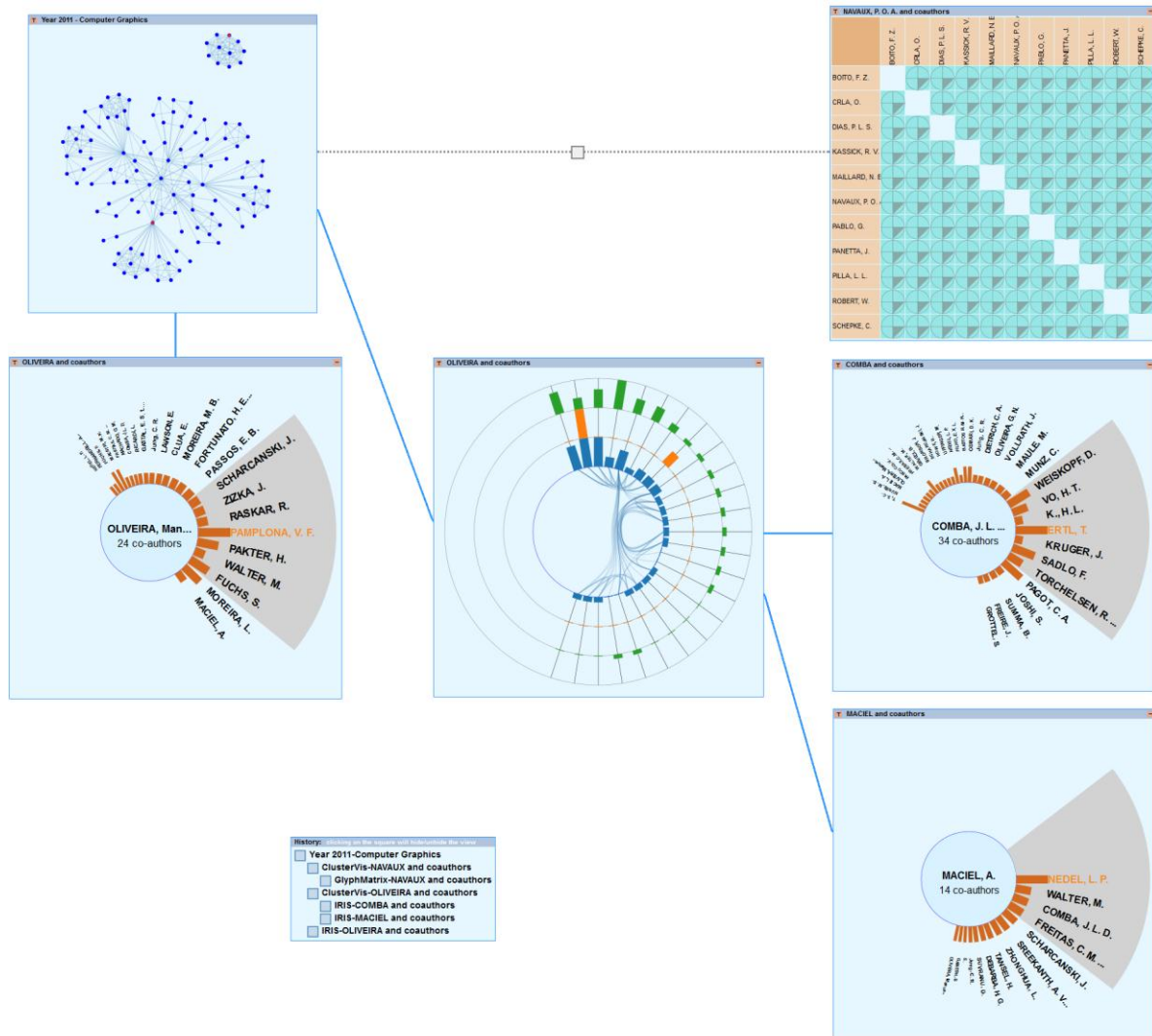


Figura 5.13 – Coautorias na linha de pesquisa Computação Gráfica

6 AVALIAÇÃO

O processo de avaliação de usabilidade utilizado neste trabalho foi desenvolvido em dois níveis distintos. O primeiro refere-se à avaliação realizada individualmente nas três técnicas de visualização para grafos multivariados. O segundo trata da avaliação do protótipo *MGExplorer*, que possibilita a utilização simultânea das técnicas durante o processo de exploração. Foram definidas três dimensões para a avaliação: utilidade, performance e satisfação. Para cada uma das dimensões foi utilizada uma técnica de avaliação distinta.

Para avaliar a dimensão utilidade foi escolhido o método de inspeção de usabilidade *CW (Cognitive Walkthrough)* (LEWIS e WHARTON, 1997). Esse método é utilizado para a avaliação de interfaces com o usuário através da análise do processo mental requerido na execução de tarefas pré-determinadas. O método consiste na análise de uma sequência de ações corretas para executar uma tarefa, identificando se um usuário hipotético seria capaz de selecionar a ação apropriada em cada ponto. Caso uma dificuldade seja constatada uma causa é atribuída a ela. Para realizar a avaliação das três técnicas e da ferramenta *MGExplorer* foi adotado o seguinte protocolo:

Inicialmente foram apresentadas as características dos usuários da ferramenta. A seguir foram descritos os mecanismos utilizados para interagir com cada técnica. E, por fim, após a definição das tarefas foi especificada a sequência de ações necessárias para executá-las. Uma análise determinando o sucesso ou falha de cada ação foi realizada.

Para a avaliação de todas as técnicas desenvolvidas neste trabalho foram considerados usuários, pessoas com experiência em sistemas que utilizam interfaces gráficas baseadas em janelas. Estes usuários também possuem alguma experiência com aplicações gráficas relacionadas à visualização de informações e são capazes de reconhecer estruturas baseadas em grafos.

Em relação à interface serão apresentados nesta seção os mecanismos utilizados por todas as técnicas. Mecanismos específicos são descritos nas seções que avaliam cada técnica individualmente. Fazem parte da interface utilizada para interagir com as janelas nas quais as técnicas de visualização são exibidas os seguintes mecanismos:

- Reposicionamento da janela na área de trabalho com o arrasto da barra de título;
- Redimensionamento das dimensões da janela com o arrasto de uma de suas bordas;

- Reposicionamento do painel de ferramentas com o arrasto do mouse;
- Botão “T” (*Tools*) utilizado para exibir/ocultar o painel de ferramentas que contém alguns mecanismos de interação (Figura 6.1);
- Botão “-“ (*Hide*) utilizado para ocultar a janela da área de trabalho (Figura 6.1).

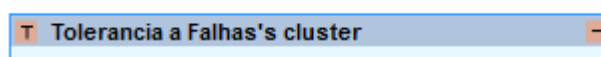


Figura 6.1 – Botões *Tools* e *Hide*

Na avaliação da dimensão performance foi utilizado o modelo *GOMS-Keystroke (KLM)* (CARD et al., 1983). *GOMS (Goals, Operators, Methods, Selection rules)* é um modelo formal cognitivo usado para avaliar rigorosamente o quanto eficientemente uma pessoa treinada pode interagir com um software. *GOMS* é um modelo construído sobre uma sequência detalhada das operações do usuário com um sistema. O *GOMS-KLM* é uma variante que inclui valores conhecidos para ações de usuários permitindo prever o tempo que usuários experientes levarão em situações aparentemente imprevisíveis. Por exemplo, a média de tempo para executar uma ação de levar o mouse a uma posição desejada é de 0,36 segundos, clicar sobre um campo é de 0,23 segundos, etc. Assim a partir de um cenário detalhado com as ações do usuário com a ferramenta é possível usar a técnica *GOMS-KLM* para prever a performance.

Para avaliar as técnicas na dimensão satisfação, foram realizados experimentos nos quais usuários executaram um determinado grupo de tarefas, utilizando aplicações que implementam cada uma das técnicas. A cada tarefa executada, foi solicitada ao usuário a resposta a uma questão sobre a facilidade ou não de uso, segundo uma escala *Likert* de 5 pontos. Após a execução do experimento, foi solicitado, também, o preenchimento de um questionário de acordo com o método *System Usability Scale (SUS)* (BROOKE,1996). Esse questionário foi aplicado apenas para as técnicas *ClusterVis* e *GlyphMatrix*.

Para executar as avaliações, foram utilizados três cenários distintos. O primeiro é baseado no grafo de coautorias dos pesquisadores do Instituto de Informática da UFRGS e foi usado para avaliação de todas as técnicas com os métodos *CW* e *KLM*. Esse cenário também foi usado na avaliação da *ClusterVis* com o método *SUS*. Um segundo cenário foi utilizado na avaliação da *Iris* na dimensão satisfação. Ele refere-se à execução de buscas auxiliares ativadas por usuários a partir de uma página *web* com objetivo de exibir os resultados na forma de uma ajuda contextual posicionada próxima da palavra-chave usada para formular a

consulta. Foi usada como exemplo a visualização de relações de coautorias extraídas da *DBLP*. E, por fim, o terceiro cenário refere-se à visualização de dados estatísticos das partidas do Campeonato Brasileiro de futebol de 2013. Esse cenário foi usado apenas para avaliar a técnica *GlyphMatrix* com o método *SUS*.

A seguir serão apresentados os protocolos e os resultados obtidos na avaliação de cada uma das três técnicas e da ferramenta *MGExplorer*.

6.1 Avaliação da *ClusterVis*

A técnica *ClusterVis* foi desenvolvida com o objetivo de visualizar as relações entre atributos de nodos pertencentes a agrupamentos em grafos multivariados. Para executar a avaliação nas dimensões utilidade e performance foram selecionadas as seguintes tarefas:

- Qual o nome do pesquisador que possui a maior quantidade de publicações na área Modelagem Conceitual e Banco de Dados?
- O autor com a maior quantidade de publicações é o mesmo que possui a maior quantidade de publicações do tipo *Journal*?
- O autor com a maior quantidade de publicações é coautor de autores classificados nas 3 categorias (discente, docente, participante externo)?

Interface: Fazem parte da interface utilizada para interagir com a técnica *ClusterVis* os seguintes mecanismos (Figura 6.2):

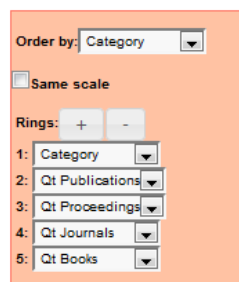


Figura 6.2 – Painel de ferramentas da *ClusterVis*

- *ComboBox (Order by)*: Proporciona a escolha de qual atributo será usado como critério de classificação;
- *CheckBox (Same Scale)*. Quando está marcado indica que todos os atributos numéricos representados nos anéis são visualizados com a mesma escala. Quando não estiver marcado indica que é usada uma escala individualizada para cada atributo;

- Botões “+” e “-“. São usados para acrescentar e remover os anéis que representam os atributos;
- Um *ComboBox* para cada anel viabiliza que o atributo representado nele possa ser alterado;
- *Tooltips* são ativados quando o cursor do mouse é posicionado sobre as barras com o objetivo de exibir informações mais detalhadas;
- Nodos adjacentes ao nodo que está sob o cursor do mouse são destacados dos demais e as arestas que os conectam são exibidas em vermelho.

Para a execução de cada uma das tarefas relacionadas anteriormente será considerado que a *ClusterVis* foi exibida na área de trabalho com os dados de coautoria do *cluster* formado por autores da linha de pesquisa Modelagem Conceitual e Banco de Dados do ano de 2011. A seguir serão apresentadas as avaliações nas dimensões utilidade, performance e satisfação.

6.1.1 Utilidade

Para avaliação na dimensão utilidade foi definida uma sequência de ações necessárias para a realização de cada tarefa. Nessa definição, foi levado em conta o estado da interface antes de cada uma delas. Considera-se que as 3 tarefas são executadas em sequência.

Tarefa 1 – Qual o nome do pesquisador que possui a maior quantidade de publicações na área Modelagem Conceitual e Banco de Dados?

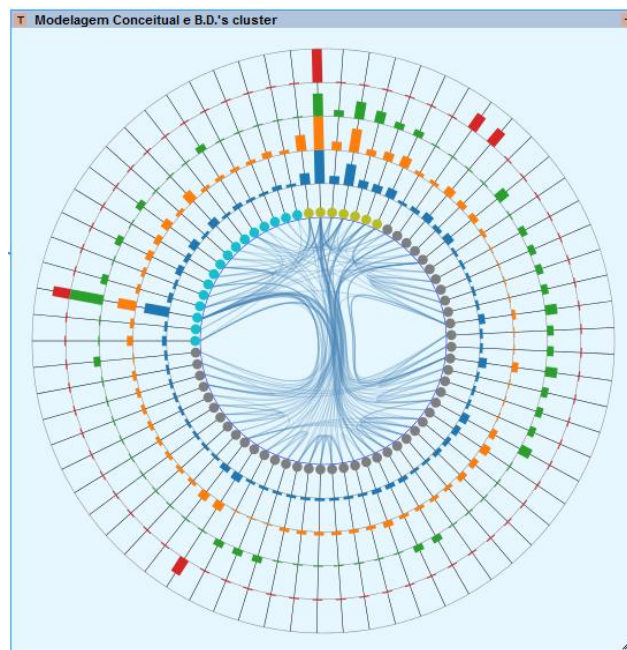


Figura 6.3 – Estado inicial da *ClusterVis* para a tarefa 1

Para a execução da tarefa 1 foi especificada a seguinte sequência de ações:

- 1) Clicar no botão **T** da barra de títulos para exibir o painel de ferramentas.

SUCESSO: No entanto existem duas observações sobre essa ação: (a) Caso a quantidade de nodos seja pequena e a diferença de tamanho entre as barras seja significativa pode não ser necessário acionar o painel de ferramentas. Neste caso pode-se passar diretamente para a ação 3. (b) Além da letra **T** o único elemento usado para identificar que um painel de ferramentas pode ser ativado é o *tooltip* vinculado ao botão. É recomendável que o usuário tenha conhecimento da existência do painel de ferramentas.

- 2) Selecionar o item *Qt. Publications* no *comboBox Order by* para classificar as barras por quantidade de publicações.

SUCESSO: A partir da exibição do painel de ferramentas a escolha do critério de classificação das barras é facilmente identificável.

- 3) Mover o painel de ferramentas para fora da área de visualização para possibilitar que todos os itens exibidos possam ser observados.

SUCESSO: O usuário pode inferir que o painel pode ser reposicionado livremente na área de trabalho observando o tipo de cursor do mouse quando ele é posicionado sobre o painel. Entretanto é exigido um conhecimento prévio do usuário do significado do cursor.

- 4) Identificar no painel de ferramentas em qual anel está representado o atributo quantidade de publicações.

FALHA: A associação do atributo com o número do anel é facilmente identificável no painel de ferramentas. Entretanto não há uma indicação da relação desse número com o anel na área de visualização. A utilização de cores para exibir os números localizados no painel de ferramentas iguais as utilizadas nas barras poderiam minimizar esse problema.

- 5) Mover o cursor do mouse sobre a maior barra do anel identificado na ação anterior para ativar o *tooltip*.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que a técnica utiliza *tooltips* para obter informações detalhadas, esta ação é executada corretamente. Observa-se que a única forma de obter informações numéricas ou textuais é com a ativação de *tooltips*.

Tarefa 2 – O autor com a maior quantidade de publicações é o mesmo que possui a maior quantidade de publicações do tipo *Journal*?

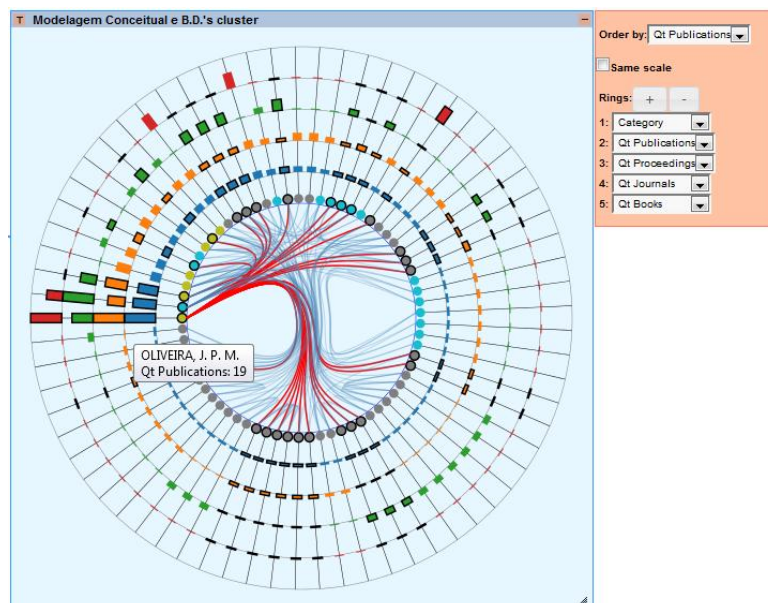


Figura 6.4 – Estado inicial da *ClusterVis* para a tarefa 2

Para a execução da tarefa 2 foi especificada a seguinte sequência de ações:

- 1) Selecionar o item *Qt. Journals* no *comboBox Order by* para classificar as barras por quantidade de *Journals*.

SUCESSO: A partir da exibição do painel de ferramentas a escolha do critério de classificação das barras é facilmente identificável.

- 2) Identificar no painel de ferramentas em qual anel está representado o atributo quantidade de publicações do tipo *Journals*.

FALHA: Esse é problema já identificado na tarefa 1, ação 4. A identificação do anel em que um determinado atributo está mapeado não é facilitada. No entanto quando se utiliza o mecanismo de ordenação em um determinado atributo, a identificação do respectivo anel é facilitada pela percepção da posição ordenada das barras.

- 3) Mover o cursor do mouse sobre a maior barra do anel identificado na ação anterior para ativar o *tooltip*.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que a técnica utiliza *tooltips* para obter informações detalhadas, esta ação é executada corretamente.

Tarefa 3 – O autor com a maior quantidade de publicações é coautor de autores classificados nas 3 categorias (discente, docente, participante externo)?

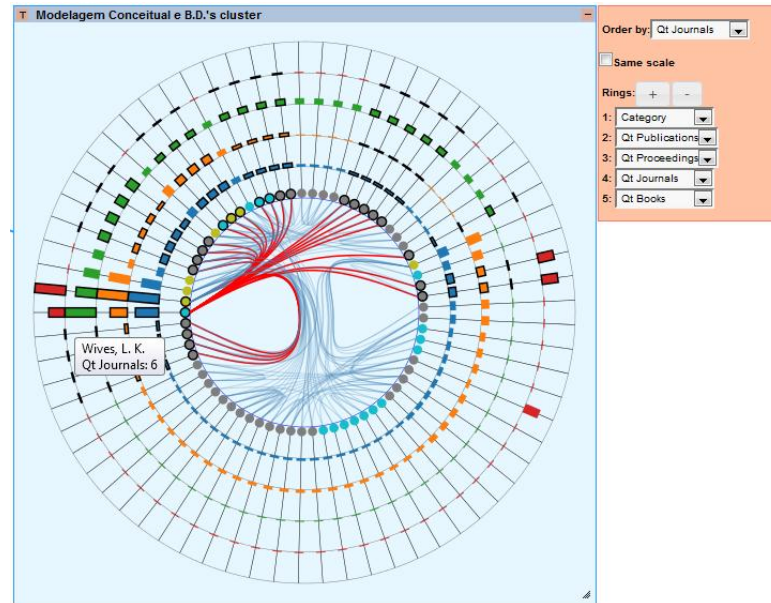


Figura 6.5 – Estado inicial da *ClusterVis* para a tarefa 3

Para a execução da tarefa 3 foi especificada a seguinte sequência de ações:

- 1) Selecionar o item *Qt. Publications* no *comboBox Order by* para classificar as barras por quantidade de publicações.

SUCESSO: A partir da exibição do painel de ferramentas a escolha do critério de classificação das barras é facilmente identificável.

- 2) Identificar no painel de ferramentas em qual anel está representado o atributo *Category*.

SUCESSO: O problema de identificação do anel em que um atributo é representado já foi mencionado na tarefa 1, ação 4. No entanto devido à representação diferenciada para atributos categóricos (círculos) essa identificação é facilitada.

- 3) Mover o cursor do mouse sobre a maior barra do anel que representa a quantidade de publicações para ativar o destaque das arestas.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que o posicionamento do cursor do mouse, sobre uma barra, ativa o destaque das arestas, a ação é executada corretamente.

- 4) Observar os nodos destacados pela ação anterior no anel que representa a categoria verificando a ocorrência das 3 categorias.

SUCESSO: Embora a observação seja facilitada pelo destaque das arestas e nodos, percebe-se a necessidade de observar os círculos que representam os tipos de categoria individualmente.

6.1.2 Performance

O mesmo conjunto de tarefas foi utilizado para a avaliação na dimensão performance. Para a definição da sequência de ações necessárias para executar cada tarefa, foi levado em conta o estado da interface antes de cada uma delas. Considera-se que as 3 tarefas são executadas em sequência.

Tarefa 1: Qual o nome do pesquisador que possui maior quantidade de publicações na área Modelagem Conceitual e Banco de Dados?

Tabela 6.1 – Detalhamento das ações para execução da tarefa 1 na *ClusterVis*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse sobre a borda da janela	P [borda]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse com o botão pressionado para redimensionar a janela	P [borda]	1,10
Mover o cursor do mouse sobre o botão <i>tools</i>	P [botão]	1,10
Clicar com o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse sobre o painel de ferramentas	P [painel]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse reposicionando o painel de ferramentas para fora da área de visualização da técnica	P [painel]	1,10
Mover o cursor do mouse para o <i>comboBox Order by</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>QtPublications</i>	P [<i>QtPublications</i>]	1,10
Clicar no item <i>QtPublications</i>	K [<i>QtPublications</i>]	0,23
Identificar o anel que representa a quantidade de publicações	M [anel]	1,35
Identificar a maior barra do anel	M [barra]	1,35
Mover o cursor do mouse sobre a barra identificada no passo anterior ativando o <i>tooltip</i>	P [barra]	1,10
Identificar o nome do autor exibido pelo <i>tooltip</i>	M [<i>tooltip</i>]	1,35
Total		14,36

Tarefa 2: O autor com a maior quantidade de publicações é o mesmo que possui a maior quantidade de publicações do tipo *Journal*?

Tabela 6.2 – Detalhamento das ações para execução da tarefa 2 na *ClusterVis*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o <i>comboBox Order by</i> do painel de ferramentas	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>QtJournals</i>	P [<i>QtJournals</i>]	1,10
Clicar no item <i>QtJournals</i>	K [<i>QtJournals</i>]	0,23
Identificar o anel que representa a quantidade de publicações do tipo <i>Journal</i>	M [anel]	1,35
Identificar a maior barra do anel	M [barra]	1,35
Mover o cursor do mouse sobre a barra identificada no passo anterior ativando o <i>tooltip</i>	P [barra]	1,10
Identificar o nome do autor exibido pelo <i>tooltip</i>	M [<i>tooltip</i>]	1,35
Identificar se o nome do autor é o mesmo do obtido na tarefa anterior	M [autor]	1,35
Total		9,52

Tarefa 3: O autor com a maior quantidade de publicações é coautor de autores classificados nas 3 categorias (discente, docente, participante externo)?

Tabela 6.3 – Detalhamento das ações para execução da tarefa 3 na *ClusterVis*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o <i>comboBox Order by</i> do painel de ferramentas	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>QtPublications</i>	P [<i>QtPublications</i>]	1,10
Clicar no item <i>QtPublications</i>	K [<i>QtPublications</i>]	0,23
Identificar o anel que representa a quantidade de publicações	M [anel]	1,35
Identificar a maior barra do anel	M [barra]	1,35
Identificar o anel que representa as categorias	M [barra]	1,35
Mover o cursor do mouse sobre a barra identificada no passo anterior destacando os <i>links</i>	P [barra]	1,10
Verificar se foram destacados círculos de três cores diferentes (categorias)	M [círculo]	1,35
Total		9,52

6.1.3 Satisfação

Na avaliação da *ClusterVis* para a dimensão satisfação, foi utilizado o mesmo conjunto de dados mencionado anteriormente. Foram disponibilizados os seguintes dados associados aos nodos: quantidade de publicações divididas em subtipos (*proceedings*, *journals* e *books*), categoria do autor (discente, docente e participante externo), área de pesquisa do autor e quantidade de coautores (*degree*). A figura 6.6 mostra o estado inicial da aplicação que foi usada para avaliar a *ClusterVis*.

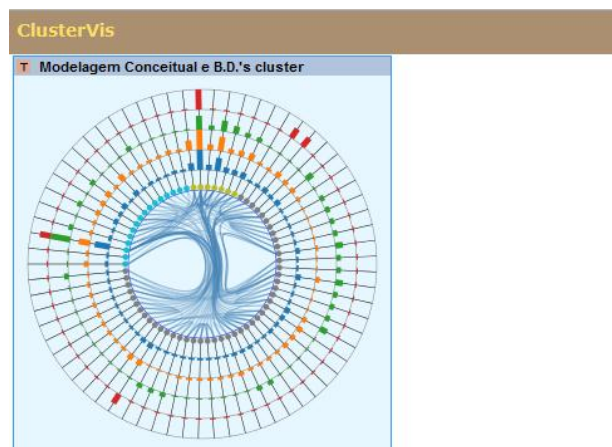


Figura 6.6 – Aplicação usada na avaliação da *ClusterVis*

6.1.3.1 Participantes

De forma voluntária, participaram do experimento 19 pessoas, todas do sexo masculino com idade distribuída nas seguintes faixas etárias: 85,7% tinham de 21 a 30 anos, 9,5% de 31 a 40 anos e somente 4,8% com idade entre 41 a 50 anos. Quanto à escolaridade,

33,3% eram alunos de graduação, 57,1% eram graduados e 9,6%, mestres. A grande maioria possuía experiência com sistemas *Web* (85,7%) e 61,9% declararam possuir experiência como usuários de sistemas interativos aplicados a jogos computacionais.

6.1.3.2 Procedimento

A avaliação foi conduzida na forma de “*survey*” remoto, com o objetivo de avaliar:

- a) os mecanismos de interação disponíveis na técnica;
- b) o poder de expressão dos atributos visualizados em cadaanel assim como dos elementos adicionais de informação (“*tooltips*”);
- c) a possibilidade de modificar a visualização segundo diversos critérios de classificação.

Os participantes, alunos da disciplina de visualização em 2016/1, foram convidados a realizar a avaliação remota através de mensagem de correio eletrônico, contendo instruções. Os participantes foram instruídos a responder inicialmente o questionário de perfil de usuário. Em sequência, os respondentes foram convidados a utilizar a técnica, durante 15 minutos, sendo orientados a exercitar todas as funcionalidades da ferramenta. Passado o tempo de experimentação, os voluntários foram convidados a responder um questionário específico sobre as funcionalidades das técnicas (ver resultados na Seção 6.1.3.3) e, em sequência, a responder o questionário de avaliação SUS. Tanto a lista de tarefas como os questionários (perfil e *SUS*) encontram-se no Apêndice A.

6.1.3.3 Resultados

Como se pode observar na tabela 6.4, todas as tarefas propostas para a avaliação da *ClusterVis* foram consideradas fáceis de serem executadas pela grande maioria (acima de 80%) dos participantes.

Tabela 6.4 – Análise das tarefas

Questões	Concordo	Neutro	Discordo
É fácil descobrir o autor com maior quantidade de publicações da área Modelagem Conceitual e Banco de Dados.	17 89,5%	2 10,5%	0 0%
É fácil descobrir se o autor com maior quantidade de publicações é o mesmo que possui maior quantidade de publicações do tipo <i>Journal</i> .	17 89,5%	2 10,5%	0 0%
É fácil descobrir se o autor com maior quantidade de publicações é coautor do autor com a segunda maior quantidade de publicações.	16 84,2%	2 10,5%	1 5,3%
É fácil descobrir se o autor com a maior quantidade de publicações é coautor de autores classificados nas 3 categorias (discente, docente, participante externo).	16 84,2%	1 5,3%	2 10,5%
É fácil descobrir entre os docentes coautores do autor com a maior quantidade de publicações, qual possui a menor quantidade de publicações.	17 89,5%	0 0%	2 10,5%

Dentre os comentários dos usuários destaca-se a dificuldade de se obter informações de nodos relacionados entre si, em razão de que um único *tooltip* é exibido por vez. O mesmo ocorre com a obtenção dos nomes de atributos categóricos (mapeados para cor), já que não há uma indicação (além do *tooltip*) do significado de cada cor.

Em relação às questões aplicadas para o método SUS foi obtido um escore médio de 76,7 com o um valor mínimo igual a 40,0 e máximo igual a 92,5.

6.2 Avaliação da *GlyphMatrix*

A técnica *GlyphMatrix* foi desenvolvida com o objetivo de visualizar os atributos de arestas pertencentes a agrupamentos em grafos multivariados. Para executar a avaliação foram selecionadas as seguintes tarefas:

- Existe algum coautor que publica com a maioria dos coautores do pesquisador Anderson Maciel?
- Qual a quantidade de cada tipo de publicação entre a coautoria dos autores FREITAS, C. M. D. S e NEDEL, L. P.?

Interface: Fazem parte da interface utilizada para interagir com a técnica *GlyphMatrix* os seguintes mecanismos (Figura 6.7):

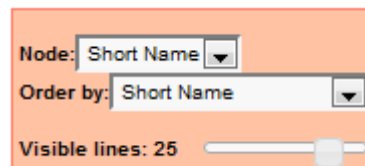


Figura 6.7 – Painel de ferramentas da *GlyphMatrix*

- *ComboBox (Node)*: Proporciona a escolha de qual atributo será exibido nas barras que representam os nodos;
- *ComboBox (Order by)*: Permite a escolha de qual atributo será usado como critério de classificação das linhas e colunas da matriz;
- *Slider (Visible lines)*: Possibilita a determinação da quantidade de células visíveis em uma linha (ou coluna). Reduzir a quantidade de linhas visíveis produz o efeito de aumentar as dimensões das células da matriz;
- *Panning*: O arrasto do mouse sobre os textos que identificam os nodos ou sobre as células da matriz permite que esta seja movida de modo que o usuário possa controlar quais nodos estarão visíveis;

- *Tooltips* são ativados quando o cursor do mouse é posicionado sobre as células com o objetivo de exibir os glifos com um maior detalhamento.

Para a execução de cada uma das tarefas relacionadas acima será considerado que a *GlyphMatrix* foi exibida na área de trabalho com os dados do *cluster* formado pelo pesquisador Anderson Maciel e seus coautores da linha de pesquisa Computação Gráfica e Processamento de Imagem no período de 2004 a 2011. A seguir serão apresentadas as avaliações nas dimensões utilidade, performance e satisfação.

6.2.1 Utilidade

Para avaliação na dimensão utilidade foi definida uma sequência de ações necessárias para a realização de cada tarefa. Nessa definição, foi levado em conta o estado da interface antes de cada uma delas. Considera-se que as 2 tarefas são executadas em sequência.

Tarefa 1 – Existe algum coautor que publica com a maioria dos coautores do pesquisador Anderson Maciel?

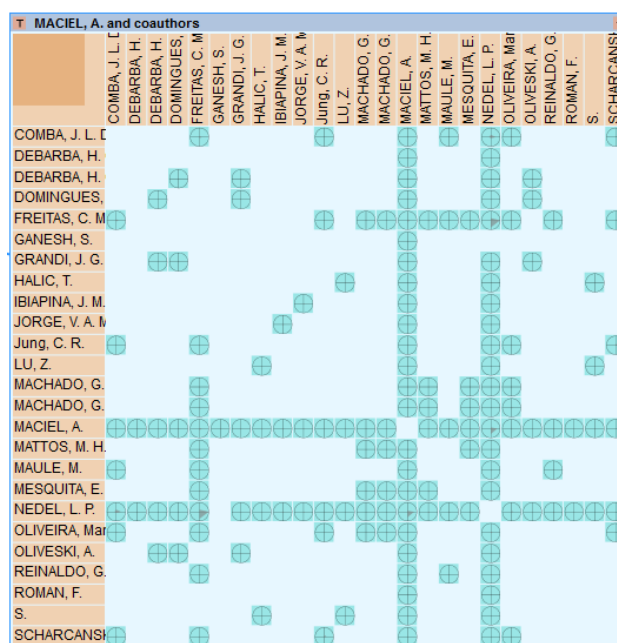


Figura 6.8 – Estado inicial da *GlyphMatrix* para as tarefas 1 e 2

Para execução da tarefa 1 foi especificada a seguinte sequência de ações:

- 1) Redimensionar a janela de visualização.

SUCESSO: O redimensionamento da janela de visualização é facilmente identificável. A execução dessa ação é necessária para aumentar as dimensões das células quando a quantidade de células da matriz alcança um valor considerável.

- 2) Clicar no botão **T** da barra de títulos para exibir o painel de ferramentas.

SUCESSO: Além da letra **T** o único elemento usado para identificar que um painel de ferramentas pode ser ativado é o *tooltip* vinculado ao botão. É recomendável que o usuário tenha conhecimento da existência do painel de ferramentas.

- 3) Tornar todas as linhas da matriz visíveis.

SUCESSO: O componente de interface (*Slider*) que permite executar essa ação é facilmente identificável.

- 4) Localizar o nome do autor Anderson Maciel na barra que contém os nomes dos nodos.

SUCESSO: A identificação do autor pode ser realizada de duas formas distintas. Observando as barras com os nomes dos autores classificados em ordem alfabética ou a linha que possui todas as células marcadas.

- 5) Observar a existência de uma linha ou coluna da matriz com as mesmas células marcadas do pesquisador Anderson Maciel

SUCESSO: A localização da linha com as características mencionadas nesta ação é facilmente observada se houver uma quantidade considerável de células em comum.

Tarefa 2 – Qual a quantidade de cada tipo de publicação entre a coautoria dos autores FREITAS, C. M. D. S e NEDEL, L. P.?

Para a execução da tarefa 2 foi especificada a seguinte sequência de ações:

- 1) Localizar a célula cuja linha e coluna pertencem aos autores FREITAS, C. M. D. S. e NEDEL, L. P.

SUCESSO: A localização da célula pode ser realizada facilmente considerando que os nomes dos autores estejam classificados em ordem alfabética.

- 2) Mover o cursor do mouse sobre a célula identificada no passo anterior para exibir o glifo que contém as informações desejadas.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que o posicionamento do cursor do mouse sobre uma célula ativa o *tooltip* com o detalhamento do glifo esta ação é executada com facilidade.

6.2.2 Performance

O mesmo conjunto de tarefas foi utilizado para a avaliação na dimensão performance. Para a definição da sequência de ações necessárias para executar cada tarefa, foi levado em conta o estado da interface antes de cada uma delas. Considera-se que as 2 tarefas são executadas em sequência.

Tarefa 1: Existe algum coautor que publica com a maioria dos coautores do pesquisador Anderson Maciel?

Tabela 6.5 – Detalhamento das ações para execução da tarefa 1 na *GlyphMatrix*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse sobre a borda da janela	P [borda]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse com o botão pressionado para redimensionar a janela	P [borda]	1,10
Verificar se existe uma linha na matriz que possui uma grande quantidade de células marcadas nas mesmas colunas do autor alvo.	M [células]	1,35
Total		4,14

Tarefa 2: Qual a quantidade de cada tipo de publicação entre a coautoria das autoras FREITAS, C.M.D. e NEDEL, L.P.?

Tabela 6.6 – Detalhamento das ações para execução da tarefa 2 na *GlyphMatrix*

Descrição	Operação	Tempo (seg.)
Identificar a linha que representa a autora FREITAS, C.M.D.	M [linha]	1,35
Identificar a linha que representa a autora NEDEL, L.P.	M [coluna]	1,35
Identificar a célula localizada na interseção da linha e coluna que representa a coautoria entre as duas autoras.	M [célula]	1,35
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para a célula ativando o <i>tooltip</i>	P [célula]	1,10
Identificar a quantidade de cada tipo de publicação exibido pelo <i>tooltip</i>	M [<i>tooltip</i>]	1,35
Total		6,86

6.2.3 Satisfação

Para a avaliação da *GlyphMatrix*, foi utilizada a implementação original da técnica, com dados do Campeonato Brasileiro de Futebol de 2013 (Figura 6.9).

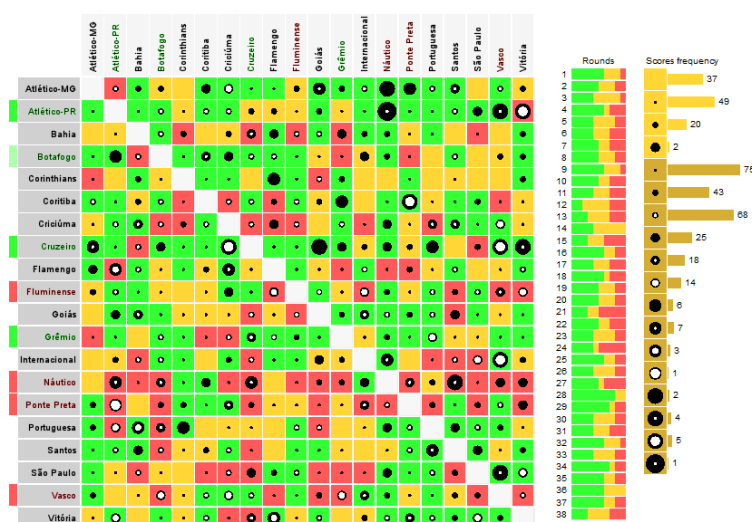


Figura 6.9 – Visualização dos resultados de partidas de futebol com a *GlyphMatrix*

Os dados disponíveis correspondem a descritores de todos os jogos nas 38 rodadas, com informações sobre: time jogando em casa, time visitante, escore da partida e, para cada time, número de faltas, número de impedimentos, número de passes errados, número de assistências e número de chutes a gol. O experimento com essa técnica foi conduzido para avaliar tanto o seu poder de expressão global como a facilidade de encontrar informações usando os glifos relativos aos escores e às estatísticas.

6.2.3.1 *Participantes*

De forma voluntária, participaram do experimento 13 alunos de pós-graduação em Computação, alunos da disciplina de visualização, todos do sexo masculino e com idades variando entre 23 e 28 anos. Além disso, todos tinham mais de 10 anos de experiência como usuários de jogos computacionais, mas tinham pouca ou nenhuma experiência em aplicações de visualização. Dos 13 participantes, apenas 2 declararam não ter experiência com desenvolvimento envolvendo técnicas de computação gráfica.

6.2.3.2 *Procedimento*

Os participantes foram reunidos num laboratório com máquinas iguais para realizar a sessão de avaliação. Inicialmente, foi aplicado o questionário de perfil e, na sequência, foram apresentados o propósito do experimento e as instruções para sua realização. Juntamente com as instruções, foi entregue uma lista de questões a serem respondidas com o uso da aplicação. Foi sugerido um tempo de experimentação livre da aplicação para conhecimento geral da interface, antes da realização das tarefas propriamente ditas. A cada questão respondida com o auxílio da técnica, o participante deveria dizer se a tarefa havia sido realizada com facilidade ou não, seguindo uma escala *Likert* de 5 pontos. Findo o experimento, foi solicitado, também, o preenchimento de um questionário SUS. Tanto a lista de tarefas como os questionários (perfil e *SUS*) encontram-se no Apêndice A.

6.2.3.3 *Resultados*

A primeira questão “Existem semelhanças na ordenação dos times na vertical segundo os diferentes critérios possíveis?” foi considerada fácil de responder pela maioria dos participantes (11 participantes, 84,6%) enquanto 1 foi neutro e outro a considerou difícil.

Para as questões envolvendo a análise do resultado geral dos jogos (vitória do time da casa, derrota do time da casa e empate), que requeriam apenas a interpretação das cores das células da matriz, a grande maioria achou fácil (Tabela 6.7).

Tabela 6.7 – Análise do resultado geral dos jogos

Questões	Concordo	Neutro	Discordo
É fácil descobrir quais são as rodadas em que ocorreu apenas 1 empate	10 76,9%	1 7,7%	2 15,4%
É fácil descobrir quais são as rodadas em quais ocorreu apenas 1 derrota de time da casa	11 84,6%	0 0%	2 15,4%
É fácil descobrir quais são as rodadas em que não ocorreram derrotas do time da casa	12 92,3%	0 0%	1 7,7%

Já a análise do glifo dos escores, questão “É fácil descobrir em quais rodadas e jogos ocorreram os resultados com maior diferença de gol?”, teve resultados mais distribuídos: 9 participantes (69,2%) acharam fácil, 3 participantes (23,1%) foram neutros e 1 participante achou difícil.

Por outro lado, o glifo utilizado para representar as estatísticas demonstrou ser mais difícil de interpretar (Tabela 6.8). As questões relativas às informações apresentadas nesses glifos foram avaliadas pelos participantes como mais difíceis.

Tabela 6.8 – Análise das estatísticas dos jogos

Questões	Concordo	Neutro	Discordo
É fácil descobrir em quais jogos o time da casa cometeu o maior número de faltas	2 15,4%	1 7,7%	10 76,9%
É fácil descobrir em quais jogos o time visitante cometeu o maior número de faltas	2 15,4%	2 15,4%	9 69,2%
É fácil descobrir em quais jogos não ocorreram impedimentos (<i>Offsides</i>)	4 30,8%	2 15,4%	7 53,8%
É fácil descobrir para um time específico em quais jogos o número de faltas supera o número de desarmes (<i>Tackles</i>).	3 23,1%	4 30,8%	6 46,1%

Com relação aos comentários do usuário destacam-se dificuldade de interpretação dos glifos usados para representar dados estatísticos (gráficos de barras) em função do tamanho reduzido das células. Cabe salientar que essa versão da *GlyphMatrix* não permitia a alteração da quantidade de linhas visíveis. Outro ponto destacado por alguns usuários está na ausência de filtros que permitam destacar as células por diferentes critérios. O único filtro disponível na implementação destacava apenas resultados de uma rodada.

Em relação às questões aplicadas para o método SUS, foi obtido um escore médio de 63,5 com o um valor mínimo igual a 27,5 e máximo igual a 87,5.

6.3 Avaliação da *Iris*

A técnica *Iris* foi desenvolvida com o objetivo de visualizar os atributos pertencentes a arestas que conectam um grupo de nodos a um determinado nodo de interesse. Para executar a avaliação foram selecionadas as seguintes tarefas:

- Qual o nome do coautor com a maior quantidade de publicações em coautoria com a autora Lucina Nedel?
- Qual a quantidade de publicações em coautoria com o autor FISCHER, L. G.?
- Qual o nome do coautor com a maior quantidade de publicações do tipo *Journals* em coautoria com a autora Luciana Nedel?

Interface: Fazem parte da interface utilizada para interagir com a técnica *Iris* os seguintes mecanismos (Figura 6.10):

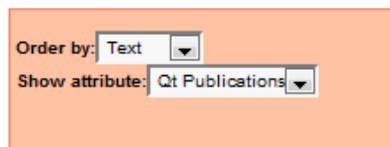


Figura 6.10 – Paineis de ferramentas da *Iris*

- *ComboBox (Order by)*: Proporciona a escolha de qual elemento será classificado (texto ou atributo);
- *ComboBox (Show attribute)*: Permite a escolha do atributo que será mapeado para as barras;
- Um clique com o botão esquerdo do mouse sobre um texto ou uma barra faz com que ela seja posicionada no foco;
- *Tooltips* são ativados quando o cursor do mouse é posicionado sobre as barras ou texto com o objetivo de exibir informações mais detalhadas.

Para a execução de cada uma das tarefas relacionadas acima será considerado que a *Iris* foi exibida na área de trabalho com os dados de coautoria do ano de 2011 da pesquisadora Luciana Nedel. A seguir serão apresentadas as avaliações nas dimensões utilidade, performance e satisfação.

6.3.1 Utilidade

Para avaliação na dimensão utilidade foi definida uma sequência de ações necessárias para a realização de cada tarefa. Nessa definição, foi levado em conta o estado da interface antes de cada uma delas. Considera-se que as 3 tarefas são executadas em sequência.

Tarefa 1 – Qual o nome do coautor com a maior quantidade de publicações em coautoria com a autora Lucina Nedel?

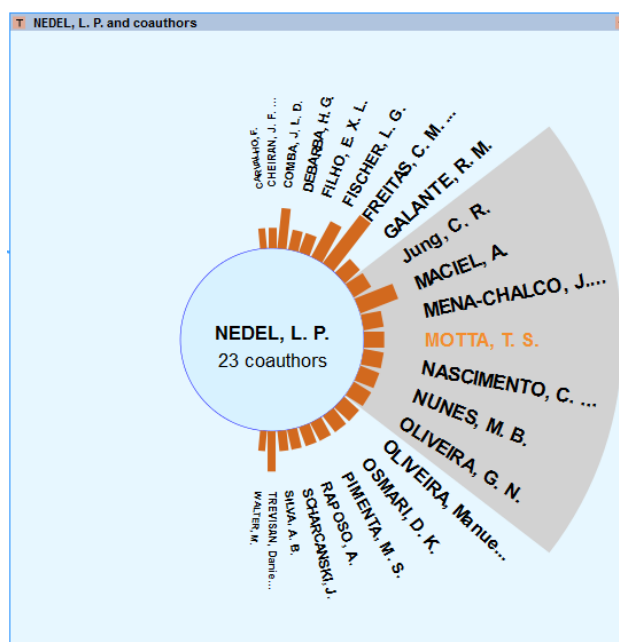


Figura 6.11 – Estado inicial da *Iris* para a tarefa 1

Para a execução da tarefa 1 foi especificada a seguinte sequência de ações:

- 1) Clicar no botão **T** da barra de títulos para exibir o painel de ferramentas.

SUCESSO: No entanto existem duas observações sobre essa ação: (a) Caso a quantidade de nodos seja pequena e a diferença de tamanho entre as barras seja significativa pode não ser necessário acionar o painel de ferramentas. Neste caso pode-se passar diretamente para a ação 3. (b) Além da letra **T** o único elemento usado para identificar que um painel de ferramentas pode ser ativado é o *tooltip* vinculado ao botão. É recomendável que o usuário tenha conhecimento da existência do painel de ferramentas.

- 2) Selecionar o item *Attribute* no *comboBox Order by* para classificar as barras por quantidade de publicações.

SUCESSO: A partir da exibição do painel de ferramentas a escolha do critério de classificação das barras é facilmente identificável.

- 3) Clicar na maior barra para posicioná-la no foco.

SUCESSO: A barra deve ser posicionada para o foco para facilitar a leitura do texto. Embora haja uma mudança do ícone do mouse quando este passa sobre uma barra mostrando que ela pode ser clicada, não há uma indicação clara que nessa situação a barra será movida para o foco. Apenas com a experimentação é possível perceber a existência deste mecanismo de interação.

6.3.2 Performance

O mesmo conjunto de tarefas foi utilizado para a avaliação na dimensão performance. Para a definição da sequência de ações necessárias para executar cada tarefa, foi levado em conta o estado da interface antes de cada uma delas. Considera-se que as 3 tarefas são executadas em sequência.

Tarefa 1: Qual o nome do coautor com maior quantidade de publicações em coautoria com a autora Lucina Nedel?

Tabela 6.9 – Detalhamento das ações para execução da tarefa 1 na *Iris*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse sobre o botão <i>tools</i>	P [botão]	1,10
Clicar com o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse para o <i>comboBox Order by</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>Attribute</i>	P [<i>Attribute</i>]	1,10
Clicar no item <i>Attribute</i>	K [<i>Attribute</i>]	0,23
Identificar a maior barra	M [barra]	1,35
Mover o cursor do mouse sobre a maior barra encontrada	P [barra]	1,10
Clicar na barra encontrada para posicioná-la no foco	K [mouse]	0,23
Mover o cursor do mouse sobre o foco ativando o <i>tooltip</i>	P [foco]	1,10
Identificar a quantidade de publicações exibidas pelo <i>tooltip</i>	M [tooltip]	1,35
Total		9,48

Tarefa 2: Qual a quantidade de publicações geradas pela coautoria das autoras Luciana Nedel e FISCHER, L. G.?

Tabela 6.10 – Detalhamento das ações para execução da tarefa 2 na *Iris*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse sobre o painel de ferramentas	P [painel]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse reposicionando o painel de ferramentas para fora da área de visualização da técnica	P [painel]	1,10
Mover o cursor do mouse para o <i>comboBox Order by</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>Text</i>	P [<i>Text</i>]	1,10
Clicar no item <i>Text</i>	K [<i>Text</i>]	0,23
Identificar um nome de autor que começa com a letra F	M [letra F]	1,35
Clicar no nome do autor para movê-lo para o foco	K [autor]	0,23
Localizar o autor alvo (FISCHER, L. G.)	M [autor]	1,35
Mover o cursor do mouse sobre o autor identificado no passo anterior ativando o <i>tooltip</i>	P [autor]	1,10
Identificar a quantidade de publicações exibidas pelo <i>tooltip</i>	M [tooltip]	1,35
Total		10,83

Tarefa 3: Qual o nome do coautor com a maior quantidade de publicações do tipo *Journal* em coautoria com a autora Luciana Nedel?

Tabela 6.11 – Detalhamento das ações para execução da tarefa 3 na *Iris*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o <i>comboBox Show attribute</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>Qt Journals</i>	P [<i>Qt Journals</i>]	1,10
Clicar no item <i>Qt Journals</i>	K [<i>Qt Journals</i>]	0,23
Mover o cursor do mouse para o <i>comboBox Order by</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>Attribute</i>	P [<i>Attribute</i>]	1,10
Clicar no item <i>Attribute</i>	K [<i>Attribute</i>]	0,23
Identificar a maior barra	M [barra]	1,35
Mover o cursor do mouse sobre a maior barra encontrada	P [barra]	1,10
Clicar na barra encontrada para posicioná-la no foco	K [mouse]	0,23
Mover o cursor do mouse sobre o foco ativando o <i>tooltip</i>	P [foco]	1,10
Identificar o nome do autor exibido pelo <i>tooltip</i>	M [tooltip]	1,35
Total		10,81

6.3.3 Satisfação

Na avaliação da dimensão satisfação, a técnica *Iris* foi usada para visualizar os resultados de buscas executadas na *Web*. O cenário utilizado no experimento de avaliação consistia de uma lista de pesquisadores exibida em uma página *web*, na qual um usuário podia consultar seus coautores em alguma base de dados bibliográfica, como, por exemplo, a DBLP. A partir da seleção do nome de um pesquisador, uma busca pré-definida (coautores) é executada, exibindo os resultados em uma *Iris* embutida dentro da página *Web* (Figura 6.14).

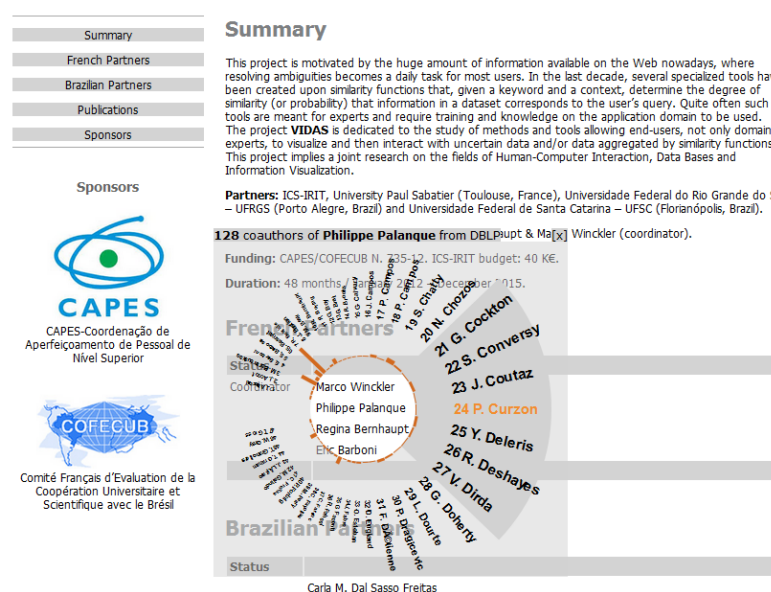


Figura 6.14 – *Iris* inserida em uma página *Web*

6.3.3.1 Participantes

Participaram do experimento 61 voluntários distribuídos entre professores, pesquisadores e estudantes envolvidos em atividades de pesquisa. A maioria dos participantes era do sexo masculino (77,1%) com uma idade média de 26,1 (DP=5,7) anos distribuídos nas seguintes faixas etárias: 55,7% de 25 a 34 anos, 24,6% de 35 a 54 anos, e 19,7% de 18 a 24 anos. Em relação às atividades profissionais, 44,3% eram estudantes, 39,3% eram pesquisadores e/ou professores e 16,4%, trabalhadores da indústria. Quanto à titulação, os participantes distribuíam-se em PhD (26,2%), Mestres (44,3%), Graduados (19,7%) e estudantes de graduação (9,8%). Os participantes estimaram gastar em média 5,3 horas (DP=4,1) por semana usando mecanismos de busca na *Web*. Desse gasto médio, ~4,3 horas por semana (DP=3,9) eram gastas na busca de autores e publicações.

6.3.2.2 Procedimento

Para a realização do experimento, foi criado um documento online com a apresentação de como uma busca podia ser executada e a forma de ativação da *Iris*. Usando a escala *Likert* para obter as respostas, os usuários foram questionados sobre o uso de uma técnica de visualização para exibir os resultados de consultas na *Web* e sobre as informações que podem ser obtidas com a *Iris*.

6.3.2.3 Resultados

Em relação ao uso de uma técnica de visualização como uma solução alternativa para exibir os resultados de consultas auxiliares na *Web*, 51 participantes (83,6%) acharam a *Iris* útil, 44 participantes (72,2%) consideraram que ela melhora o desempenho na realização da consulta e 50 (82%) apreciaram utilizar a *Iris*. Quanto à forma de exibir os resultados da consulta, 38 participantes (62,3%) disse preferir ver os resultados com a *Iris* ao invés de uma página HTML. A tabela 6.12 apresenta cada uma das questões e seus respectivos resultados.

Tabela 6.12 –Técnica de visualização para exibir resultados de uma consulta

Questões	Concordo	Neutro	Discordo
Eu acho útil mostrar os resultados de uma consulta sobre a <i>Web</i> através de uma técnica de visualização interativa	51 83,6%	8 13,1%	2 3,3%
Eu acho que mostrar os resultados de uma consulta sobre a <i>Web</i> através de uma técnica de visualização interativa melhora meu desempenho (com respeito aos mecanismos de busca tradicionais).	44 72,2%	13 21,3%	4 6,6%
Eu gosto de poder usar uma técnica de visualização interativa para explorar resultados de busca sobre a <i>Web</i>	50 82%	9 14,7	2 3,3%
Eu prefiro ver meus resultados da consulta como uma página HTML ao invés de usar uma técnica de visualização interativa	7 11,5%	16 26,2%	38 62,3%

Tabela 6.13 – Uso da *Iris* para a visualização das informações de coautoria

Questões	Concordo	Neutro	Discordo
Com a técnica <i>Iris</i> eu posso obter facilmente a quantidade de coautores do pesquisador selecionado.	49 77%	9 14,8%	5 8,2%
Com a técnica <i>Iris</i> eu posso obter facilmente a quantidade de publicações geradas pela coautoria entre o pesquisador selecionado e um dado coautor.	36 59%	9 14,8%	16 26,2%
Com a técnica <i>Iris</i> eu posso localizar facilmente um determinado coautor do pesquisador selecionado.	40 65,6%	10 16,4%	11 18%
Com a técnica <i>Iris</i> eu posso localizar facilmente o coautor com o maior número de publicações em coautoria com o pesquisador selecionado.	45 73,8%	6 9,8%	10 16,4%

Considerando as informações das relações existentes entre um pesquisador e seus coautores, exibidas pela *Iris*, a maioria dos participantes considerou fácil obtê-las. Pode-se citar como exemplo informações como a quantidade de coautores do pesquisador selecionado e a quantidade de publicações geradas pela coautoria entre ele e um autor dado qualquer. A tabela 6.13 mostra as questões e seus respectivos resultados.

6.4 Avaliação da *MGExplorer*

A ferramenta *MGExplorer* foi desenvolvida como prova de conceito da abordagem proposta neste trabalho, proporcionando a instanciação de múltiplas técnicas de visualização em janelas redimensionáveis. Nesta seção será apresentada a avaliação da ferramenta apenas nas dimensões utilidade e performance.

Para executar as avaliações foram selecionadas as seguintes **tarefas**:

- Qual o coautor do autor Marcelo Soares Pimenta possui mais publicações em 2011?
- Qual a quantidade de publicações geradas pela coautoria entre Marcelo Soares Pimenta e o autor obtido na tarefa anterior?
- Considerando as relações de coautoria entre os elementos do *cluster* formado pelo autor Claudio Jung e seus coautores, qual coautoria (2 autores) gerou maior quantidade de publicações?

6.4.1 Utilidade

A seguir será apresentada a avaliação da *MGExplorer* na dimensão utilidade. Inicialmente são apresentados os mecanismos usados para interagir com o diagrama de nodos e arestas, com a visualização do histórico e os mecanismos de uso geral. E, por fim, a descrição das ações necessárias para a execução das tarefas.

6.4.1.1 Mecanismos de interação com o diagrama de nodos e arestas:

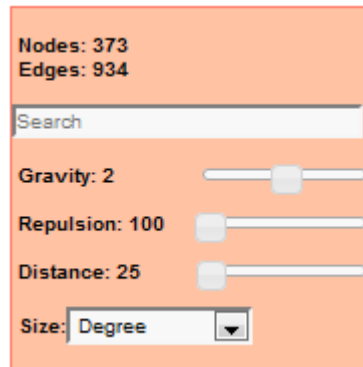


Figura 6.15 – Painel de ferramentas da técnica que exibe o diagrama de nodos e arestas

- Caixa de texto: Possibilita a execução de uma busca para localizar no grafo o nodo que possui como atributo o nome do autor informado;
- *Slider (Gravity)*: Permite ajustar o valor da gravidade usado na geração do *layout* do grafo. Quanto maior o valor, mais próximos estarão os nodos entre si;
- *Slider (Repulsion)*: Permite ajustar o valor da repulsão usada na geração do *layout* do grafo. Quanto maior o valor, mais afastados estarão os nodos entre si.
- *Slider (Distance)*: Possibilita determinar o comprimento das arestas utilizadas na geração do *layout* do grafo.
- *ComboBox (Size)*: Proporciona a escolha de qual atributo será usado para mapear o diâmetro dos círculos que representam os nodos.

6.4.1.2 Mecanismos de interação com o histórico:

- O arrasto do mouse sobre o conteúdo do histórico permite move-lo dentro da janela na qual está contido;
- Um clique com o botão esquerdo do mouse sobre os quadrados permitem ocultar/exibir a respectiva janela de visualização.

6.4.1.3 Mecanismos de interação de uso geral:

- *ComboBox (Select to load)*: Proporciona a escolha do conjunto de dados que será carregado no diagrama de nodos e arestas (Figura 6.16);
- *ComboBox (Cluster by)*: Proporciona a escolha do algoritmo a ser usado na clusterização do grafo exibido através do diagrama de nodos e arestas (Figura 6.16);

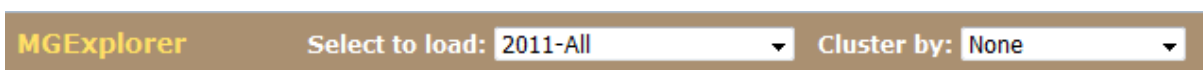


Figura 6.16 – Barra de seleção

- Um clique com um botão direito sobre um nodo ou aresta aciona um menu de contexto (Figura 6.17) que permite escolher a técnica que será utilizada para exibir um subgrafo. Esse subgrafo é constituído pelo autor selecionado e seus coautores;

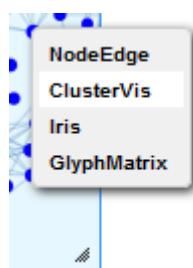


Figura 6.17 – Menu de contexto para ativações de outras técnicas de visualização

- O arrasto do mouse sobre o quadrado cinza que representa uma janela oculta permite reposicioná-lo na área de trabalho.

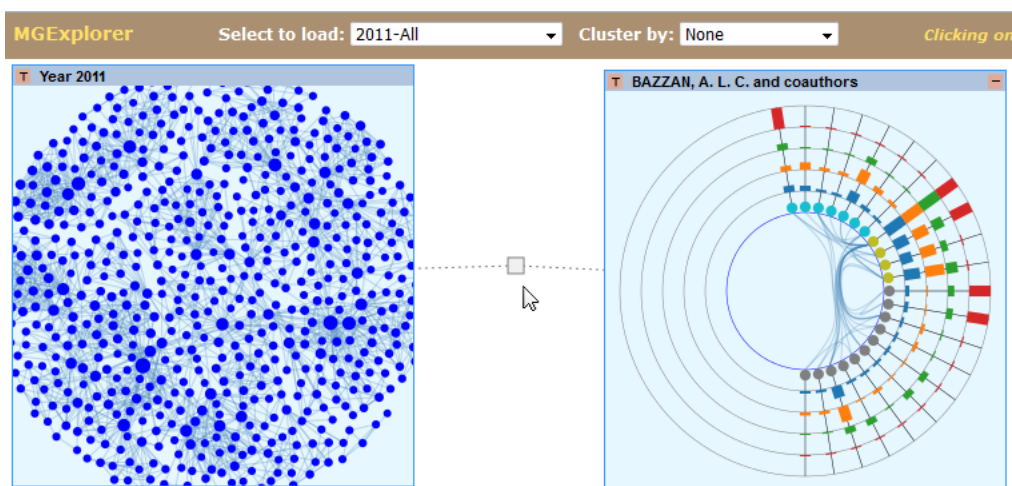


Figura 6.18 – Representação de uma janela oculta

6.4.1.4 Avaliação

Para a execução de cada uma das tarefas relacionadas anteriormente será considerado que a *MGExplorer* foi carregada com o grafo de coautoria das publicações realizadas no ano de 2011. Na definição da sequência de ações necessárias para a realização de cada tarefa foi levado em conta o estado da área de trabalho antes de cada uma delas. Considera-se que as 4 tarefas são executadas em sequência.

Tarefa 1 – Qual o coautor do autor Marcelo Soares Pimenta possui mais publicações em 2011?

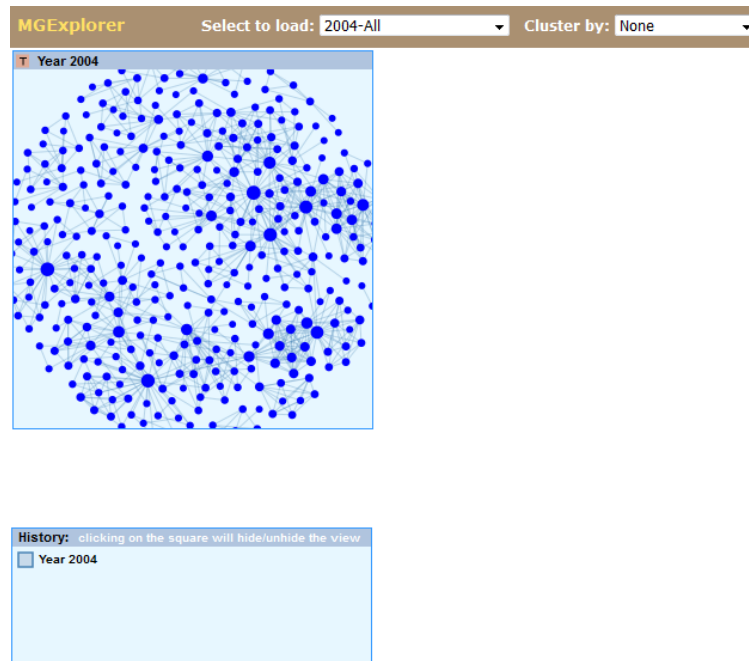


Figura 6.19 –Estado inicial da *MGExplorer* para a tarefa 1

Para a execução da tarefa 1 foi especificada a seguinte sequência de ações:

- 1) Selecionar o grafo de coautoria correspondente aos dados de 2011.

SUCESSO: A partir da barra de seleção a escolha do conjunto de dados a ser carregado é facilmente identificável.

- 2) Clicar no botão **T** da barra de títulos para exibir o painel de ferramentas.

SUCESSO: Além da letra **T** o único elemento usado para identificar que um painel de ferramentas pode ser ativado é o *tooltip* vinculado ao botão. É recomendável que o usuário tenha conhecimento da existência do painel de ferramentas.

- 3) Mover o painel de ferramentas para fora da área de visualização para possibilitar que todos os itens exibidos possam ser observados.

SUCESSO: O usuário pode inferir que o painel pode ser reposicionado livremente na área de trabalho observando o tipo de cursor do mouse quando ele é posicionado sobre o painel. Entretanto é exigido um conhecimento prévio do usuário.

- 4) Digitar o nome Marcelo Soares Pimenta na caixa de texto do painel de ferramentas.

SUCESSO: O local para digitação do texto é facilmente identificável. A caixa de texto utiliza um mecanismo de autocompletar que exhibe a relação de autores disponíveis à medida que o nome (ou parte dele) é digitado. O nodo correspondente ao autor procurado é marcado no diagrama de nodos e arestas.

- 5) Ativar o menu de contexto a partir do nodo marcado como Marcelo Soares Pimenta para instanciar uma *ClusterVis*.

SUCESSO: A indicação que essa ação pode ser executada está na barra de seleção da ferramenta. No entanto, deve haver um conhecimento prévio do usuário de que a *ClusterVis* exhibe atributos de nodos de um agrupamento.

- 6) Reposicionar a *ClusterVis* para fora da área de visualização do diagrama de nodos e arestas.

SUCESSO: A ação de reposicionamento é facilmente identificável.

- 7) Clicar no botão **T** da barra de títulos da *ClusterVis* para exhibir o painel de ferramentas.

SUCESSO: Vide comentário da ação 2.

- 8) Mover o painel de ferramentas para fora da área de visualização para possibilitar que todos os itens exibidos possam ser observados.

SUCESSO: Vide comentário da ação 3

- 9) Selecionar o item *Qt. Publications* no *comboBox Order by* para classificar as barras por quantidade de *Publications*.

SUCESSO: A partir da exibição do painel de ferramentas a escolha do critério de classificação das barras é facilmente identificável.

- 10) Identificar no painel de ferramentas em qualanel está representado o atributo quantidade de *Publications*.

FALHA: Esse é problema já foi identificado na tarefa 1, ação 4 na avaliação da *ClusterVis*.

- 11) Mover o cursor do mouse sobre a maior barra doanel identificado na ação anterior para ativar o *tooltip*. Caso a maior barra pertença ao autor Marcelo Soares Pimenta mover o cursor para a segunda maior barra.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que a técnica se utiliza de *tooltips* para obter informações detalhadas, esta ação é executada

corretamente. No entanto, o uso de *tooltips* para obter o nome dos autores faz que sejam necessárias múltiplas ativações até encontrar a informação desejada.

Tarefa 2 – Qual a quantidade de publicações geradas pela coautoria entre Marcelo Soares Pimenta e o autor obtido na tarefa anterior?

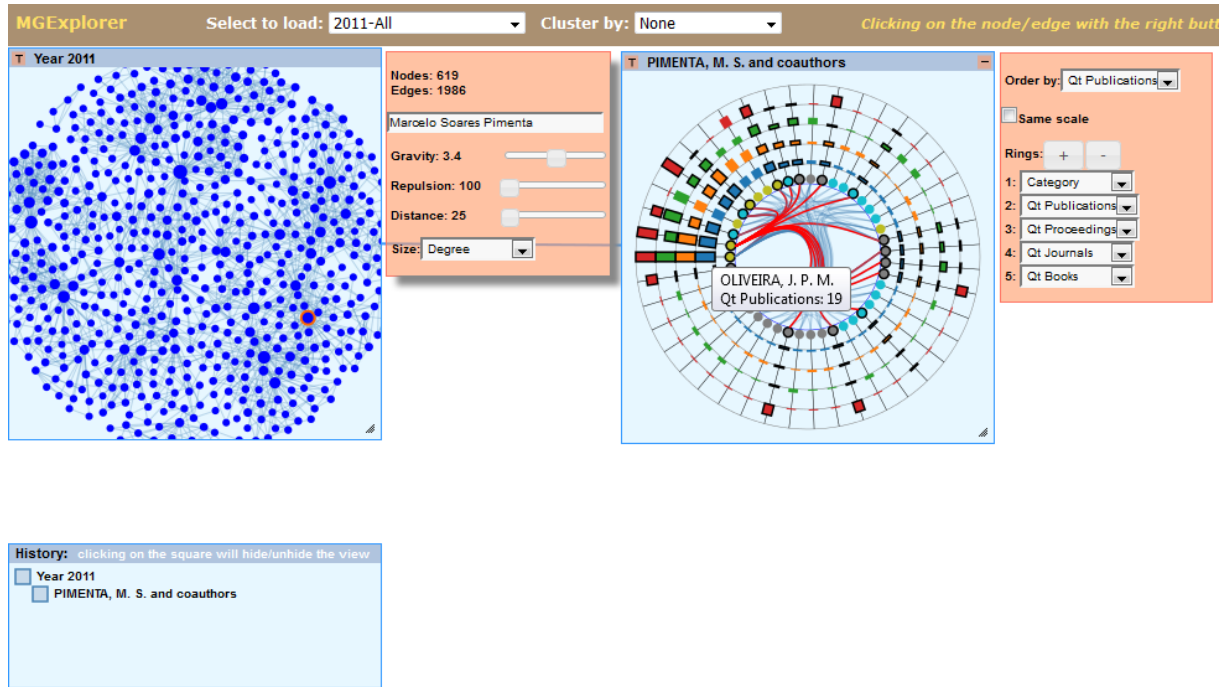


Figura 6.20 –Estado inicial da *MGExplorer* para a tarefa 2

Para a execução da tarefa 2 foi especificada a seguinte sequência de ações:

- 1) Localizar o autor identificado na tarefa anterior.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que a técnica se utiliza de *tooltips* para obter informações detalhadas, esta ação é executada corretamente.

- 2) Ativar o menu de contexto a partir dos barras que representam o autor identificado na ação anterior para instanciar uma *Iris*.

SUCESSO: A indicação que essa ação pode ser executada está na barra de seleção da ferramenta. No entanto, deve haver um conhecimento prévio do usuário de que a *Iris* exibe atributos de arestas entre um nodo de interesse e seus adjacentes.

- 3) Remover o painel de ferramentas da *ClusterVis* para disponibilizar espaço na área de trabalho.

SUCESSO: No entanto é necessário um conhecimento prévio do usuário que o clique no botão “T” exibe/esconde o painel de ferramentas.

- 4) Reposicionar a *Iris* para fora da área de visualização do diagrama de nodos e arestas.

SUCESSO: A ação de reposicionamento é facilmente identificável.

- 5) Redimensionar a *Iris* para facilitar a leitura dos nomes dos autores.

SUCESSO: A ação de redimensionamento é facilmente identificável.

- 6) Clicar em algum autor cujo nome comece com a letra “P”.

SUCESSO: O nome deve ser posicionado no foco para facilitar a leitura do texto e a identificação do autor Marcelo Pimenta. Embora a alteração do cursor do mouse indique que uma barra pode ser clicada deve existir um conhecimento prévio do usuário de que o clique posiciona o respectivo nodo no foco

- 7) Identificar o autor Marcelo Pimenta e ativar a *tooltip* para obter as informações mais detalhadas.

SUCESSO: Considerando que o usuário possui um conhecimento prévio de que a técnica utiliza *tooltips* para obter informações detalhadas, esta ação é executada corretamente.

Tarefa 3 – Considerando as relações de coautoria entre os elementos do *cluster* formado pelo autor Claudio Jung e seus coautores, qual coautoria (2 autores) gerou maior quantidade de publicações?

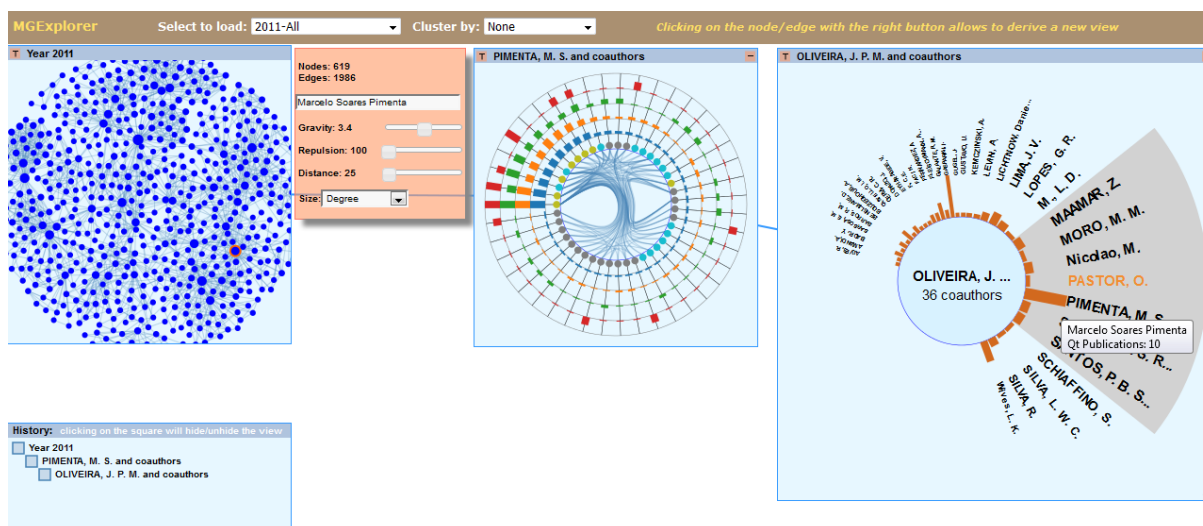


Figura 6.21 –Estado inicial da *MGExplorer* para a tarefa 2

Para a execução da tarefa 3 foi especificada a seguinte sequência de ações:

- 1) Digitar o nome Cláudio Jung na caixa de texto do painel de ferramentas.

SUCESSO: O local para digitação do texto é facilmente identificável. A caixa de texto utiliza um mecanismo de autocompletar que exhibe a relação de autores disponíveis à medida que o nome (ou parte dele) é digitado. O nodo correspondente ao autor procurado é marcado no diagrama de nodos e arestas.

- 1) Ativar o menu de contexto a partir do nodo marcado como Cláudio Jung para instanciar uma *GlyphMatrix*.

SUCESSO: A indicação que essa ação pode ser executada está na barra de seleção da ferramenta. No entanto, deve haver um conhecimento prévio do usuário de que a *GlyphMatrix* exhibe atributos de arestas de um grafo.

- 2) Ocultar a janela que contém a *ClusterVis* para disponibilizar mais espaço na área de trabalho.

SUCESSO: Essa ação é facilmente identificável.

- 3) Reposicionar e redimensionar a *GlyphMatrix* para fora da área de visualização do diagrama de nodos e arestas.

SUCESSO: A ações de reposicionamento e redimensionamento são facilmente identificáveis.

- 4) Clicar no botão **T** da barra de títulos da *Iris* para exibir o painel de ferramentas.

SUCESSO: Vide comentário apresentado na ação 2 da tarefa 1.

- 5) Mover o painel de ferramentas para fora da área de visualização para possibilitar que todos os itens exibidos possam ser observados.

SUCESSO: Vide comentário apresentado na ação 3 da tarefa 1

- 6) Selecionar o item *Qt. publications* no *comboBox Order by* para classificar as barras por ordem de quantidade de publicações.

SUCESSO: A partir da exibição do painel de ferramentas a escolha do critério de classificação das barras é facilmente identificável.

- 7) Mover o cursor do mouse para a célula que contém glifo com maior valor de quantidade de publicações para ativar o *tooltip* que exhibe as informações detalhadas.

SUCESSO: Quando existe uma diferença significativa entre o maior e os demais a célula é facilmente identificável. Entretanto para valores próximos pode ser necessária uma exploração individualizada (por meio dos *tooltips*) em cada célula.

6.4.2 Performance

A seguir são descritas as ações necessárias para a execução de cada tarefa com a respectiva média de tempo de execução para avaliação na dimensão performance.

Tarefa 1: Qual o coautor do autor Marcelo Soares Pimenta possui mais publicações em 2011?

Tabela 6.14 – Detalhamento das ações para execução da tarefa 1 na *MGExplorer*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o <i>comboBox Select to load</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>2011-All</i>	P [<i>2011-All</i>]	1,10
Clicar no item <i>2011-All</i>	K [<i>2011-All</i>]	0,23
Mover o cursor do mouse sobre o botão <i>tools</i>	P [botão]	1,10
Clicar com o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse sobre o painel de ferramentas	P [painel]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse reposicionando o painel de ferramentas para fora da área de visualização da técnica	P [painel]	1,10
Mover o cursor do mouse para o campo <i>Search</i>	P [<i>Search</i>]	1,10
Clicar sobre o campo <i>Search</i>	K [mouse]	0,23
Colocar a mão sobre o teclado	H [teclado]	0,36
Digitar o nome Marc	M4K[teclado]	2,15
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o item Marcelo Soares Pimenta	P [item]	1,10
Clicar no item Marcelo Soares Pimenta	K [mouse]	0,23
Localizar o círculo marcado em vermelho no diagrama de nodos e arestas	M [círculo]	1,35
Mover o cursor do mouse para o círculo identificado no passo anterior	P [item]	1,10
Clicar com o botão direito no círculo para ativar o menu de contexto	K [mouse]	0,23
Mover o cursor do mouse sobre o item <i>ClusterVis</i>	P [item]	1,10
Clicar no item <i>ClusterVis</i>	K [mouse]	0,23
Mover o cursor do mouse sobre a barra de título da janela que foi aberta	P [título]	1,10
Pressionar o botão esquerdo mouse	K [mouse]	0,23
Reposicionar a janela com a técnica <i>ClusterVis</i> na área de trabalho	P [janela]	1,10
Mover o cursor do mouse sobre o botão <i>tools</i> da <i>ClusterVis</i>	P [botão]	1,10
Clicar com o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse sobre o painel de ferramentas	P [painel]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse reposicionando o painel de ferramentas para fora da área de visualização da técnica	P [painel]	1,10
Mover o cursor do mouse para o <i>comboBox Order by</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>QtPublications</i>	P [<i>Attribute</i>]	1,10
Clicar no item <i>QtPublications</i>	K [<i>Attribute</i>]	0,23
Identificar o anel que representa a quantidade de publicações	M [anel]	1,35
Identificar a maior barra do anel	M [barra]	1,35
Mover o cursor do mouse sobre a barra identificada no passo anterior ativando o <i>tooltip</i>	P [barra]	1,10
Identificar o nome do autor que possui a maior quantidade de publicações	M [<i>tooltip</i>]	1,35
Se o autor identificado no passo anterior for Marcelo Soares Pimenta identificar a segunda maior barra	M [barra]	1,35
Mover o cursor do mouse sobre a barra identificada no passo anterior ativando o <i>tooltip</i>	P [barra]	1,10
Identificar o nome do autor que possui a maior quantidade de publicações	M [<i>tooltip</i>]	1,35
Total		31,97

Tarefa 2: Qual a quantidade de publicações geradas pela coautoria entre Marcelo Soares Pimenta e o autor obtido na tarefa anterior?

Tabela 6.15 – Detalhamento das ações para execução da tarefa 2 na *MGExplorer*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o mouse sobre a barra que representa o autor identificado na tarefa anterior	P [barra]	1,10
Clicar com o botão direito na barra para ativar o menu de contexto	K [mouse]	0,23
Mover o cursor do mouse sobre o item <i>Iris</i>	P [item]	1,10
Clicar no item <i>Iris</i>	K [mouse]	0,23
Mover o cursor do mouse sobre o botão <i>tools</i> da <i>ClusterVis</i>	P [botão]	1,10
Clicar com o botão esquerdo do mouse para ocultar o painel	K [mouse]	0,23
Mover o cursor do mouse sobre a barra de título da janela que foi aberta	P [título]	1,10
Pressionar o botão esquerdo mouse	K [mouse]	0,23
Reposicionar a janela com a técnica <i>Iris</i> na área de trabalho	P [janela]	1,10
Mover o cursor do mouse sobre a borda da janela	P [borda]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse com o botão pressionado para redimensionar a janela	P [borda]	1,10
Identificar um nome de autor que começa com a letra P	M [letra P]	1,10
Clicar no nome do autor para movê-lo para o foco	K [autor]	0,23
Localizar o autor alvo Pimenta, M. S.	M [autor]	1,35
Mover o cursor do mouse sobre o autor identificado no passo anterior ativando o <i>tooltip</i>	P [autor]	1,10
Identificar a quantidade de publicações exibidas pelo <i>tooltip</i>	M [<i>tooltip</i>]	1,35
Total		14,34

Tarefa 3: Considerando as relações de coautoria entre os elementos do *cluster* formado pelo autor Cláudio Jung e seus autores, qual coautoria (2 autores) gerou maior quantidade de publicações?

Tabela 6.16 – Detalhamento das ações para execução da tarefa 3 na *MGExplorer*

Descrição	Operação	Tempo (seg.)
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o campo <i>Search</i> do painel de ferramentas do diagrama de nodos e arestas	P [<i>Search</i>]	1,10
Clicar sobre o campo <i>Search</i>	K [mouse]	0,23
Colocar a mão sobre o teclado	H [teclado]	0,36
Digitar o nome Clau	M4K[teclado]	2.15
Mover a mão sobre o mouse	H [mouse]	0,36
Mover o cursor do mouse para o item Cláudio Jung	P [item]	1,10
Clicar no item Cláudio Jung	K [mouse]	0,23
Localizar o círculo marcado com a borda vermelha no diagrama de nodos e arestas	M [círculo]	1,35
Mover o cursor do mouse para o círculo identificado no passo anterior	P [círculo]	1,10
Clicar com o botão direito no círculo para ativar o menu de contexto	K [mouse]	0,23
Mover o cursor do mouse sobre o item <i>GlyphMatrix</i>	P [item]	1,10
Clicar no item <i>GlyphMatrix</i>	K [mouse]	0,23
Mover o cursor do mouse para o botão <i>Hide</i> da <i>ClusterVis</i>	P [<i>Hide</i>]	1,10
Clicar com o botão esquerdo do mouse para ocultar a <i>ClusterVis</i>	K [mouse]	0,23
Mover o cursor do mouse sobre a barra de título da janela que foi aberta	P [título]	1,10
Pressionar o botão esquerdo mouse	K [mouse]	0,23
Reposicionar a janela com a técnica <i>GlyphMatrix</i> na área de trabalho	P [janela]	1,10
Mover o cursor do mouse sobre a borda da janela	P [borda]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse com o botão pressionado para redimensionar a janela	P [borda]	1,10
Mover o cursor do mouse sobre o botão <i>tools</i> da <i>GlyphMatrix</i>	P [botão]	1,10
Clicar com o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse sobre o painel de ferramentas	P [painel]	1,10
Pressionar o botão esquerdo do mouse	K [mouse]	0,23
Mover o cursor do mouse reposicionando o painel de ferramentas para fora da área de visualização da técnica	P [painel]	1,10
Mover o cursor do mouse para a <i>comboBox Order by</i>	P [<i>comboBox</i>]	1,10
Clicar na seta do <i>comboBox</i>	K [<i>comboBox</i>]	0,23
Mover o cursor para o item <i>Qt.Publications</i>	P [<i>Qt.Publications</i>]	1,10
Clicar no item <i>Qt.Publications</i>	K [<i>Qt.Publications</i>]	0,23
Localizar a célula cujo glifo identifica a maior quantidade de publicação	M [Célula]	1,35
Mover o cursor do mouse sobre a célula para ativar o <i>tooltip</i>	P [<i>tooltip</i>]	1,10
Identificar as informações do <i>tooltip</i>	M [Célula]	1,35
Total		15,04

6.5 Comentários finais

Esta seção apresenta alguns comentários sobre a avaliação de cada uma das três técnicas propostas neste trabalho e da ferramenta *MGExplorer*.

O resultado da avaliação da *ClusterVis* foi considerado satisfatório nas três dimensões avaliadas. As tarefas propostas foram consideradas fáceis de serem executadas pela maioria dos usuários. Entretanto alguns problemas foram observados. Foi detectada uma falha na identificação dos anéis que exibem os atributos dos nodos. Na área de visualização os atributos se diferenciam pelas cores e posição dos anéis, mas não há qualquer indicação que os relacione com nomes que são exibidos no painel de ferramentas. Um usuário que

participou da avaliação sugere que cores idênticas às usadas nos anéis sejam incluídas no painel de ferramentas para facilitar a identificação dos atributos.

Outro ponto a ser considerado está relacionado à forma como as informações mais detalhadas são exibidas. A estratégia usada (*tooltips*) mostra uma única informação por vez. No entanto, em algumas situações, é necessário que as informações de vários nodos possam ser exibidas simultaneamente. Um usuário comenta sobre a dificuldade de se obter informações de nodos relacionados (destacados pelas arestas que os ligam), pois ao mover o cursor do mouse de um nodo para o outro (para ativar o *tooltip*) a relação é perdida devido ao ocultamento da aresta em destaque. Uma alternativa de solução para esse problema é ativar a exibição das informações para todos os nodos em destaque.

A técnica *GlyphMatrix* foi avaliada por meio de duas aplicações distintas. Para as dimensões utilidade e performance foi usada uma implementação similar à empregada na *MGEplorer* e a dimensão satisfação foi avaliada com a *SoccerMatches* (CAVA e FREITAS, 2013), uma aplicação projetada para visualizar dados estatísticos de partidas de futebol. De forma similar à *ClusterVis* o resultado da avaliação foi considerado satisfatório. A *GlyphMatrix* propicia uma visão geral dos atributos de arestas de um grafo, que pode ser prejudicada quando a quantidade de nodos aumenta, em consequência da redução das dimensões de suas células. Um mecanismo que diminui a quantidade de células visíveis é usado para minimizar esse problema, mas possui o inconveniente de reduzir a percepção da visão geral da estrutura. Observa-se também que nem todo tipo de glifo tem sua percepção facilitada. Enquanto os glifos (dois círculos concêntricos) que mostram as diferenças de gols foram facilmente interpretados pelos usuários, os gráficos de barras que mostram as estatísticas das partidas foram considerados de difícil percepção.

De forma similar a *GlyphMatrix* a técnica *Iris* também foi avaliada por meio de duas aplicações distintas. Para as dimensões utilidade e performance foi usada uma implementação similar à empregada na *MGEplorer* e a dimensão satisfação foi avaliada com uma aplicação que proporciona a visualização dos resultados de buscas executadas na *Web* (CAVA et al. 2014). Em linhas gerais a *Iris* atende o objetivo para a qual foi proposta: a visualização de atributos de arestas que conectam um grupo de nodos a um determinado nodo de interesse. No entanto, a estratégia adotada (foco+contexto), que possibilita que uma maior quantidade de nodos seja exibida, produz uma redução no tamanho do texto e das barras dificultando a leitura. Como consequência dessa dificuldade, a exploração da área de contexto pode acarretar no uso de vários cliques para mover o nodo procurado para a área de foco.

A avaliação da ferramenta *MGExplorer* foi considerada satisfatória para a execução das tarefas propostas nas duas dimensões em que foi avaliada: utilidade e performance. No entanto observa-se que a quantidade de ações para executar as tarefas pode ser reduzida. Cada vez que o painel de ferramentas é ativado é necessário movê-lo para fora da área de visualização. O mesmo ocorre quando uma técnica é instanciada. Ela é apresentada sobre a janela a partir da qual foi ativada ocultando as informações exibidas. Além disso, em algumas situações é necessário realizar o redimensionamento para permitir uma melhor interpretação do que está sendo visualizado. Uma estratégia de posicionamento e redimensionamento automático poderia minimizar esse problema.

Outro ponto a ser considerado é que embora as linhas que conectam as janelas representem a origem dos dados, não há mecanismos de coordenação que os destaquem na janela pai. Em algumas situações pode ser interessante que algum tipo de coordenação possa ser ativado, conforme a necessidade, para localizar a origem dos dados.

7 CONCLUSÃO

A visualização de grafos multivariados, os quais representam itens e relacionamentos caracterizados por apresentarem múltiplos atributos, é de interesse em muitos domínios de aplicação. No capítulo 2 desse trabalho, foram apresentadas diversas técnicas desenvolvidas na última década para esse propósito. No entanto, observa-se que com a aplicação de uma única técnica de visualização não é possível apoiar as diversas tarefas que podem surgir da necessidade de exploração da topologia do grafo e das relações de seus múltiplos atributos. Embora o uso de múltiplas visões do conjunto de dados de interesse não seja uma estratégia realmente nova, a abordagem proposta nesse trabalho possui a característica particular de depender de várias visões que podem ser construídas a partir de consultas iniciadas com a seleção de nodos e/ou arestas em diferentes visualizações. Além disso, as visões são encadeadas para representar a história do processo exploratório e as dependências entre elas. É importante notar que as visões não são estáticas, mas sim interativas, de modo que os usuários podem continuar a explorar múltiplos caminhos nas instâncias, ao longo do processo exploratório.

A abordagem proposta (Capítulo 3) para apoiar as tarefas exploratórias sobre os grafos multivariados foi baseada nas seguintes premissas:

- 1) Durante o processo de exploração de um grafo multivariado pode ser necessário que os usuários façam uso de diferentes instâncias de visualização do conjunto de dados;
- 2) Os itens de informação armazenados em grafos multivariados não podem ser facilmente exibidos todos de uma única vez. Portanto, um processo exploratório deve permitir que os usuários inspecionem e comparem subconjuntos de dados por meio de visões interligadas;
- 3) A exploração de grafos multivariados não exige apenas a exploração de nodos, mas também pode exigir a inspeção de arestas, subgrafos e relacionamentos criados em função do valor dos atributos de nodos e arestas.

Como prova de conceito para demonstrar a abordagem proposta, foi desenvolvida uma ferramenta (Capítulo 5) que, além de atender aos 11 requisitos definidos na seção 3.2, integra quatro técnicas de visualização: diagrama de nodos e arestas tradicional, *ClusterVis*, *GlyphMatrix* e *Iris*. As três últimas (Capítulo 4) foram desenvolvidas para a visualização de agrupamentos em grafos multivariados e possuem os seguintes objetivos:

- Visualizar as relações entre atributos de nodos (*ClusterVis*);

- Visualizar os atributos pertencentes as arestas (*GlyphMatrix*);
- Visualizar os atributos pertencentes a arestas que conectam um grupo de nodos a um determinado nodo de interesse (*Iris*).

A abordagem apresentada aqui, no entanto, não deve ser entendida como limitada a essas técnicas de visualização. De fato, foi assumido que a ferramenta é extensível e poderá dispor de outras técnicas de visualização no futuro. Enquanto cada uma das técnicas de visualização por si só é limitada pelas suas escolhas de *design*, a ferramenta permite desencadear qualquer outra visualização aplicada a partir de qualquer elemento, seja nodo ou aresta, que tem sido definido como selecionável na visualização que o usuário está usando no momento. Tais elementos selecionáveis são usados para consultar o grafo multivariado e coletar os dados necessários para a nova visualização. Outro ponto a destacar é o fato de que embora tenha sido usado como exemplo a visualização de grafos de coautoria, a abordagem é suficientemente genérica para permitir a exploração de outros domínios.

7.1 Publicações

Durante o desenvolvimento desse trabalho, foram produzidas as seguintes publicações:

- STEIN, A. CAVA R. A., DORNELES, C. F., FIRMENICH, S., FREITAS, C.M.S., PALANQUE, P., WINCKLER, M. Proposta de um Framework para Visualização de Dados agregados por Similaridade para Auxiliar Consultas durante a Navegação na *Web*. In: IHC 2013 – Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais, 2013, Manaus, AM, Brasil. IHC 2013 – Simpósio Brasileiro de Fatores Humanos em Sistemas Computacionais. Porto Alegre, RS: SBC, 2013, p. 148-157.
- CAVA, R. A., FREITAS, C. M. S., Glyphs in Matrix Representation of Graphs for Displaying Soccer Games Results. In: IEEE VIS 2013 Workshop: 1st Workshop on Sports Data Visualization, 2013, Atlanta: Georgia Institute of Technology 2013. P. 1-6.
- CAVA, R. A., FREITAS, C. M. S., BARBONI, E., PALANQUE, P., WINCKLER, M. In Search: an alternative for performing ancillary search tasks on the Web. In: 9th Latin American Web Congress 2014, Ouro Preto. 9th Latin American Web Congress. Los Alamitos: IEEE CPS, 2014. P. 1-8.
- WINCKLER, M., CAVA, R. A., BARBONI, E., PALANQUE, P., FREITAS, C.M.S. Usability aspects of the inside-in approach for ancillary search tasks on

the web. IFIP TC13. Conference on Human-Computer Interaction 2015. LNC v.9297, p.207-226. Springer, 2015.

- CAVA, R. A., FREITAS, C.M.S., WINCKLER, M.– ClusterVis: Visualizing Nodes Attributes in Multivariate Graphs. In: 32nd ACM Symposium on Applied Computing, ACM SAC' 17, 2017 (aceito).
- CAVA, R. A., FREITAS, C.M.S, WINCKLER, M. – Synergistic approach for interactive explorarion of multidimensional graphs. In: Information Visualization (a ser submetido).

7.2 Trabalhos futuros

Embora os objetivos deste trabalho tenham sido atingidos com a abordagem e as técnicas propostas, considera-se que outros pontos podem ser ainda investigados. Na concepção das técnicas *Iris* e *ClusterVis*, não foi levada em consideração a direção das arestas, fazendo com que apenas grafos não dirigidos possam ser representados. Com isso, pode-se considerar como um novo aspecto a ser explorado a investigação de alternativas para a representação de aresta dirigidas. Outro ponto a ser considerado é a inclusão de alguma forma de identificação dos vários tipos de relacionamentos que podem existir em um grafo multivariado. Por exemplo, em uma relação de coautoria pode ser interessante identificar a relação existente entre os coautores (professor-aluno, orientador-orientando, professor-professor, entre outras).

Outras técnicas podem ser especialmente concebidas para visualizar algumas estruturas particulares, como, por exemplo, a representação de atributos associados a caminhos, ou a representação de redes heterogêneas, as quais são constituídas por vários tipos de nodos (por exemplo, autor-artigo-evento). Outro ponto a ser considerado e a exploração de atributos de geolocalização que contém informações espaciais. Além disso, a representação do atributo tempo, que acrescenta uma maior complexidade ao grafo, é outro ponto a ser investigado. Pode-se pensar na passagem do tempo como que exercendo transformações na topologia (inclusão, alteração e remoção de novos nodos ou arestas) ou alterando o valor de atributos associados a nodos e/ou arestas. Encontrar visualizações que permitam explorar o aspecto dinâmico, introduzido pelo atributo tempo em um grafo multivariado, é um grande desafio.

REFERÊNCIAS

- AHN, J.; PLAISANT, C.; SHNEIDERMAN, B.: A task taxonomy of network evolution analysis. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 20, n. 3, p. 365-376, 2014.
- AMAR, R.A.; EAGAN, J.; STASKO, J.T.: Low level components of analytic activity in information visualization. In: IEEE SYMPOSIUM ON INFORMATION VISUALIZATION (INFOVIS), 2005, **Proceedings...**, Washington, DC, USA: IEEE Computer Society, 2005. p. 111-117.
- ANDERY, G. **Integrando projeções multidimensionais a análise visual de redes sociais**. Dissertação (Mestrado em Ciência da Computação). ICMC-USP, São Carlos, 2010.
- BALZER, M.; DEUSSEN, O. Voronoi treemaps. In: IEEE SYMPOSIUM ON INFORMATION VISUALIZATION (INFOVIS), 2005, **Proceedings...**, Washington, DC, USA: IEEE Computer Society, 2005. p. 7-14.
- BATTISTA, G.; EADES, P.; TAMASSIA, R.; TOLLIS, I. **Graph Drawing: Algorithms for the Visualization of Graphs**. New Jersey, Prentice Hall, 1999.
- BEDERSON, B. Fisheye Menus. In: ACM SYMPOSIUM ON USER INTERFACE SOFTWARE AND TECHNOLOGY (UIST), 2000, **Proceedings...**, New York, NY, USA: ACM, 2000. p. 217-225.
- BECK, F.; BURCH, M.; DIEHL, S.; WEISKOPF, D. The state of the art in visualizing dynamic graphs. In: **EuroVis STAR**, [S.l], Eurographics Association, p.83-103,2014.
- BEZERIANOS, A.; CHEVALIER, F.; DRAGICEVIC, P.; ELMQVIST, N.; FEKETE, J. Graphdice. A System for Exploring Multivariate Social Networks. In: EUROGRAPHICS / IEEE VGTC CONFERENCE ON VISUALIZATION (EUROVIS), 2010, **Proceedings...**, Chichester, UK, The Eurographs Association & John Wiley & Sons, Ltd, v.29, n. 3, p. 863-871, 2010.
- BOSTOCK M.; OGIEVETSKY, V.; HEER, J. D3: datadriven documents. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 17, n. 12, p. 2301-2309, 2011.
- BREHMER, M.; MUNZNER, T. A multi-level typology of abstract visualization tasks. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 19, n. 12, p. 2376-2385, 2013.
- BROOKE, J. **SUS: A 'quick and dirty' usability scale**. In: JORDAN, P.; THOMAS, B.; McCLELLAND, I.; WEERDMEESTER, B. Usability Evaluation in Industry. London, Taylor and Francis, 1996.
- CAO, N.; LIN, Y-R.; LI, L.; TONG, H. G-Miner: Interactive visual group mining on multivariate graphs, In: ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI), 2015, **Proceedings....** New York, NY, USA: ACM, 2015. p 279-288.

CARD, S.; MORAN, T.; NEWELL, A. **The psychology of human-computer interaction**. Hillsdale, NJ: Lawrence Erlbaum Associates. 448 p, 1983.

CARD, S. K.; MACKINLAY, J.; SHNEIDERMAN, B. Information Visualization. In: CARD, S. K.; MACKINLAY, J. D.; SHNEIDERMAN, B. **Readings in Information Visualization: Using Visualization to Think**. San Francisco, CA, USA: Morgan Kaufmann Publishers, 1999a. p. 1-34.

CARD, S.; MACKINLAY, J.; SHNEIDERMAN, B. Network. In: CARD, S.; MACKINLAY, J.; SHNEIDERMAN, B. **Readings in Information Visualization: Using Visualization to Think**. CA, USA: Morgan Kaufmann Publishers, 1999b. p. 187-189.

CARD, S.; MACKINLAY, J.; SHNEIDERMAN, B. Focus+Context. In: CARD, S.; MACKINLAY, J.; SHNEIDERMAN, B. **Readings in Information Visualization: Using Visualization to Think**. CA, USA: Morgan Kaufmann Publishers, 1999c. p. 307-309.

CAVA, R. A.; FREITAS, C. M. S. Glyphs in Matrix Representation of Graphs for Displaying Soccer Games Results. In: **IEEE VIS 2013 Workshop: 1st Workshop on Sports Data Visualization**, 2013, Atlanta. 1st Workshop on Sports Data Visualization. Atlanta: Georgia Institute of Technology, 2013. p.1-6

CAVA, R. A.; FREITAS, C. M. S.; BARBONI, E.; PALANQUE, P.; WINCKLER, M. Inside-In Search: an alternative for performing ancillary search tasks on the Web. In: **9th Latin American Web Congress**, 2014, Ouro Preto, MG. 9th Latin American Web Congress. Los Alamitos, USA: IEEE CPS, 2014. p. 1-8.

COLLINS, C.; PENN, G.; CARPENDALE, S. Bubble sets: Revealing set relations with isocontours over existing visualizations. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 15, n. 6, p. 1009-1016, 2009.

DIEHL, S.; TELEA, A. Multivariate Network in Software Engineering. In: KERREN, A.; PURCHASE, H.; WARD, M. (Eds). **Multivariate Network Visualization**. [S.l]: Springer, 2014, (Lecture Notes in Computer Science v. 8380). p. 13-36. ISBN 978-3-319-06792-6.

DUNNE, C.; RICHE, N.; LEE, B.; METOYER, R.; ROBERTSON, G. GraphTrail: Analyzing large multivariate, heterogeneous networks while supporting exploration history. In: ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI), 2012, **Proceedings...**, New York, NY, USA: ACM, 2012. p. 1663-1672.

EADES, P. A heuristic for graph drawing. In: **Congressus Numerantium**, [S.l:s.n]; v. 42, p.149-160, 1984.

ELMQVIST, N.; DRAGICEVIC, P.; FEKETE, J. Rolling the dice: Multidimensional visual exploration using scatterplot matrix navigation., **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 14, n. 6, p. 1141-1148, 2008.

ELZEN, S.; WIJK, J. Multivariate Network Exploration and Presentation: From Detail to Overview via Selections and Aggregations. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 20, n. 12, p. 2310-2319, 2014.

FEKETE, J.; PLAISANT, C. Excentric labeling: dynamic neighborhood labeling for data visualization. In: ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI), 1999, **Proceedings**, New York, NY, USA: ACM, 1999. p. 512-519.

FISHER, D. Using egocentric networks to understand communication. **IEEE Internet Computing**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 9, n. 5, p. 20-28, 2005.

GERSHON, N.; EICK, S. Information Visualization. **IEEE Computer Graphics and Applications**, IEEE Computer Society Press, Los Alamitos, CA, USA, v. 17, n.4, p. 29-31, 1997.

GHANI, S.; KWON, B.; LEE, S.; YI, J.; ELMQVIST, N. Visual analytics for multimodal social network analysis: A design study with social scientists. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 19, n. 12, p. 2032-2041, 2013.

GÖRG, C.; POHL, M.; QELI, E.; XU, K. Visual Representations. In: KERREN, A.; EBERT, A., MEYER, J. **Human-Centered Visualization Environments**. Springer, Heidelberg (Lecture Notes in Computer Science v. 4417), p. 163-230, 2007.

GOU, L.; ZHANG. TreeNetViz: revealing patterns of networks over tree structures. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 17, n. 12, p. 2449-2458, 2011.

GOU, L.; ZHANG, X.; LUO, A.; ANDERSON, P. SocialNetSense: supporting sensemaking of social and structural features in networks with interactive visualization. In: IEEE CONFERENCE ON VISUAL ANALYTICS SCIENCE AND TECHNOLOGY (VAST), 2012, **Proceedings...**, Seattle, WA, USA [s.n.]. p. 133-142, 2012.

HEER, J.; ROBERTSON, G. Animated Transitions in Statistical Data Graphics. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 13, n. 6, p. 1240-1247, 2007.

HERMAN, I.; MELANÇON, G.; MARSHALL, M. Graph Visualization and Navigation in Information Visualization: a survey. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 6, n. 1, p. 24-43, 2000.

HOLTEN, D. Hierarchical edges bundles: Visualization of adjacency relations in hierarchical data. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 12, n. 5, p. 741-748, 2006

JUSUFI, I.; DINGJIE, Y.; KERREN, A. The Network Lens: Interactive exploration of multivariate networks using visual filtering. In: INTERNATIONAL CONFERENCE ON

INFORMATION VISUALIZATION, **Proceedings...**, Washington, DC, USA: IEEE Computer Society, 2010. p. 35-42.

JUSUFI, I.; KERREN, A.; ZIMMER, B. Multivariate network exploration with JauntyNets. In: INTERNATIONAL CONFERENCE ON INFORMATION VISUALIZATION, **Proceedings...**, Washington, DC, USA: IEEE Computer Society, 2013. p. 19-27.

JUSUFI, I.; KERREN, A.; LIU, J.; ZIMMER, B. Visual Exploration of Relationships between Document Clusters. In: INTERNATIONAL CONFERENCE ON INFORMATION VISUALIZATION THEORY AND APPLICATION (IVAPP), 2014, **Proceedings...**, SciTePress, Lisbon, Portugal, p. 195-203, 2014.

KEIM, D. Designing Pixel-Oriented Visualization Techniques: Theory and Applications. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 6, n. 1, p. 59-78, 2000.

KITCHENHAM, B.; CHARTERS, S. **Guidelines for performing systematic literature reviews in software engineering**, EBSE Technical report, Keele University, [S.l.: s.n], 2007.

KERREN, A.; JUSUFI, I.; ALEKSAKHIN, V.; SCHREIBER, F. Cluma-GO: Bring Gene Ontologies and Hierarchical Clustering Together. In: IEEE SYMPOSIUM ON BIOLOGICAL DATA VISUALIZATION (BioVis), 2011: **Proceedings...** Extended Abstract Providence, RI, USA, 2011.

KERREN, A.; KÖSTINGER, H.; ZIMMER, B. VINCENT – Visualization of network centralities. In: INTERNATIONAL CONFERENCE ON INFORMATION VISUALIZATION THEORY AND APPLICATION (IVAPP), 2012, **Proceedings...**, Springer, [S.l.], p. 703-712, 2012.

KERREN, A.; PURCHASE, H.; WARD, M. Introduction to Multivariate Network Visualization. In: KERREN, A.; PURCHASE, H.; WARD, M. (Eds). **Multivariate Network Visualization**. [S.l]: Springer, 2014, (Lecture Notes in Computer Science v. 8380). p. 1-9. ISBN 978-3-319-06792-6.

KO, S.; AFZAL, S.; WALTON, S.; YANG, Y.; CHAE, J.; MALIK, A.; JANG, Y.; CHEN, M.; EDBERT, D. Analyzing high-dimensional multivariate network links with integrated anomaly detection, highlighting and exploration. In: IEEE CONFERENCE ON VISUAL ANALYTICS SCIENCE AND TECHNOLOGY (VAST), 2014, **Proceedings...**, Paris, France, IEEE, p.83-92, 2014.

KOHLBACHER, O.; SCHREIBER, F.; WARD, M. Multivariate Network in the Life Sciences. In: KERREN, A.; PURCHASE, H.; WARD, M. (Eds). **Multivariate Network Visualization**. [S.l]: Springer, 2014, (Lecture Notes in Computer Science v. 8380). p. 61-73. ISBN 978-3-319-06792-6.

LANDESBERGER, T.; KUIJPER, A.; SCHRECK, T.; KOHLHAMMER, J.; WIJK, J.; FEKETE, J.; FELLNER, D. Visual Analysis of Large Graphs: State-of-the-Art and Future Research Challenges. **Computer Graphics Forum**, [S.l.:s.n] v. 30, n. 6, p. 1719-1749, 2011.

LEE, B.; PLAISANT, C.; PARR, C. S.; FEKETE, J. D.; HENRY, N. Task taxonomy for graph visualization. In: AVI WORKSHOP ON BEYOND TIME AND ERRORS (BELIV), 2006, **Proceedings...** New York, NY, USA: ACM, 2006. p. 1-5.

LEWIS, C.; WHARTON, C. Cognitive Walkthroughs. In: HELANDER, M.; LANDAUER, T. K.; PRABHU, P. (Eds). **Handbook of Human-Computer Interaction**. New York: Elsevier Press. p. 717-732, 1997.

LYMAN, P.; VARIAN, H. How Much Storage is Enough? **ACM Queue** [S.l:s.n] v. 1, n. 4, 2003.

MUELDER, C.; GOU, L.; MA, K-L.; ZHOU, M. Multivariate Social Network Visual Analytics. In: KERREN, A.; PURCHASE, H.; WARD, M. (Eds). **Multivariate Network Visualization**. [S:l]: Springer, 2014, (Lecture Notes in Computer Science v. 8380). p. 37-59. ISBN 978-3-319-06792-6.

MUNZNER, T. A Nested Model to Visualization Design and Validation. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v.15, n.6, p.921-928, 2009.

MUNZNER, T. **Visualization Analysis and Design**, A. K. Peters/CRC Press, 2014.

OLIVEIRA, M.; LEVKOWITZ, H., From visual data exploration to visual data mining: a survey. **IEEE Transaction on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 9, n. 3, p. 378-397, 2003.

PARTL, C.; LEX, A.; STREIT, M.; KALKOFEN, D.; KASHOFER, K.; SCHMALSTIEG, D. enRoute: dynamic path extraction from biological pathway maps for indepth experimental data analysis. In: IEEE SYMPOSIUM ON BIOLOGICAL DATA VISUALIZATION (BIOVIS), 2012, **Proceedings...**, IEEE Computer Society, Washington, DC, USA, p. 107-114, 2012.

PINZGER, M.; GALL, H.; FISCHER, M.; LANZA, M. Visualizing Multiple Evolution Metrics. In: ACM SYMPOSIUM ON SOFTWARE VISUALIZATION (SOFTVIS), 2005, **Proceedings...**, ACM, New York, NY, USA, p. 354-362, 2005.

PRETORIUS, A. J.; WIJK, Van J. J. Visual analysis of multivariate state transition graphs. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 12, n. 5, p. 685-692, 2006. ISSN 1077-2626.

PRETORIUS, A. J.; WIJK, Van J. J. Visual inspection of multivariate graphs. In: EUROGRAPHICS / IEEE-VGTG SYMPOSIUM ON VISUALIZATION (EUROVIS), 2008, **Proceedings...**, Washington, DC, USA: IEEE Computer Society, 2008. p. 967-974.

PRETORIUS, J.; PURCHASE, H.; STASKO, J.: Tasks for Multivariate Network Analysis. In: KERREN, A.; PURCHASE, H.; WARD, M. (Eds). **Multivariate Network Visualization**. [S:l]: Springer, 2014, (Lecture Notes in Computer Science v. 8380). p. 77-95. ISBN 978-3-319-06792-6.

SHANNON, R.; HOLLAND, T.; QUIGLEY, A. **Multivariate graph drawing using parallel coordinate visualisations**. Dublin: [s.n.], Technical Report 2008-6, University College Dublin, 2008.

SHI, L.; LIAO, Q.; TONG, H.; HU, Y.; ZHAO, Y.; LIN, C. Hierarchical Focus+Context Heterogeneous Network Visualization. In: IEEE PACIFIC VISUALIZATION SYMPOSIUM (PacificVis), 2014, **Proceedings...**, IEEE, Hangzhou, p. 89-96, 2014.

SHNEIDERMAN, B. The Eyes Have It: A task by data type taxonomy for information visualizations. In: IEEE VISUAL LANGUAGES, College Park, Maryland 20742, U.S.A. **Proceedings...** [S.l.:s.n], 1996. n.UMCP-CSD CS-TR-3665, p.336–343.

SHNEIDERMAN, B.; ARIS, A. Network Visualization by Semantic Substrates. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v.12, n.5, p.733-740, 2006.

TAMASSIA, R. **Handbook of Graph Drawing and Visualization (Discrete Mathematics and Its Applications)**. Chapman and Hall/CRC Press, 2013.

TUKEY, J. **Exploratory data analysis**. Reading, Mass., Addison-Wesley, 1977.

ULRIK, Brandes; BOBO Nick. Asymmetric relations in longitudinal social networks. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 17, n. 12, p. 2283–2290, 2011.

VALIATI, E.R.A.; PIMENTA, M. S.; FREITAS, C. M. D. S. A taxonomy of tasks for guiding the evaluation of multidimensional visualizations. In: AVI WORKSHOP ON BEYOND TIME AND ERROR (BELIV), 2006, **Proceeding...**, ACM, New York, NY, USA, p. 1-6, 2006.

VIAU, C.; MCGUFFIN, M.; CHIRICOTA, Y.; JURISICA, I. The FlowVizMenu and Parallel Scatterplot Matrix: Hybrid Multidimensional Visualizations for Network Exploration. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v.16, n.1, p.1100-1108, 2010.

WANG, Y.; SHEN, Q.; ARCHAMBAULT, D.; ZHOU, Z.; ZHU, M.; YANG S.; QU H. AmbiguityVis: Visualization of Ambiguity in Graph Layouts. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department Piscataway, NJ, USA, v.22, n.1, p.359-368, 2016.

WARD, M. A Taxonomy of Glyph Placement Strategies for Multidimensional Data Visualization. **Information Visualization**, Palgrave Macmillan, v. 1, n. 3-4, p. 194-210, 2002.

WARD, M.; GRINSTEIN, G.; KEIM, D. **Interactive Data Visualization: Foundations, Techniques and Applications**, Natick, MA, USA: A. K. Peters, Ltd, 2010.

WATTENBERG, M. Arc diagrams: Visualizing structure in strings. In: IEEE SYMPOSIUM ON INFORMATION VISUALIZATION (INFOVIS), 2002, **Proceedings...**, DC, USA: IEEE Computer Society, 2002. p. 110-116.

WATTENBERG, M. Visual exploration of multivariate graphs. In: ACM CONFERENCE ON HUMAN FACTORS IN COMPUTING SYSTEMS (CHI), 2006, **Proceedings....** New York, NY, USA: ACM Press, 2006. p. 811-819.

WONG, P.; BERGERON, D. 30 years of multidimensional multivariate visualization. In: **Scientific Visualization, Overview, Methodologies, and Techniques**. IEEE Computer Society, Washington, DC, USA, p. 3-33, 1994.

WU, Y.; WU, W.; YANG, S.; YAN, Y.; QU, H. Interactive visual summary of major communities in a large network. In: IEEE PACIFIC VISUALIZATION SYMPOSIUM (PacificVis), 2015, **Proceedings...**, IEEE, Hangzhou, p. 47-54, 2015.

WU, Y.; PITIPORNVIVAT, N.; ZHAO, J.; YANG, S.; HYANG, G.; QU, H. EgoSlider: Visual Analysis of Egocentric Network Evolution. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 22, n. 1, p. 260-269, 2016

WYBROW, M.; ELMQVIST, N.; FEKETE, J.; LANDESBERGER, T.; WIJK, J.; ZIMMER, B. Introduction to Multivariate Network Visualization. In: KERREN, A.; PURCHASE, H.; WARD, M. (Eds). **Multivariate Network Visualization**. [S.l]: Springer, 2014, (Lecture Notes in Computer Science v. 8380). p. 97-125. ISBN 978-3-319-06792-6.

XU, K.; CUNNINGHAM, A.; HONG, S. H.; THOMAS, B. H. Graphscape: Integrated multivariate network visualization. In: INTERNATIONAL ASIA-PACIFIC SYMPOSIUM ON VISUALIZATION (APVIS), 2007, **Proceedings...**, IEEE, Sydney, p. 33-49, 2007.

YANG, J.; LIU, Y.; ZHANG, X.; YUAN, X.; ZHAO, Y.; BARLOWE, S.; LIU, S. PIWI: Visually explorint graphs based on their community structure. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 19, n. 6, p. 1034-1047, 2013.

YI, J.S.; KANG, Y.; STASKO, J.T.; JACKO, J.A. Toward a deeper understanding of the role of interaction in information visualization. **IEEE Transactions on Visualization and Computer Graphics**, IEEE Educational Activities Department, Piscataway, NJ, USA, v. 13, n. 6, p. 1224-1231, 2007.

APÊNDICE A – DOCUMENTOS USADOS NAS AVALIAÇÕES**Perfil do usuário**

Idade:

- 15-20 21-30 31-40 41-50 mais de 50

Sexo:

- Feminino masculino

Escolaridade: escolha o nível mais recente:

- Ensino médio completo
 Aluno de graduação
 Graduado
 Mestre
 Doutor

Área da atividade principal:

- (...) Ciência da Computação/Engenharia da Computação/Informática
 Engenharias/Matemática/Física/Química
 Ciências Humanas ou Ciências Sociais
 Ciências Biológicas e correlatas
 Ciências da Saúde

Atividade principal:

- Aluno de graduação
 Aluno de mestrado
 Aluno de doutorado
 Docente
 Profissional liberal
 Profissional em empresa de sua área de formação
 Profissional em empresa de área diferente da formação

Experiência como usuário de sistemas interativos (marque todas as aplicáveis):

- Sim, com sistemas *web* usuais (sites de compra, redes sociais, sites bancários, etc)
 Sim, com sistemas necessários às minhas atividades profissionais
 Sim, com jogos computacionais
 Não, apenas navego e faço consultas na *web*

Questionário SUS - System Usability Scale

São utilizadas afirmações que devem ser respondidas usando uma escala Likert de 5 opções: "Discordo totalmente", "Discordo", "Neutro", "Concordo" e "Concordo totalmente".

Id participante:

Eu acho que gostaria de usar essa aplicação com frequência.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu acho a aplicação desnecessariamente complexa.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu achei a aplicação fácil de usar.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu acho que precisaria de ajuda de uma pessoa com conhecimentos técnicos para usar a aplicação.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu acho que as várias funções da aplicação estão muito bem integradas.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu acho que a aplicação apresenta muita inconsistência.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu imagino que as pessoas aprenderão como usar a aplicação rapidamente.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu achei a aplicação confusa de usar.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu me senti confiante ao usar a aplicação.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Eu precisei aprender várias coisas novas antes de conseguir usar a aplicação.

	1	2	3	4	5	
Discordo totalmente	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	Concordo totalmente

Questionário aplicado na avaliação da *ClusterVis*

Este questionário deve ser respondido utilizando a técnica *ClusterVis*. A técnica está em desenvolvimento e, portanto, problemas detectados podem ser relatados no campo final de observações.

Participante (coloque o número recebido):

1) Qual o autor com maior quantidade de publicações da área Modelagem Conceitual e Banco de Dados (responda colocando o nome)?

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

2) O autor com maior quantidade de publicações é o mesmo que possui maior quantidade de publicações do tipo *Journal*?

Sim Não

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

3) O autor com maior quantidade de publicações é coautor do autor com a segunda maior quantidade de publicações?

Sim Não

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

4) O autor com a maior quantidade de publicações é coautor de autores classificados nas 3 categorias (discente, docente, participante externo)?

Sim Não

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

5) Entre os docentes coautores do autor com a maior quantidade de publicações, qual possui a menor quantidade de publicações (responda colocando o nome)?

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

Faça aqui os comentários que julgar necessário, inclusive informação sobre erros detectados.

Questionário aplicado na avaliação da *SoccerMatches (GlyphMatrix)*

Responda as questões indicando tanto a resposta como o seu grau de concordância com a afirmação.

Participante (coloque o número recebido):

1) Existem semelhanças na ordenação dos times na vertical segundo os diferentes critérios possíveis?

Sim Quais?

Não existem

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

2) Quais são as rodadas em que ocorreu apenas 1 empate?

Rodadas: _____

Não houve

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

3) Quais são as rodadas nas quais ocorreu apenas 1 derrota de time jogando em casa?

Rodadas: _____

Não houve

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

4) Quais são as rodadas em que não ocorreram derrotas dos times jogando em casa?

Rodadas: _____

Não houve

É fácil chegar à resposta da questão anterior.

Concordo fortemente Concordo Neutro Discordo Discordo fortemente

5) Em quais rodadas e jogos ocorreram os escores com maior diferença de gols?

Rodadas: _____

 Não houve

É fácil chegar à resposta da questão anterior.

 Concordo fortemente Concordo Neutro Discordo Discordo fortemente**6) Em quais jogos o time da casa cometeu o maior número de faltas?**

Jogos: _____

 Não houve

É fácil chegar à resposta da questão anterior.

 Concordo fortemente Concordo Neutro Discordo Discordo fortemente**7) Em quais jogos o time visitante cometeu o maior número de faltas?**

Jogos: _____

 Não houve

É fácil chegar à resposta da questão anterior.

 Concordo fortemente Concordo Neutro Discordo Discordo fortemente**8) Em quais jogos não ocorreram impedimentos (*Offsides*)?**

Jogos: _____

 Não houve

É fácil chegar à resposta da questão anterior.

 Concordo fortemente Concordo Neutro Discordo Discordo fortemente**9) Escolha um time específico: em quais jogos o número de faltas supera o número de desarmes (*Tackles*)?**

Time: _____

Jogos: _____

É fácil chegar à resposta da questão anterior.

 Concordo fortemente Concordo Neutro Discordo Discordo fortemente

10) Analisando os dois confrontos entre os times, é possível encontrar padrão semelhante nos dois jogos ao olhar as estatísticas?

() Sim Quais?_____

() Não existem

É fácil chegar à resposta da questão anterior.

() Concordo fortemente () Concordo () Neutro () Discordo () Discordo fortemente