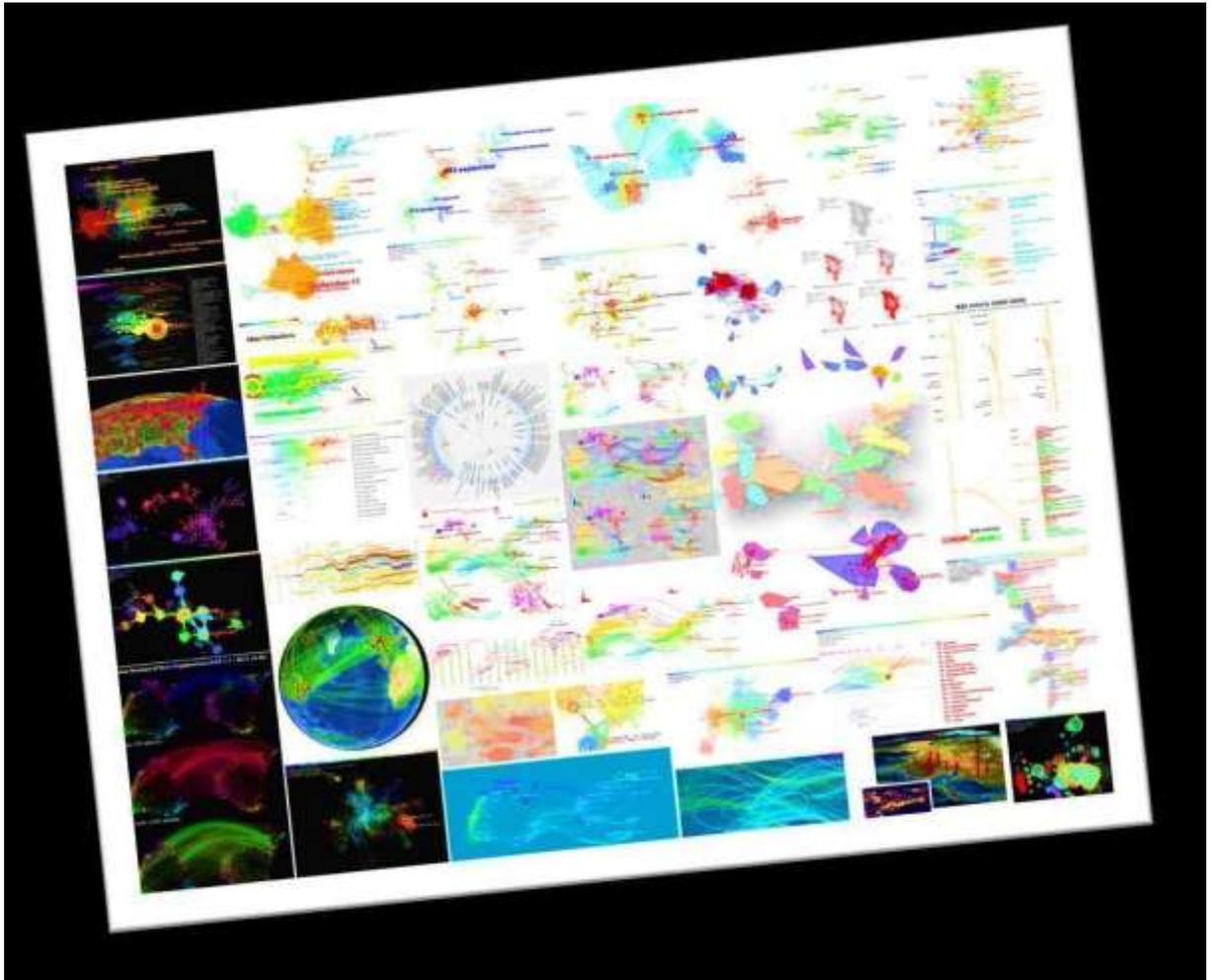


Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Manual do *CiteSpace*

Versão 1.01

Chaomei Chen
College of Computing and Informatics
Drexel University



Como cita:

Chen, Chaomei (2014) The *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

**Equipe de Tradução Livre do Manual do CiteSpace para o português brasileiro.
Grupo de Pesquisa de Recuperação da Informação e Tecnologias Avançadas (RITA) da
Universidade Federal de Santa Catarina (UFSC).**

Prof^o. Dr. Angel Freddy Godoy Viera. Líder do Grupo de Pesquisa RITA.

Msc. Alexandre Lucas. Membro do RITA.

Msc. Charles Rodrigues. Membro do RITA.

Fabrcio Foresti. Membro do RITA.

Maria Luiza Jurema Cassotta. Membro do RITA.

Paulo Roberto Freitas da Silva. Membro do RITA.

Florianópolis, Santa Catarina, Brasil, abril de 2016.

SUMÁRIO

1	Como posso encontrar a versão mais recente do manual do CiteSpace?	6
2	Para que eu posso usar o CiteSpace? Qual a finalidade do CiteSpace?	6
2.1	O que acontece se eu tiver dúvidas	8
2.2	Como devo fazer a citação do CiteSpace?	9
2.3	Onde estão os Usuários do CiteSpace?	10
3	Requisitos para executar o CiteSpace	12
3.1	Java Runtime Environment (JRE).....	12
3.2	Como posso verificar se o Java está instalado em meu computador?	13
3.3	Tenho um Computador de 32-bit ou 64-bit?.....	15
4	Como Instalar e Configurar CiteSpace.....	15
4.1	Onde posso baixar CiteSpace da Web?	15
4.2	Qual é o número máximo de registros que posso tratar com CiteSpace?	17
4.3	Como configurar a alocação de memória para CiteSpace?	17
4.4	Como desinstalar o CiteSpace.....	18
4.5	No Mac ou sistemas baseados em Unix	19
5	Começar com CiteSpace	24
5.1	Experimente com um conjunto de dados (dataset) de demonstração	24
5.1.1	O projeto de demonstração	25
5.1.2	<i>Clustering</i>	28
5.1.3	Gerar Etiquetas do <i>cluster</i>	30
5.1.4	Onde estão as principais áreas de pesquisa com base no <i>dataset</i> de entrada? 32	
5.1.5	Como estão conectadas essas grandes áreas?	33
5.1.6	Onde estão as áreas mais ativas?	34
5.1.6.1	Detecção de explosão (burst)	34
5.1.6.2	Detecção de explosão – Controles Adicionais	38
5.1.7	O que é importante sobre cada área? Qual /onde estão os artigos-chave para uma determinada área?.....	43
5.1.8	Visão da Linha do Tempo	45
5.2	Tente com seu próprio conjunto de dados (dataset)	46

5.2.1	Coleta de dados	46
5.2.1.1	Como construir meus próprios dados utilizando a Web of Science.....	46
5.2.1.2	Baixar registros para arquivos	49
5.2.2	Trabalhando com um projeto no CiteSpace	50
5.2.2.1	Criar um projeto CiteSpace	50
5.2.2.2	Editar um projeto existente	50
5.2.2.3	Limpar um Projeto	52
5.2.2.4	Remover um projeto	52
5.2.3	Fontes de Dados em Chinês	52
5.2.4	Como lidar com resultados da pesquisa contendo tópicos irrelevantes	53
6	Configurar a execução do CiteSpace.....	55
6.1	Time Slicing (intervalo de tempo).....	56
6.2	Processamento de texto (Text Processing).....	56
6.3	Configurando as redes	57
6.3.1	Acoplamento bibliográfico	57
6.4	Critérios de Seleção de Nó	58
6.4.1	Eu tenho a rede certa?	59
6.5	Pruning (Poda), ou Link Reduction (Redução de Link)	60
6.6	Visualização	60
7	Interagindo com o CiteSpace	61
7.1	Como Mostrar ou Ocultar a Intensidade dos Links.....	61
7.2	Adicionando um Rótulo Persistente ao Nó.....	61
7.3	Usando Pseudônimos para Fundir Nós	62
7.4	Como Excluir um Nó da Rede.....	65
7.5	Como utilizar o Controle Deslizante de Visualização Fisheye	65
7.6	Como Configurar Quando Calcular Automaticamente as Pontuações de Centralidade (Centrality Scores).....	66
7.7	Como Salvar a Visualização Como um Arquivo PNG	68
7.8	Filtros: Correspondência dos Registros na Base Pubmed	69
8	FUNÇÕES ADICIONAIS.....	73
8.1	Menu: Data (dados)	73
8.1.1	Banco de Dados incorporado no <i>CiteSpace</i>	74

8.1.2	Funções de Utilitário para o Formato <i>Web of Science</i>	77
8.1.3	Base de Dados Scopus	78
8.1.4	Base de Dados PubMed	80
8.2	Menu: Network	82
8.2.1	Lote de Exportação para arquivos Pajek.net	82
8.3	Menu: Geographical	82
8.3.1	Gerar mapas no <i>Google Earth</i>	82
8.4	Menu: Overlay Maps	86
8.4.1	Adicionar uma Sobreposição (<i>Overlay</i>)	87
8.4.2	Outras Leituras e Termos de Uso	90
8.5	Menu: Text	90
8.5.1	Árvores conceituais e árvores de predicado	90
8.5.2	Lista de termos por Propriedades de Agrupamento	94
8.5.3	Análise Semântica Latente	94
9	Exemplos Selecionados	96
10	MÉTRICAS E INDICADORES	97
10.1	Informações Teóricas	97
10.1.1	Entropia Informacional	97
10.2	Estrutural	98
10.2.1	Centralidade de Intermediação	98
10.2.2	Modularidade	98
10.2.3	Silhouette (silhueta)	98
10.3	Temporal	98
10.3.1	<i>Burstness</i> (explosão)	98
10.4	Combinado	99
10.4.1	Sigma	99
10.5	Rotulagens do Cluster	99
10.5.1	Frequência de Termo por frequência Invertida de Documento	99
10.5.2	Log Relação Probabilidade	99
10.5.3	Informação mútua	99
11	REFERÊNCIAS	100

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

1 Como posso encontrar a versão mais recente do manual do CiteSpace?

A última versão do Manual do *CiteSpace* está disponível em:

<http://cluster.cis.drexel.edu/~cchen/citespace/>

Você também pode acessar o manual no menu do *CiteSpace*: **Help ► How to Use CiteSpace** (eBook). Ele irá abrir o endereço onde pode ser comprado o manual.



Figura 1. A última versão do Manual do *CiteSpace* é acessível a partir do menu *Help* do *CiteSpace*.

2 Para que eu posso usar o CiteSpace? Qual a finalidade do CiteSpace?

CiteSpace é projetado para responder às perguntas sobre o domínio do conhecimento, o que é uma definição ampla que abrange um campo científico, uma área de investigação, ou de uma disciplina científica. O domínio do conhecimento é geralmente representado por um conjunto de registros bibliográficos de publicações relevantes. É de sua responsabilidade preparar os mais adequados e representativos conjuntos de dados (*dataset*) que contenham a informação adequada para responder às suas perguntas.

CiteSpace é projetado para facilitar as respostas das perguntas sobre a estrutura e a dinâmica de um domínio de conhecimento. Aqui estão algumas perguntas frequentes:

- Quais são as principais áreas de investigação com base no conjunto de dados de entrada?
- Como essas grandes áreas estão ligadas, por exemplo, através de quais artigos específicos?
- Onde estão as áreas mais ativas?
- O que é importante para cada área? Qual/onde estão os principais artigos para uma determinada área?
- Existem transições críticas na história do desenvolvimento do campo de estudo? Onde estão os “pontos de mudança”?

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

O *design* do *CiteSpace* é inspirado na estrutura das revoluções científicas de Thomas Kuhn. A ideia central é que o foco dos centros de pesquisa muda com o passar do tempo, às vezes ‘incrementalmente’ e outras vezes drasticamente. O desenvolvimento da ciência pode ser rastreado estudando seus caminhos revelados pelas publicações acadêmicas.

Membros da comunidade científica contemporânea dão sua contribuição. As suas contribuições formam um dinâmico e auto-organizado sistema de conhecimento. O sistema contém consenso, conflitos, incertezas, hipóteses, mistérios e problemas não resolvidos e perguntas não respondidas. Não é suficiente o estudo de uma única escola de pensamento. Na verdade, uma melhor compreensão de um tópico específico, muitas vezes depende de uma compreensão do modo como ela está relacionada a outros temas.

A fundamentação do *CiteSpace* é análise de rede e visualização. Por meio de uma rede de modelagem e visualização pode-se explorar a paisagem intelectual de uma área de conhecimento e discernir quais perguntas os investigadores têm tentado responder e quais os métodos e as ferramentas eles desenvolveram para atingir seus objetivos.

Esta não é uma tarefa simples. Em vez disso, é muitas vezes conceitualmente exigente e complexa. Se você está prestes a escrever um romance, o processador de texto ou um editor de texto pode tornar a tarefa mais fácil, mas ele não pode ajudá-lo a criar o enredo ou enriquecer o caráter de seu herói. Da mesma forma e provavelmente em uma maior extensão, *CiteSpace* pode gerar uma raio-X de uma área do conhecimento, mas para interpretar esse raio-X, é necessário ter algum conhecimento de vários elementos envolvidos.

O papel do *CiteSpace* é assumir algumas das tradicionais sobrecargas de trabalho através dos algoritmos computacionais e visualizações interativas, para que se possa concentrar no que os usuários humanos são melhores: na resolução de problemas e encontrar a verdade. No entanto, é provavelmente mais fácil gerar algumas misteriosas visualizações com *CiteSpace* do que, compreender plenamente o que estas visualizações mostram e quem pode se beneficiar de tais descobertas.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Também se podem postar perguntas no blog *sciencenet*:

<http://blog.sciencenet.cn/home.php?mod=space&uid=496649>

Por favor, evite enviar *e-mails* porque há uma possibilidade muito maior de conseguir a resposta no *Facebook* ou no *blog sciencenet*.

Em geral, perguntas significativas são respondidas rapidamente. Perguntas que você pode ser capaz de descobrir a resposta por si mesmo, se você pensar um pouco mais sobre o assunto tem uma menor prioridade na fila do atendimento e é possível que algumas delas nunca sejam respondidas.

2.2 Como devo fazer a citação do *CiteSpace*?

As três publicações a seguir representam as ideias fundamentais do *CiteSpace*.

O artigo 2004 *PNAS* é a primeira publicação sobre o *CiteSpace* (CHEN, 2004). Em retrospectiva, ele poderia ter sido chamado *CiteSpace* I. A página 19 do artigo *JASIST* 2006 dá a mais completa e profunda descrição das principais funções do *CiteSpace* II (CHEN, 2006) e um acompanhamento de especialistas nos domínios identificados nas visualizações. O *JASIST* 2010 é um artigo com mais 24 páginas (CHEN, IBEKWE-SANJUAN, HOU, 2010), que é o terceiro da trilogia. São descritos detalhes técnicos sobre como as etiquetas dos agrupamentos (*clusters*) são selecionadas e como cada um dos três algoritmos de seleção em comparação com etiquetas que foram selecionadas por especialistas da área.

Citações (<i>Google Scholar</i>)	Referência
800	Chen, C. (2006). "CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature." <i>Journal of the American Society for Information Science and Technology</i> 57(3): 359-377.
394	Chen, C. (2004). "Searching for intellectual turning points: Progressive Knowledge Domain Visualization." <i>Proc. Natl. Acad. Sci. USA</i> 101(Suppl.): 5303-5310.
157	Chen, C., et al. (2010). "The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis." <i>Journal of the American Society for Information Science and Technology</i> 61(7): 1386-1409.

O mais recente estudo de caso de um tópico fora da área de Ciência da Informação e da Cientometria é um estudo cientométrico de medicina regenerativa (CHEN, HU, LIU, TSENG, 2012).

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Chen, C., et al. (2012). "Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace." *Expert Opinions on Biological Therapy* **12**(5): 593-608.

2.3 Onde estão os Usuários do CiteSpace?

Em termos das cidades onde *CiteSpace* foi utilizado, a China, os Estados Unidos e a Europa são proeminentes. Brasil, Turquia e Espanha também têm muitas cidades no gráfico.



Figura 3. Cidades com usuários de *CiteSpace* entre Agosto de 2013 e Março de 2014 são mostrados no mapa. As cores dos marcadores retratam o nível de intensidade de usuários: o verde (1 - 10), amarelo (10-100), vermelho (100-1000) e o marcador em forma de gota vermelha (1000+).



Figura 4. O uso de *CiteSpace* na China (Agosto de 2013 - Março de 2014).



Figura 5. O uso de *CiteSpace* nos Estados Unidos (Agosto de 2013 - Março de 2014).



Figura 6. O uso do *CiteSpace* na Europa (Agosto de 2013 - Março de 2014).

3 Requisitos para executar o *CiteSpace*

3.1 *Java Runtime Environment (JRE)*

O *CiteSpace* é escrito em Java. Trata-se de uma aplicação Java. Você deve ser capaz de executá-lo em um computador que suporte Java, incluindo Windows ou Mac.

CiteSpace está otimizado para o Windows 64-bits Java 7 (ou seja, Java 1.7).

Para executar uma aplicação Java em seu computador, é necessário ter o *Java Runtime (JRE)* instalado.

3.2 Como posso verificar se o Java está instalado em meu computador?

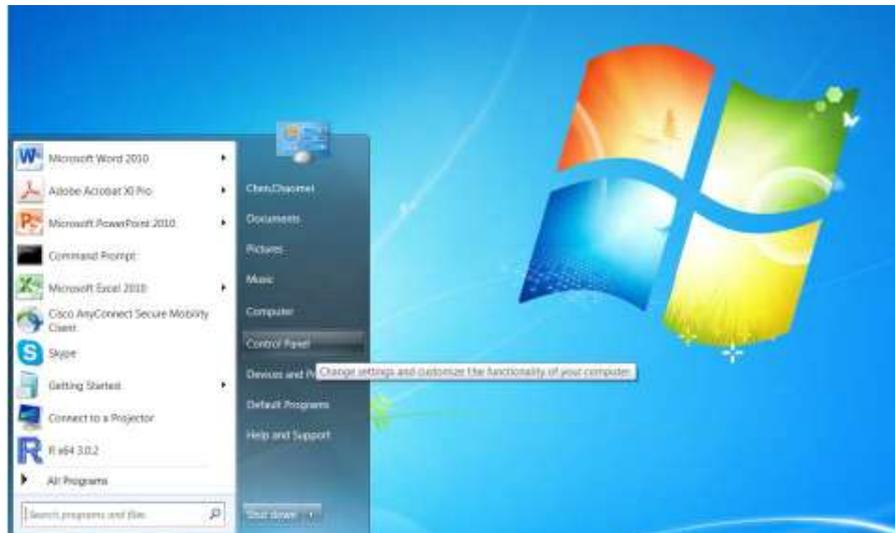


Figura 7. Selecione Painel de controle.



Figura 8. Clique em Programas para encontrar o painel de controle Java.

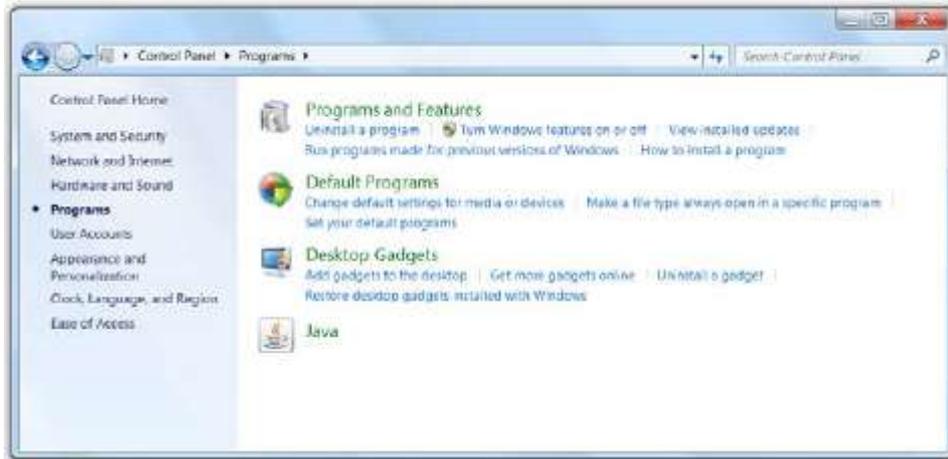


Figura 9. Localize o painel de controle Java.



**Figura 10. Painel de controle Java.
Escolha a guia Java e pressione o botão *View* para ver mais detalhes.**

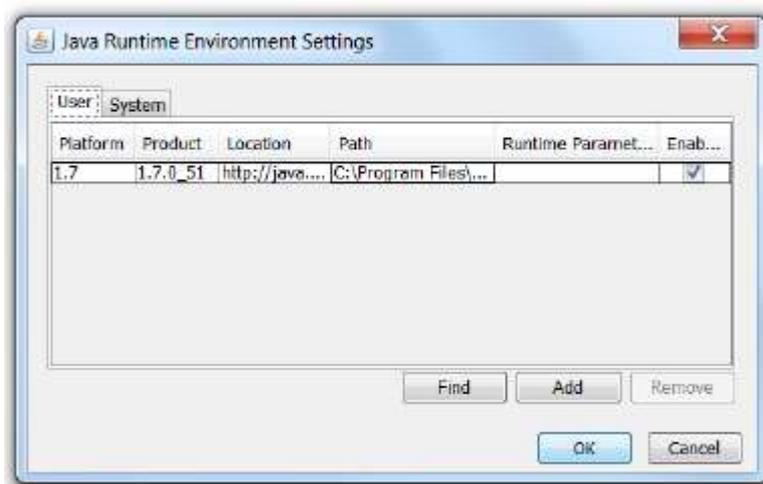


Figura 11. Java *Runtime* 1.7 está instalado.

3.3 Tenho um Computador de 32-bit ou 64-bit?

É necessário saber se o seu computador possui um sistema operacional de 32 bits ou de 64 bits.

Vá ao **Painel de Controle** ► **Sistema e Segurança** ► **Sistema**. Você irá ver vários detalhes sobre o computador. No item tipo de sistema, você verá se possui um sistema operacional de 32 bits ou de 64 bits.

Siga o *link* abaixo para obter mais instruções sobre como instalar o Java. http://www.java.com/en/download/help/index_installing.xml

Uma vez que se tenha o Java configurado em seu computador, pode-se prosseguir com a instalação do *CiteSpace*.

4 Como Instalar e Configurar *CiteSpace*

CiteSpace é fornecido como um arquivo zip para computadores de 64-bits e 32-bits. Para usuários Mac, é preciso baixar a versão de 64 bits.

4.1 Onde posso baixar *CiteSpace* da Web?

Pode-se fazer o *download* da versão mais recente do *CiteSpace* no seguinte *site*: <http://cluster.ischool.drexel.edu/~cchen/citespace/download/>

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Date	Version	WebStart	Download	Download	Video
January 29, 2014	3.8.R1 (64-bit)	1 GB RAM Required	zip	64-bit / Windows x64	video
January 29, 2014	3.8.R1 (32-bit)		zip	32-bit / Windows x86	
September 27, 2013	3.7.R7 (64-bit)	launch	zip	64-bit / Windows x64	
September 27, 2013	3.7.R7 (32-bit)	launch	zip	32-bit / Windows x86	

Figura 12. Página de *download* do *CiteSpace*.

Depois de “baixado” o arquivo zip em seu computador, descompacte o arquivo zip para uma pasta de sua escolha.

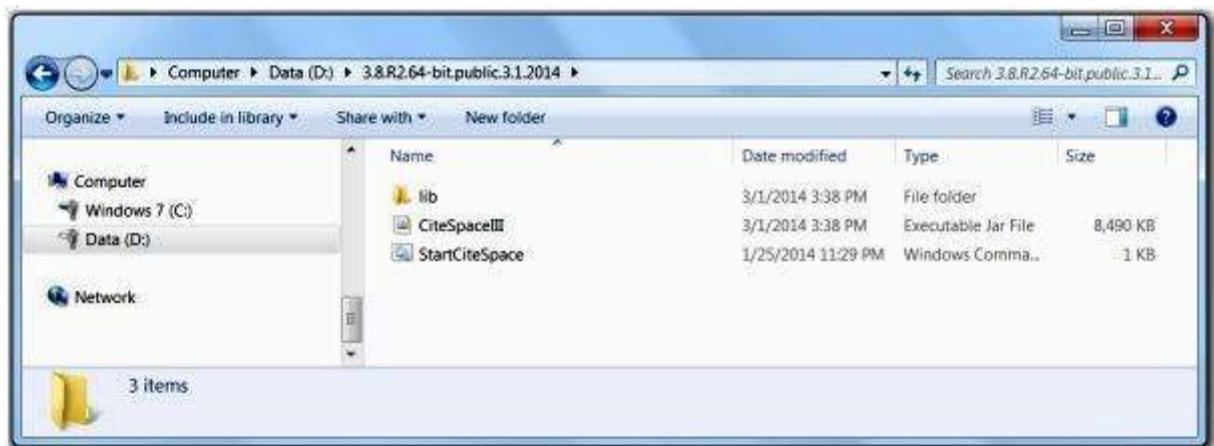


Figura 13. *CiteSpace* é descompactado para a unidade D em um computador.

Agora se pode iniciar *CiteSpace*, clicando duas vezes sobre o arquivo *StartCiteSpace*.

Se precisar modificar a quantidade de memória alocada para *CiteSpace* (mais precisamente para a Máquina Virtual Java no qual *CiteSpace* é executado), pode-se editar *StartCiteSpace* como um arquivo de texto simples com qualquer editor de texto.

4.2 Qual é o número máximo de registros que posso tratar com *CiteSpace*?

Esta pergunta precisa ser respondida em dois níveis: o número de registros processados pelo *CiteSpace* e o número de nós visualizados, ou seja, pode-se ver e interagir com eles no *CiteSpace*.

O primeiro número é o número total de registros no conjunto de dados (*dataset*) que foi baixado. O *CiteSpace* lê por meio de cada registro em seus arquivos de *download*.

O segundo número é determinado pelos critérios de seleção que você especificar e pela quantidade de memória RAM, disponível no seu computador. Quanto mais memória RAM você disponibilizar para *CiteSpace*, maior o porte da rede que você pode visualizar com uma taxa de resposta mais rápida.

A velocidade de processamento também é afetada por alguns poucos e “pesados” algoritmos computacionais como *Pathfinder network scaling* (guia de dimensionamento da rede) e *cluster labeling* (Rotulagem de Clusters). Empiricamente, as melhores opções para *Pathfinder network scaling* seria 50~500 nós por fatia. Com computadores mais rápidos ou se você pode esperar um pouco mais, é possível aumentar o número de forma correspondente.

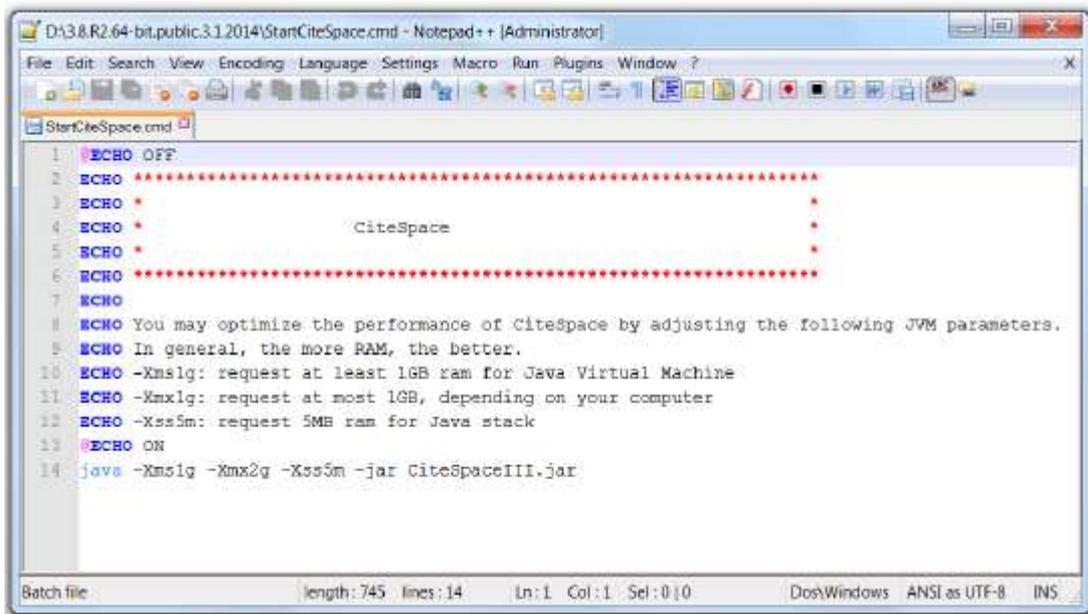
O tempo de conclusão da Rotulagem de *Clusters* está relacionado ao tamanho do seu *dataset*. Se todo o período de vigência do seu *dataset* é de 100 anos, mas se for considerado os 10 anos mais recentes, será uma boa ideia configurar um *dataset* muito menor abrangendo somente os 10 anos de interesse. Isso irá reduzir o tempo de processamento consideravelmente.

4.3 Como configurar a alocação de memória para *CiteSpace*?

O desempenho do *CiteSpace* é influenciado pela quantidade de memória disponível para a Máquina Virtual Java (JVM) sobre o qual *CiteSpace* está funcionando. Para analisar uma grande quantidade de registros, deve-se considerar atribuir mais memória o quanto for necessária para o *CiteSpace* utilizar.

Você pode modificar o arquivo *StartCiteSpace.cmd* para otimizar a configuração. Mais especificamente, modificar a linha 14 do arquivo. Por exemplo, `-Xmx2g` significa que *CiteSpace* pode obter um máximo de 2GB de RAM para seu trabalho. Após a alteração, salvar o arquivo e reiniciar o *CiteSpace*.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

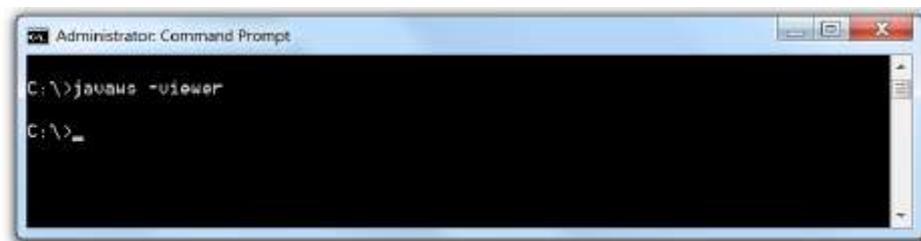


```
1: ECHO OFF
2: ECHO *****
3: ECHO *
4: ECHO *          CiteSpace
5: ECHO *
6: ECHO *****
7: ECHO
8: ECHO You may optimize the performance of CiteSpace by adjusting the following JVM parameters.
9: ECHO In general, the more RAM, the better.
10: ECHO -Xms1g: request at least 1GB ram for Java Virtual Machine
11: ECHO -Xmx1g: request at most 1GB, depending on your computer
12: ECHO -Xss5m: request 5MB ram for Java stack
13: ECHO ON
14: java -Xms1g -Xmx2g -Xss5m -jar CiteSpaceIII.jar
```

Figura 14. Configurar a memória para Java na linha 14.

4.4 Como desinstalar o *CiteSpace*

Pode-se usar as seguintes etapas para extrair os arquivos em *cache* do *CiteSpace* do seu computador.



```
Administrator: Command Prompt
C:\>javaws -viewer
C:\>_
```

Figura 15. Em uma janela de *Prompt* de Comando, digite `javaws -viewer`

Quando você ver a lista de cópias em *cache* do *CiteSpace* no *Java Cache Viewer*, selecione os itens que você deseja remover e, em seguida, clique sobre o botão com o X em vermelho.

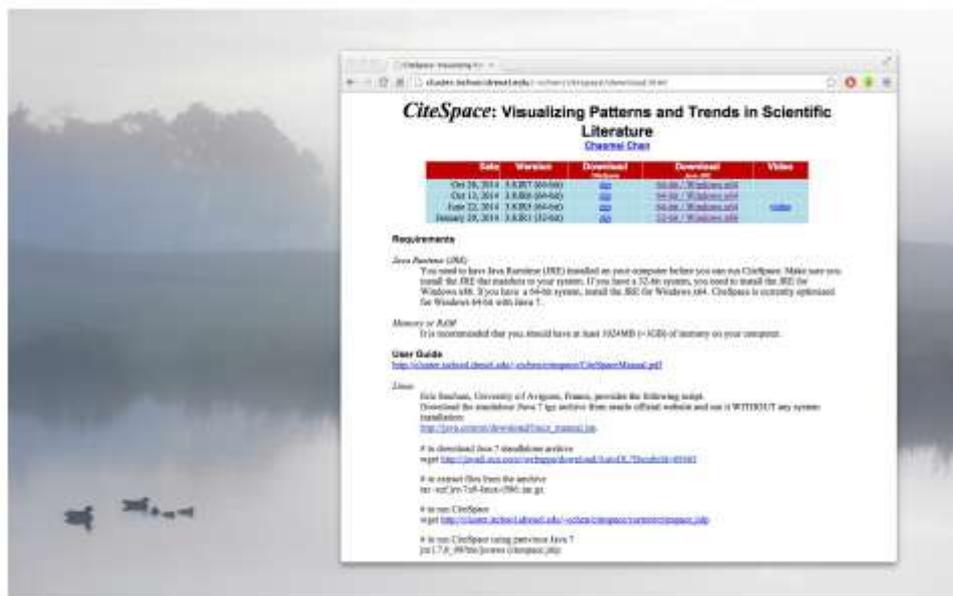
Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 16. Selecione a cópia em *cache* de *CiteSpace* e remova o item.

4.5 No Mac ou sistemas baseados em Unix

O exemplo a seguir mostra os passos básicos para começar a trabalhar com *CiteSpace* em um Mac. Primeiro, vá para a *homepage* do *CiteSpace* no seu navegador, como Chrome, e baixe a última versão de 64 bits.



Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Figura 17. Em um Mac, vá para a *homepage* do *CiteSpace* no navegador, como Chrome e baixe a última versão de 64 bits.

Assim que o *download* for concluído, siga a opção "*Show in Finder*". Ela irá levá-lo a uma lista de arquivos baixados para o seu Mac. O arquivo mais recente deve ser o arquivo zip para *CiteSpace*.

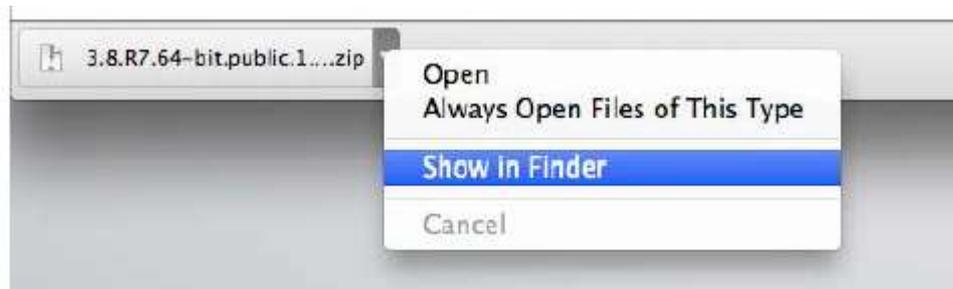


Figura 18. Selecione a opção "*Show in Finder*".

Name	Date Modified	Size
3.8.R7.64-bit.public.10.26.2014.zip	Today, 8:27 PM	25.4 MB

Figura 19. O arquivo zip baixado é mostrado em seu *Finder*.

Clique duas vezes sobre o arquivo zip para descompactar o arquivo para a pasta atual.

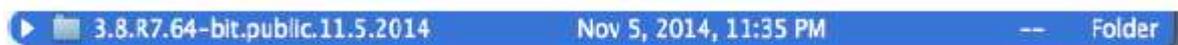


Figura 20. O arquivo zip é descompactado para uma nova pasta na lista.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

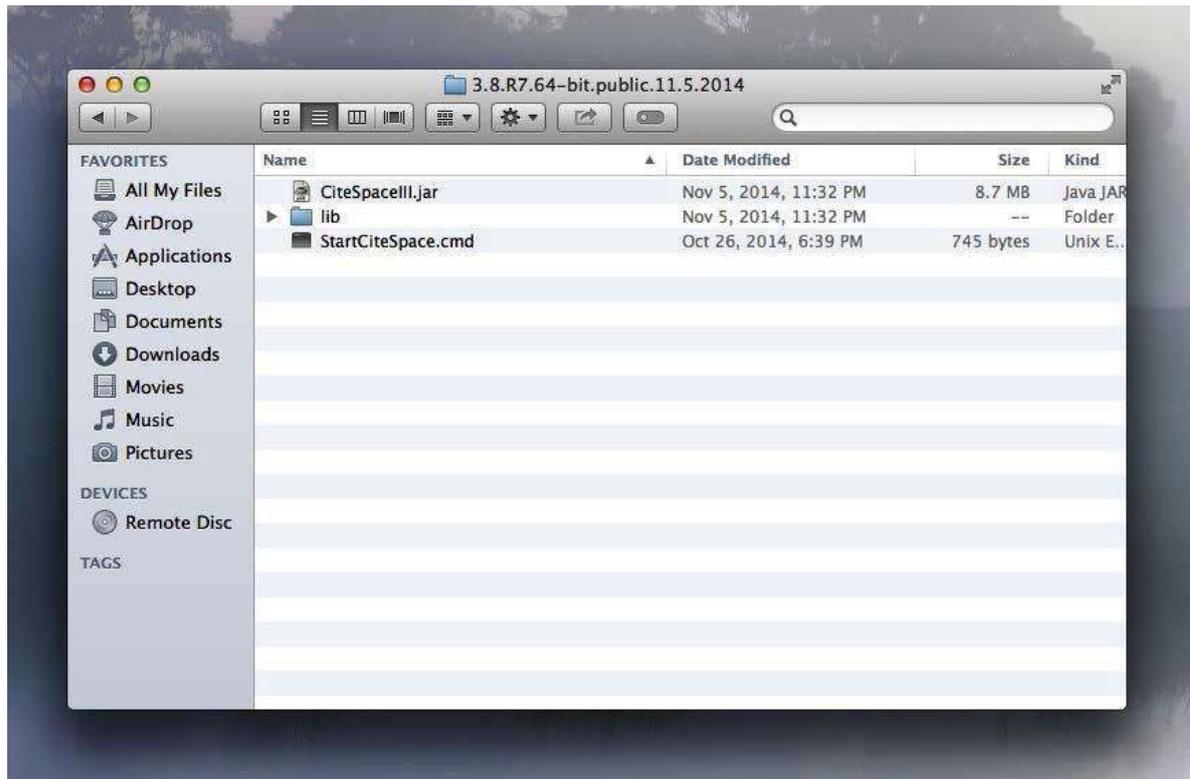


Figura 21. A nova pasta contém *CiteSpaceII.jar* e uma pasta *lib*.

A forma mais simples de começar o *CiteSpace* é abrir a *CiteSpaceII.jar*, clicando sobre ele, segurando a tecla "*Control*" no Mac. Selecione *Open* no Menu *pop-up*.

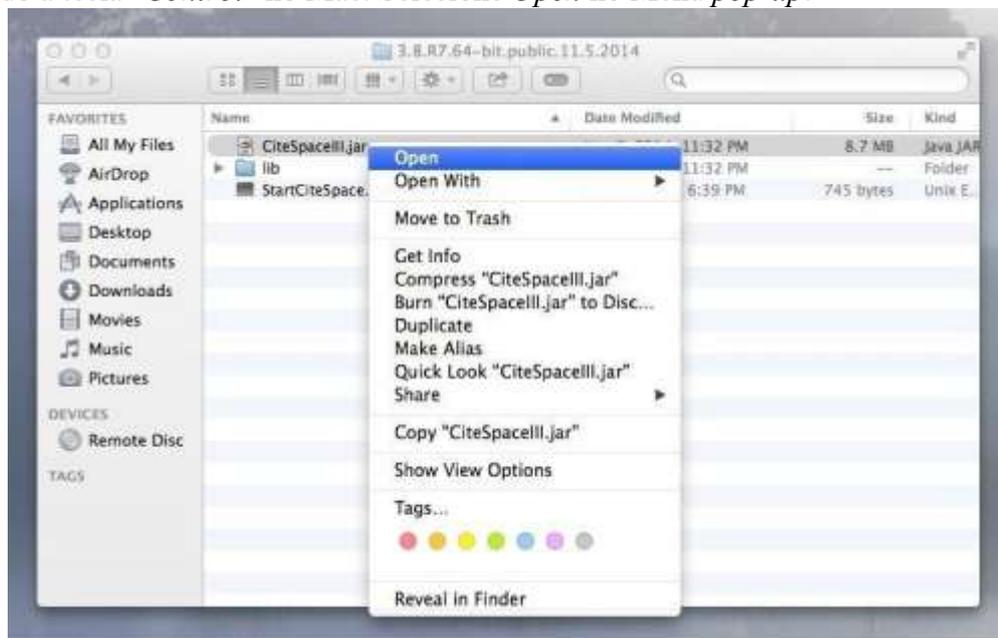


Figura 22. Clique sobre o *CiteSpaceII.jar*, segurando a tecla "*Control*" e selecione "*Open*".

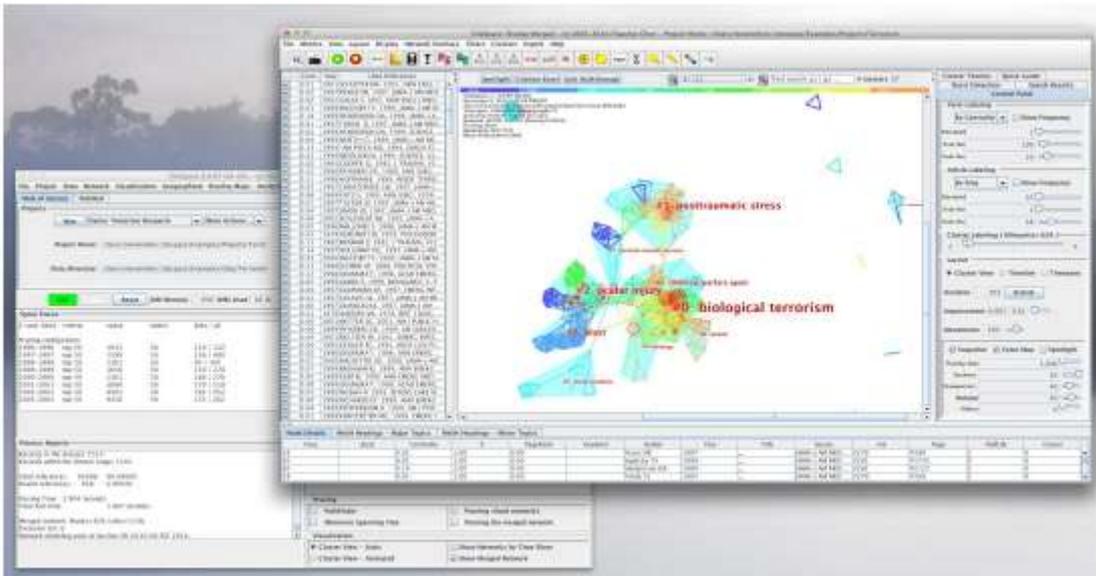


Figura 25. Imagens da execução do projeto de demonstração de *CiteSpace* no Mac.

É uma boa ideia para se familiarizar com as funções básicas de *CiteSpace* dar uma visualizada por meio do projeto de demonstração sobre o terrorismo, que está incluído no arquivo zip.

Se você deseja configurar diversos parâmetros da Máquina Virtual Java com mais detalhes do que é mostrado no exemplo acima, você pode gerar um arquivo *bash* para o seu Mac como se segue.

O Mac equivalente do *StartCiteSpace.cmd* seria um arquivo *bash*, que deve ter uma extensão de arquivo *.sh* e deve ser executável. Vamos nomear o arquivo como *StartCiteSpace.sh* para ser coerente.

1. O conteúdo do arquivo *StartCiteSpace.sh* deve ter as duas linhas seguintes:

```
#!/bin/bash
java -Xms1g -Xmx4g -Xss5m -jar CiteSpaceIII.jar
```

2. As instruções a seguir converte o arquivo *StartCiteSpace.sh* para um arquivo executável.

```
chmod +x StartCiteSpace.sh
```

3. Para chamar o arquivo executável, basta digitar seu nome ou clicar duas vezes sobre ele.

```
StartCiteSpace
```

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

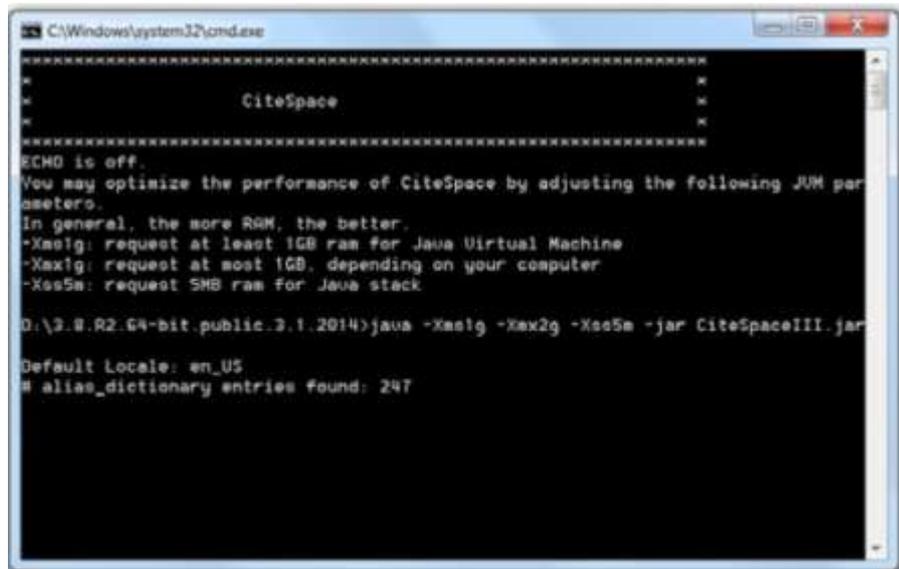
5 Começar com CiteSpace

5.1 Experimente com um conjunto de dados (dataset) de demonstração

Quando você instalou o *CiteSpace* pela primeira vez, um conjunto de dados demonstrativo sobre a pesquisa terrorismo foi configurado para você brincar e se familiarizar com as principais funções analíticas do *CiteSpace*.

Se você nunca usou o *CiteSpace*, recomenda-se que você comece com este conjunto de dados de demonstração.

Para iniciar o *CiteSpace*, clique duas vezes no arquivo *StartCiteSpace.cmd*. Abrirá uma janela do *prompt* de comando primeiro. Esta janela também exibirá várias informações sobre o *status* e eventuais erros.



```
C:\Windows\system32\cmd.exe
CiteSpace
ECHO is off.
You may optimize the performance of CiteSpace by adjusting the following JVM parameters.
In general, the more RAM, the better.
-Xms1g: request at least 1GB ram for Java Virtual Machine
-Xmx1g: request at most 1GB, depending on your computer
-Xss5m: request 5MB ram for Java stack
D:\3.8.R2.64-bit.public.3.1.2014>java -Xms1g -Xmx2g -Xss5m -jar CiteSpaceIII.jar
Default Locale: en_US
# alias_dictionary entries found: 247
```

Figura 26. A janela de *prompt* de comando.

Abrirá outra janela "*About CiteSpace*" - ela exibe informações do sistema do seu computador, incluindo a versão Java.

Para continuar, é necessário clicar no botão *Agree* (Concordo). O *CiteSpace* pode coletar eventos dirigidos de usuários para fins de pesquisa.

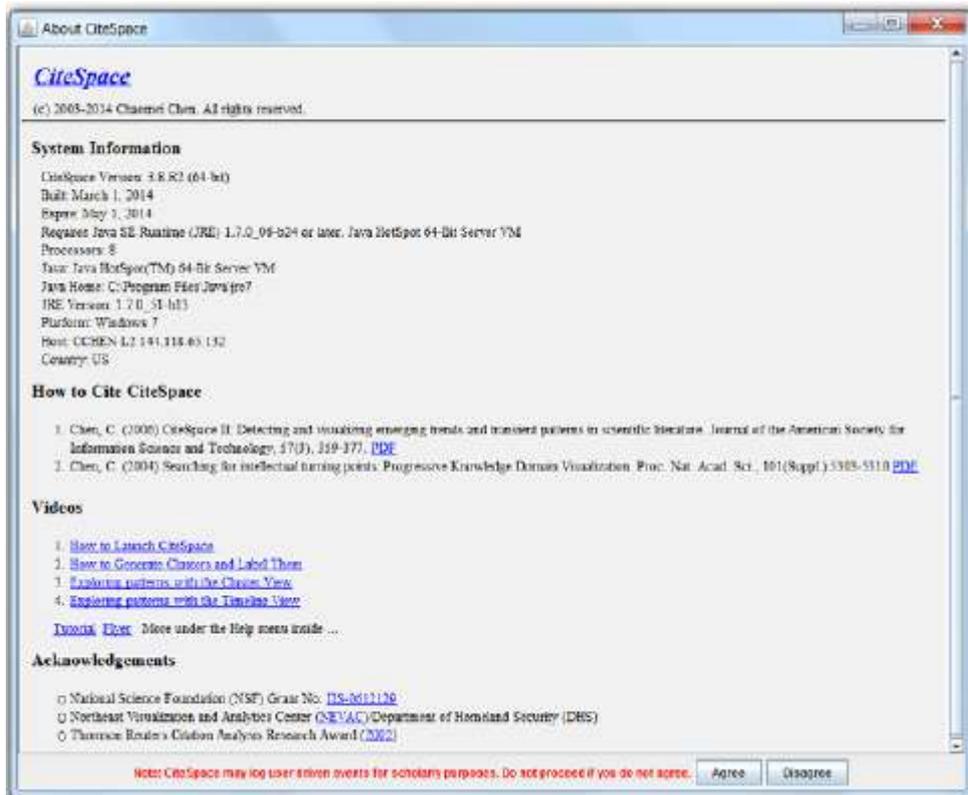


Figura 27. A janela "AboutCiteSpace". Para prosseguir, clique no botão *Agree*.

Em seguida, você verá a interface principal do usuário do *CiteSpace*.

A interface do usuário está dividida em duas metades: esquerda e direita. O lado esquerdo contém controles de projetos (ou seja, *datasets* de entrada) e janelas de relatório de progresso. O lado direito contém vários painéis para a configuração do processo com vários parâmetros.

O processo em *CiteSpace* leva um *dataset* de entrada especificada no projeto atual, construção de modelos de rede de entidades bibliográficas, e visualiza as redes para exploração interativa para as tendências e padrões identificados a partir do *dataset*.

O projeto de demonstração contém um dataset em publicações sobre a pesquisa do terrorismo. Estes registros bibliográficos foram recuperados da *Web of Science*. Veja os capítulos posteriores sobre dicas de como construir o seu próprio *dataset*.

5.1.1 O projeto de demonstração

Vamos iniciar o processo e explicar como *CiteSpace* é projetado para ajudá-lo a responder a algumas das questões-chave sobre uma área de conhecimento, ou seja, um campo de estudo, uma área de pesquisa, ou um conjunto de publicações definido pelo usuário.

Pressione o botão verde "GO!" para iniciar o processo.

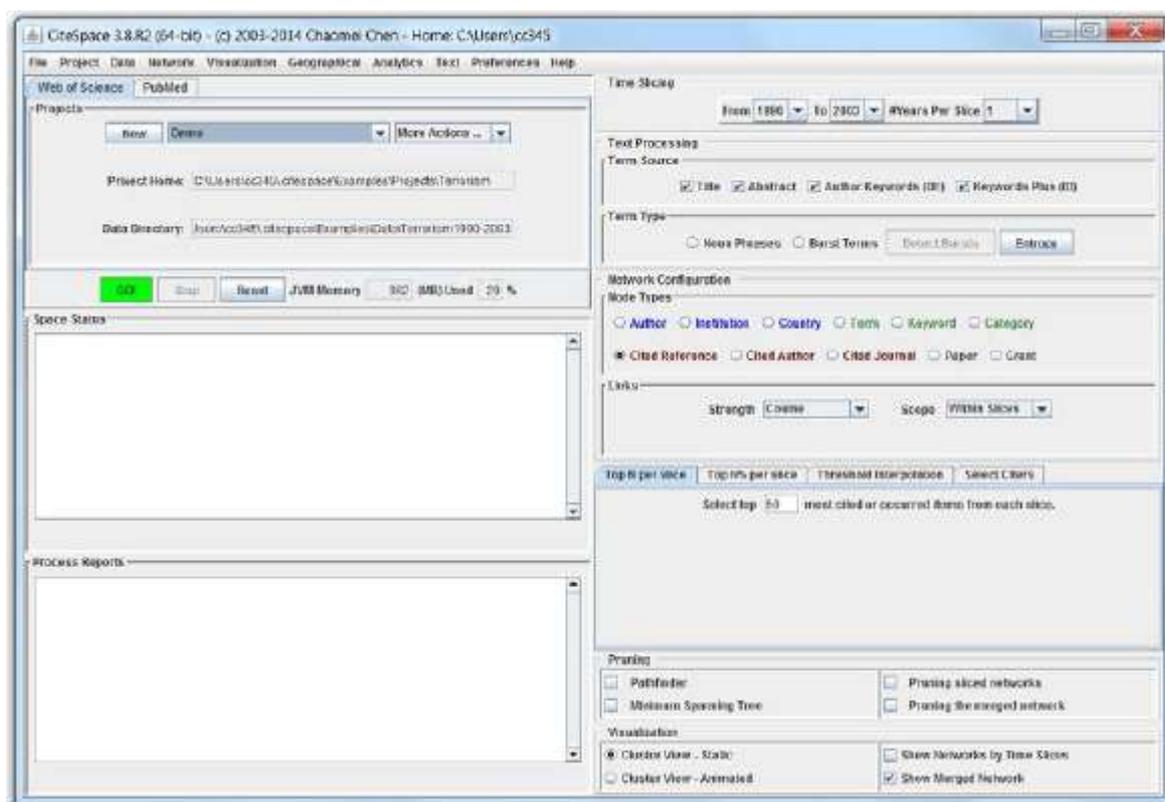


Figura 28. A interface principal do usuário do *CiteSpace*.

O *CiteSpace* irá ler os arquivos de dados do projeto atual (Demo) e relatar seu progresso nas duas janelas no lado esquerdo da interface do usuário. Quando o processo de modelagem for concluído, haverá três opções para escolher: *Visualize*, *Save As GraphML*, ou *Cancel*.

Visualize:

Esta opção abrirá uma janela de visualização para uma maior exploração interativa.

Save As GraphML:

Esta opção salvará a rede construída num arquivo de um formato de gráfico comum. Nenhuma visualização.

Cancel:

Esta opção não gerará qualquer visualização interativa nem salvará os arquivos. Ela permite que você reconfigure e repita o processo.

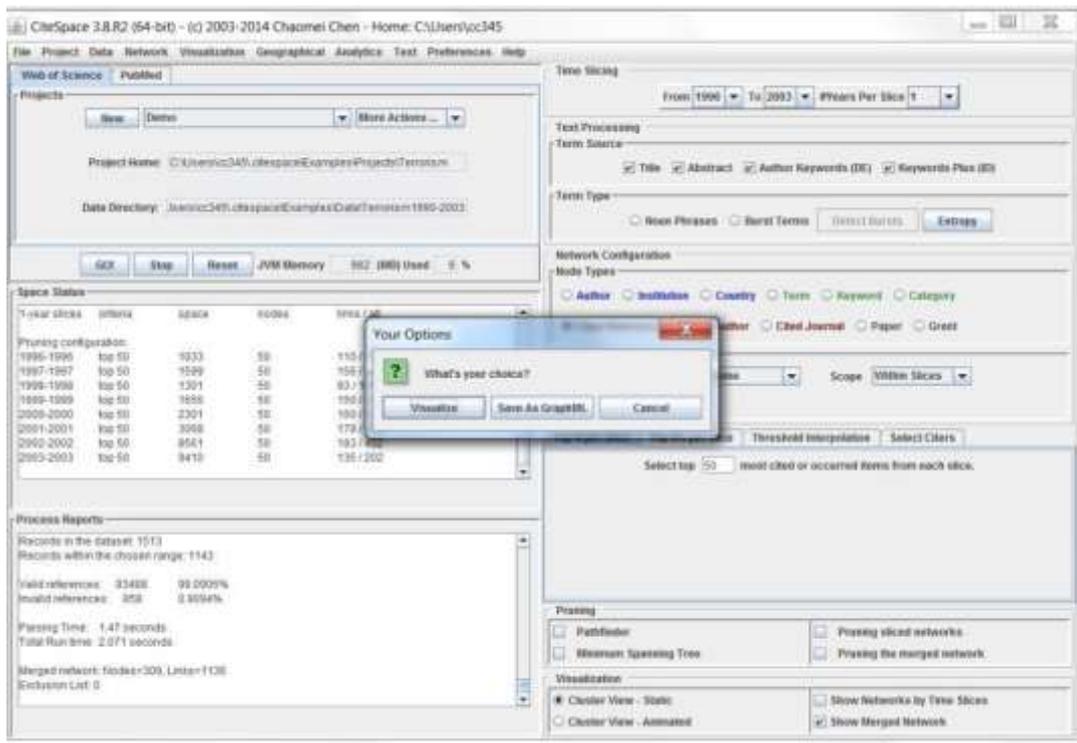


Figura 29. *CiteSpace* está pronto para visualizar a rede construída.

Se você clicar no botão *Visualize*, uma nova janela aparecerá. Esta é a *Visualization Window*. Inicialmente você verá alguns movimentos na tela com um fundo preto. Uma vez que os movimentos são resolvidos, o fundo preto torna-se branco.

Vamos nos concentrar no que a visualização inicial nos diz e, em seguida, explorar o que mais podemos encontrar usando as funções adicionais.

Primeiro, o *CiteSpace* visualiza uma rede mesclada com base em várias redes correspondentes aos instantâneos de anos consecutivos. No exemplo do projeto de demonstração, a extensão total do tempo é de 1996 a 2003. A rede resultante da fusão caracteriza o desenvolvimento do campo ao longo do tempo, mostrando os rastros mais importantes das atividades de pesquisa relacionadas. Cada ponto (*dot*) representa um nó na rede. No caso da demonstração, os nós são citações. O *CiteSpace* pode gerar redes de outros tipos de entidades. Aqui, por enquanto, vamos nos concentrar nas citações. Linhas que conectam os nós são *links* de co-citações, novamente, o *CiteSpace* pode gerar redes de outros tipos de *links*. As cores destas linhas são projetadas para mostrar quando uma conexão foi feita pela primeira vez. Note-se que isto é influenciado pelo escopo e pela profundidade do *dataset* analisado.

A codificação de cores torna mais fácil para identificar qual parte da rede é antiga e qual é a nova.

Se aparecerem algumas referências com rótulos, então saberemos que essas referências são altamente citadas, sugerindo que elas são, provavelmente, artigos de referência no campo. A lista no lado esquerdo da janela mostra todos os nós apresentados na visualização. A lista

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

nos dizem sobre as propriedades estruturais globais da rede. Por exemplo, a modularidade Q de 0,8149 é relativamente elevada, o que significa que a rede está razoavelmente dividida em *clusters* de baixo acoplamento. A silhueta significativa de pontuação de 0,39 sugere que a homogeneidade destes aglomerados, em média, não é muito elevada, e levemente baixa.

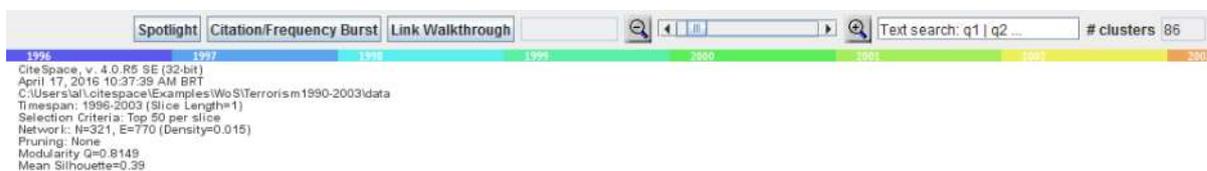


Figura 32. O processo de *clustering* está concluído. São identificados 86 *clusters* (#clusters mostrado no canto superior direito).

Resultados da Modularidade e silhueta são mostrados na assinatura de rede do lado esquerdo.

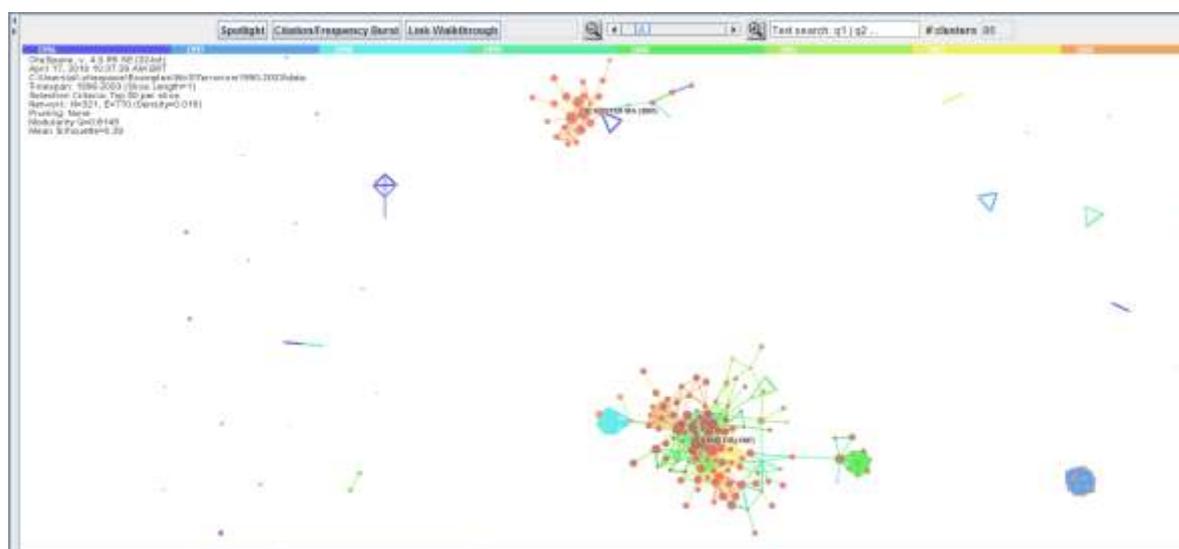


Figura 33. Membros de diferentes *clusters* são mostrados em diferentes cores.

É possível inspecionar várias medidas de cada *cluster* em um quadro-resumo de todos os *clusters* usando: **Clusters** ► **4. Summarization of Clusters**. Isto só funciona após ter gerado a etiqueta do *cluster*, item seguinte. A coluna da silhueta mostra a homogeneidade de um aglomerado. Quanto maior a pontuação da silhueta, mais consistentes são os membros do *cluster*, desde que os *clusters* em comparação tenham tamanhos semelhantes. Se o tamanho do *cluster* é pequeno, então a alta homogeneidade não significa muito. Por exemplo, o *cluster* # 10 tem 8 membros e uma silhueta de 1,00, isso é mais provável, devido à possibilidade de que todos as 8 referências são as referências de citação do mesmo autor subjacente. Em outras palavras, *cluster* # 10 pode refletir o comportamento de citação ou preferências de um único artigo, então, é menos representativo.

O ano médio de publicação de um *cluster* indica se ele é formado por artigos geralmente recentes ou antigos. Este é um indicador simples e útil.

Cluster	Cluster ID	Size	Silhouette	Modularity	Weighted Mean Silhouette	Top Terms (tf*idf weights)	Top Terms (log-likelihood ratio test)	Terms (linked information)
0	0	54	0.892	0.892	0.892	emergencies (17.41), emergency (17.41), eme...	emergencies (17.41), emergency (17.41), eme...	emergencies
1	1	21	0.84	0.84	0.84	procedures (16.1), emergency (16.1), eme...	procedures (16.1), emergency (16.1), eme...	procedures
2	2	20	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
3	3	20	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
4	4	14	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
5	5	11	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
6	6	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
7	7	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
8	8	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
9	9	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
10	10	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
11	11	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
12	12	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies
13	13	10	0.892	0.892	0.892	agencies (16.1), emergency (16.1), eme...	agencies (16.1), emergency (16.1), eme...	agencies

Figura 34. Uma tabela resumo dos *clusters*.

5.1.3 Gerar Etiquetas do *cluster*

Para caracterizar a natureza de um *cluster* identificado, o *CiteSpace* pode extrair frases nominais dos títulos (T no ícone a seguir), listas de palavras-chave (K), ou resumos (A) de artigos que citavam o *cluster* em particular.

Vamos pedir ao *CiteSpace* para localizar frases nominais por títulos (ou seja, selecione o ícone T). Este processo pode demorar um pouco, pois o *CiteSpace* precisa calcular várias métricas de seleção. Uma vez que o processo for concluído, os rótulos escolhidos serão exibidos. Por padrão, rótulos com base em um dos três algoritmos de seleção serão mostrados, a saber, $tf*idf$. Nosso estudo descobriu que LLR geralmente dá o melhor resultado em termos de exclusividade e cobertura.



Figura 35. Ícones para executar funções de *Clustering* e rotulagem.

Etiquetas de *cluster* são exibidas uma vez que o processo é concluído. Os *clusters* são numerados na ordem decrescente do tamanho do *cluster*, começando pelo maior *cluster* # 0, o segundo maior # 1, e assim por diante.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

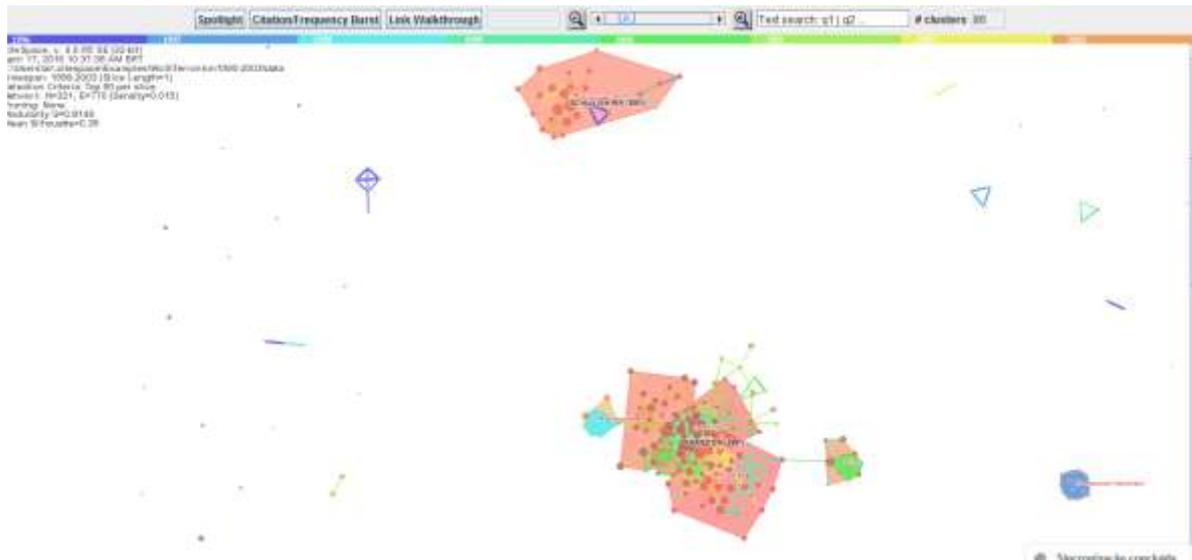


Figura 36. Etiquetas dos *clusters* são geradas e exibidas.

Para tornar mais fácil a verificação de qual *cluster* é o maior, pode-se optar por alterar o tamanho da fonte dos rótulos do uniformizado para proporcional:

Display ► Label Font Size ► Cluster: Uniformed/Proportional

Esta é uma função de alternância. Isso significa que há dois estados. Sua seleção irá alternar entre os dois estados, ou seja, quer utilizando um tamanho de fonte uniformizado ou proporcional.

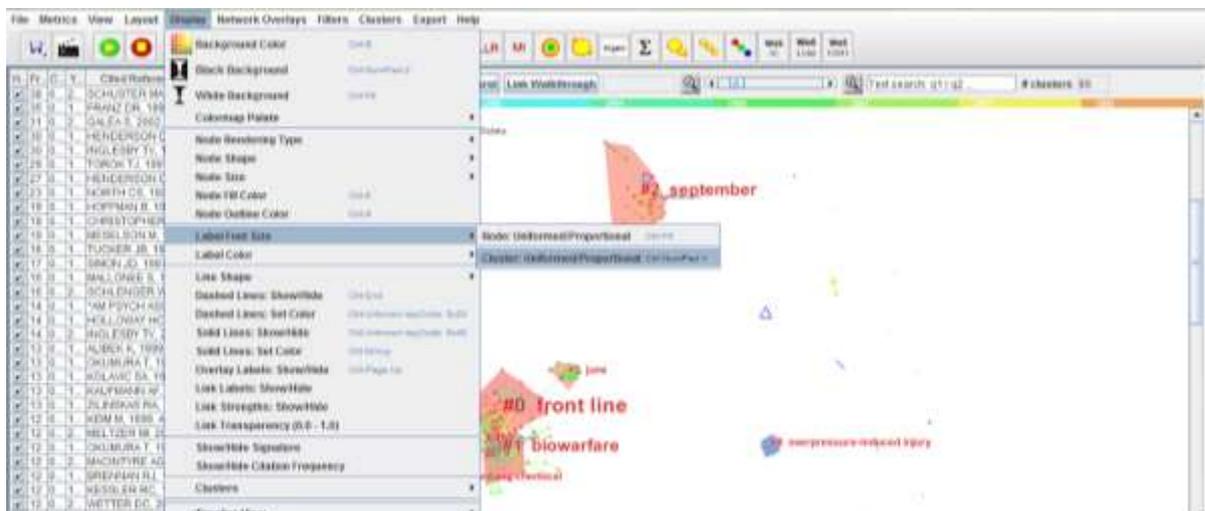


Figura 37. Definir o tamanho da fonte dos rótulos do *cluster* proporcional ao seu tamanho.

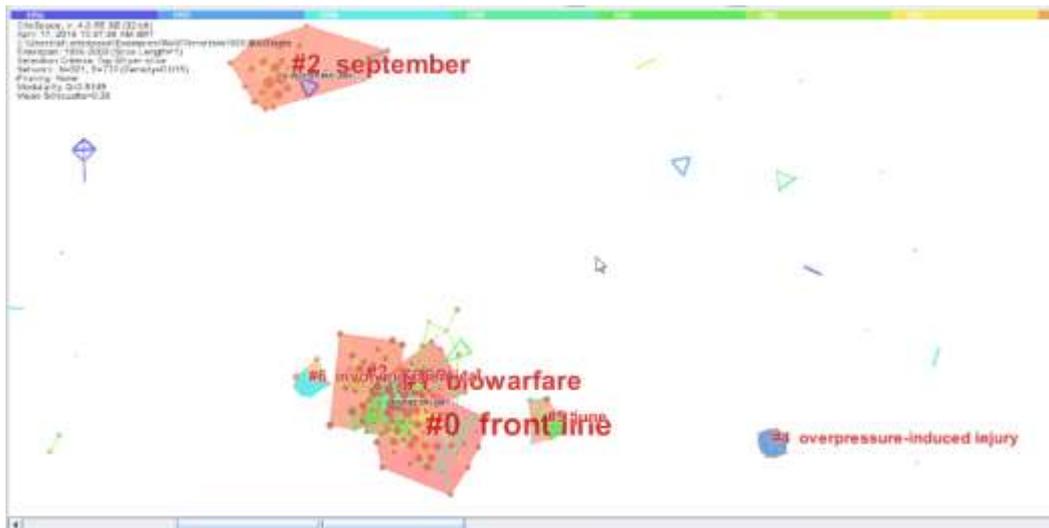


Figura 38. Etiquetas do *Cluster* são proporcionais ao tamanho de um *cluster*.
O maior *cluster* é #0 sobre o terrorismo biológico.

5.1.4 Onde estão as principais áreas de pesquisa com base no *dataset* de entrada?

Esta é uma das questões primárias para que o *CiteSpace* foi projetado para responder. Para responder a essa pergunta, vamos nos concentrar na grande figura da coleção de publicações representadas por seu conjunto de dados. Vamos fazer alguns ajustes com os controles deslizantes no painel de controle à direita para que as informações de nosso interesse sejam exibidas claramente e a informação que é menos relevante para esta questão agora seja temporariamente escondida.

1. Node Size

A este nível, nós realmente não precisamos ver o tamanho de um nó, embora ele forneça informações ricas sobre a história de um nó. Use o controle deslizante nos termos do **Article Labeling ► Node Size ► [Slide to 0]** (marcado pelo ponteiro #1 na figura a seguir).

2. Cluster Label Size

O tamanho da fonte dos rótulos de *cluster* é controlado por um controle deslizante com dois controles: um controle do limiar para mostrar ou esconder uma etiqueta com base no tamanho do *cluster* (ou seja, para garantir que os *clusters* maiores sejam sempre rotulados) e o outro controle o tamanho da fonte dos rótulos de *cluster* (marcado pelo ponteiro # 2 na imagem).

3. Transparency of Links

Os *links* detalhados podem ser utilizáveis mais tarde, mas podemos ignorá-los agora usando o controle deslizante de transparência para definir a transparência de todos os *links* ao nível mais baixo, ou seja, invisível. Em retrospectiva, um termo mais preciso seria completamente transparente.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Depois de fazer esses ajustes menores, será fácil de responder à pergunta: Onde estão as principais áreas de pesquisa? Evidentemente, a maior área (*cluster* #0 com o maior número de referências de membros) é o terrorismo biológico. O segundo maior é o estresse pós-traumático (*cluster* #1), ou seja, PTSD (*Post Traumatic Stress*). O terceiro é a lesão ocular (*cluster* #2). A quarta é explosão (*cluster*#3). E existem alguns *clusters* menores. Portanto, agora temos uma ideia geral do que constituía pesquisa sobre terrorismo durante o período de 1996 e 2003. Você pode repetir o processo em um conjunto de dados atual para obter uma grande imagem de atualização.

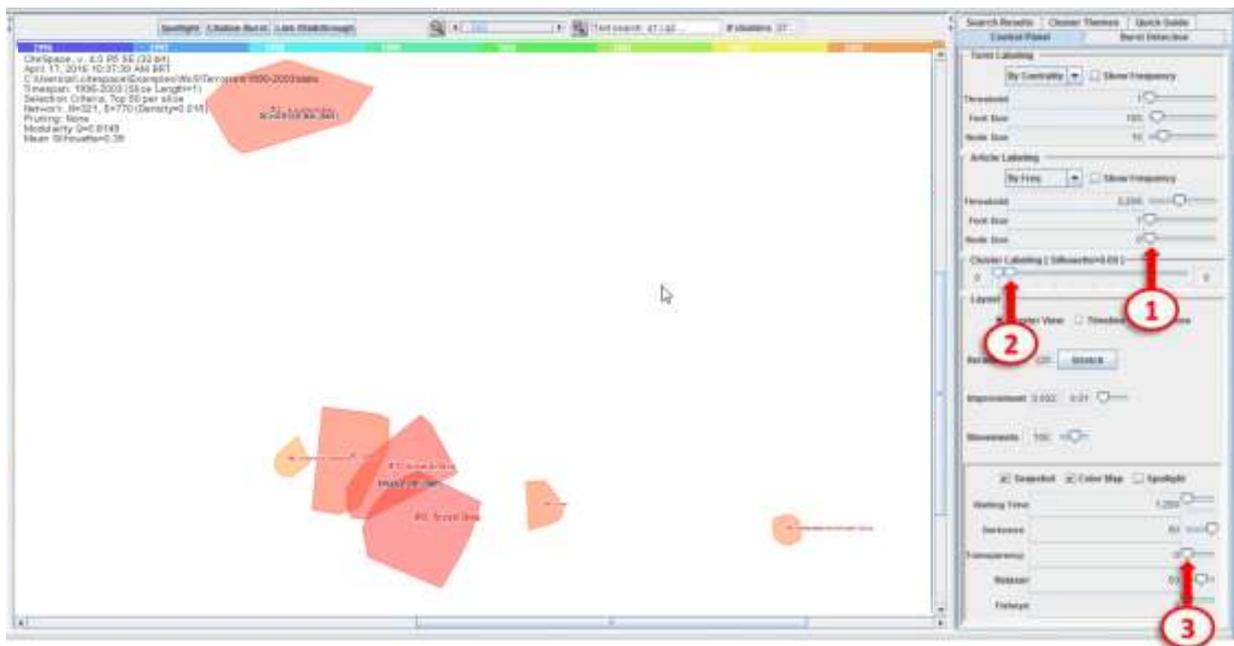


Figura 39. Ajustar a aparência da visualização com algumas barras deslizantes. Ponteiros: 1) controle deslizante do tamanho Node, 2) tamanho de etiqueta do *cluster* e 3) transparência dos *links*.

5.1.5 Como estão conectadas essas grandes áreas?

Para responder a essa pergunta, precisamos trazer de volta as linhas que ligam os nós. Ajuste o controle deslizante de transparência (*transparency*) para tornar as linhas visíveis.

Um indicador útil de como diferentes *clusters* são ligados, é um tipo de nó que têm altas pontuações de centralidade de intermediação (*betweenness centrality scores*). Em *CiteSpace*, pontuações de centralidade de intermediação são normalizadas para a unidade de intervalo de [0, 1]. Um nó de alta centralidade de intermediação é geralmente um que liga dois ou mais grandes grupos de nós com o próprio nó no meio, daí o termo intermediação. O *CiteSpace* destaca nós com alta centralidade de intermediação com realces roxos. A espessura de um realce em roxo indica o quão forte é a sua centralidade de intermediação. Quanto mais espessa mais forte. Ocasionalmente, um nó com alta centralidade de intermediação pode aparecer no

centro de um componente da rede, porém, o nosso interesse são os nós que estão de fato fazendo conexão entre os *clusters*.

Para tornar visíveis os anéis roxos, mude o nó de modo renderização para anéis de árvores , no primeiro ícone mostrado na figura 40, ou seja, anéis concêntricos de citação representam quantas citações foram feitas para o nó nos anos pesquisados. Lembre-se que cores representam quando as citações foram realmente feitas.



Figura 40. Ícones de controles de renderização de nó.

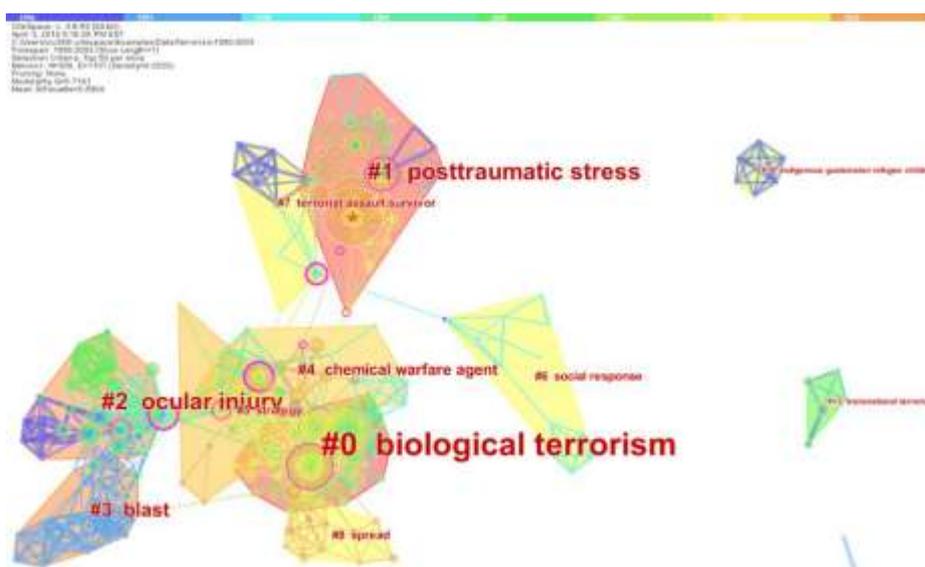


Figura 41. Nós com anéis roxos são importantes na ligação de diferentes *clusters*.
Figura obtida com dados antigos, não correspondem aos dados do exemplo da versão de 29/12/2015.

5.1.6 Onde estão as áreas mais ativas?

5.1.6.1 Detecção de explosão (*burst*)

A explosão de citações é um indicador de uma área de pesquisa mais ativa. Explosão de citação é uma detecção de uma explosão de evento, que pode durar por vários anos, bem como um único ano. Uma explosão de citação fornece evidências de que uma publicação específica está associada com um aumento de citações. Em outras palavras, a publicação, evidentemente, tem atraído um extraordinário grau de atenção de sua comunidade científica. Além disso, se um *cluster* contém numerosos nós com fortes explosões de citação, então, o *cluster* como um todo capta uma área ativa de pesquisa, ou uma tendência emergente.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

A detecção de explosão em *CiteSpace* é baseada no algoritmo do Kleinberg (Kleinberg, 2002).

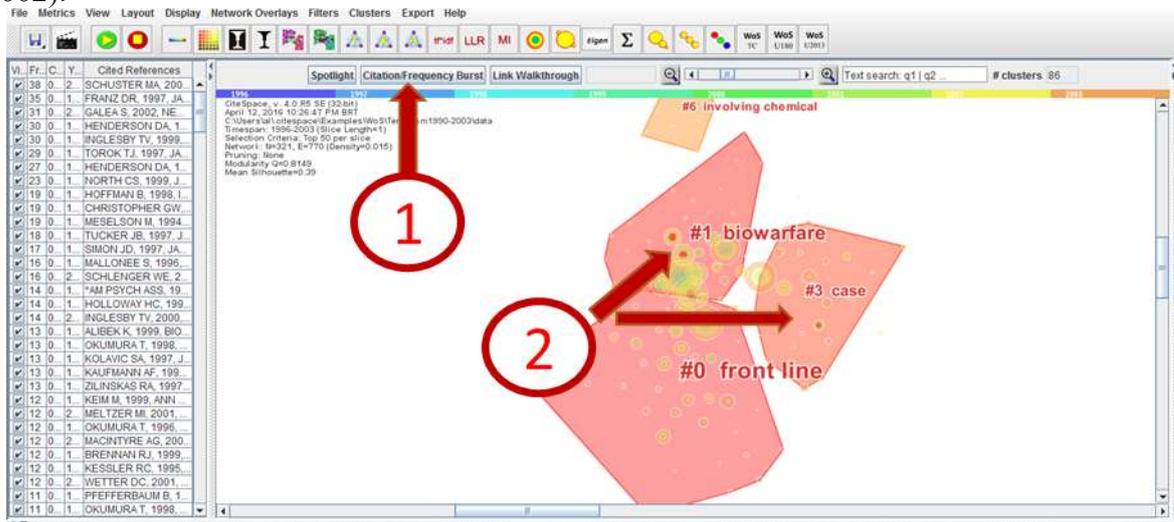


Figura 42. Explosões de citação são indicadores das áreas mais ativas.

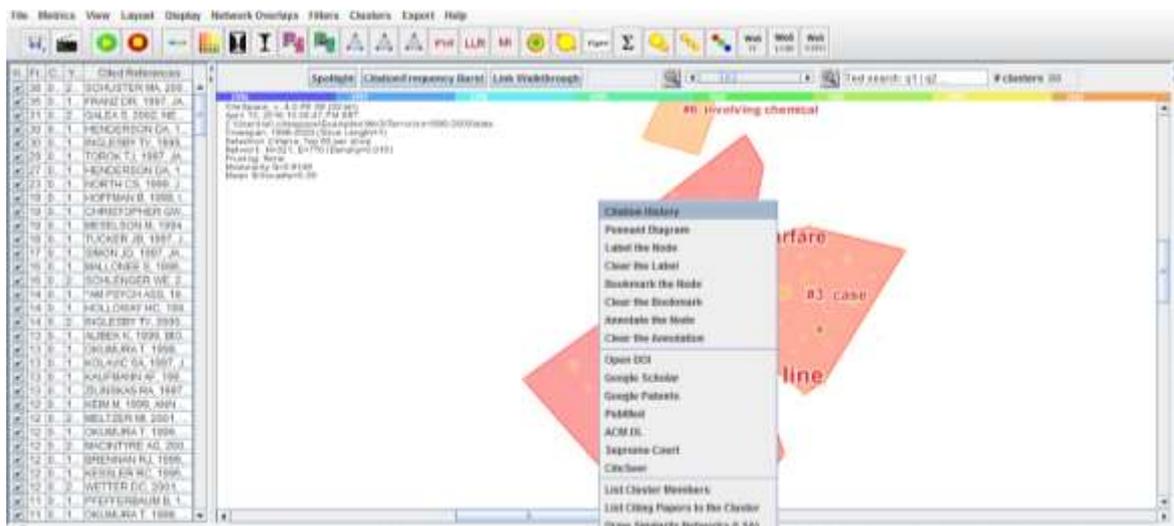


Figura 43. Clique com o botão direito do mouse no nó de interesse e escolha a *Citation History* do nó.

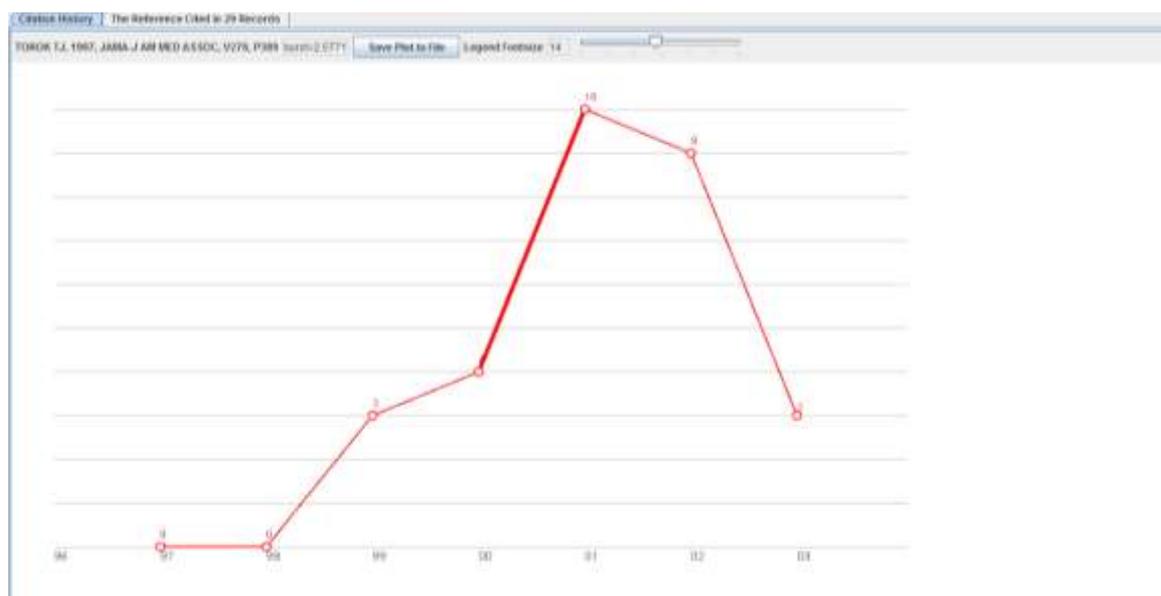


Figura 44. Esta é a *Citation History* de um artigo com explosão de citação.

Usando **View ► Citation Burst History** pode-se gerar uma lista resumida dos artigos que estão associados com explosões de citação. Esta visualização mostra quais referências têm explosão de citações mais fortes e em quais períodos de tempo que ocorreram. Por exemplo, na lista, podemos dizer que KATZ E. e ANN SURG, (1989) tem as explosões de citação mais fortes (Strength = 3.4633) entre os artigos publicados desde os ataques terroristas em 2001, no período de 1996 a 1997. Também é interessante observar que o CARTER A. e FOREIGN AFF., (1998) tem a segunda explosão de citação mais forte (Strength = 3.8914) no período de 1999 e 2001.

Top 9 References with Strongest Citation Bursts

References	Year	Strength	Begin	End	1996 - 2003
ENDERS W, 1993, AM POLIT SCI REV, V87, P829	1993	2.3817	1996	2000	██████████
KATZ E, 1989, ANN SURG, V209, P484	1989	3.4633	1996	1997	██████████
MALLONEE S, 1996, JAMA-J AM MED ASSOC, V276, P382	1996	2.7259	1997	2000	██████████
SHALEV AY, 1992, J NERV MENT DIS, V180, P505	1992	2.3202	1997	2000	██████████
SIDELL FR, 1996, ANN EMERG MED, V28, P223	1996	2.2151	1998	1999	██████████
CARTER A, 1998, FOREIGN AFF, V77, P80	1998	2.8414	1999	2001	██████████
MESELSON M, 1994, SCIENCE, V266, P1202	1994	2.5491	1999	2001	██████████
SIMON JD, 1997, JAMA-J AM MED ASSOC, V278, P428	1997	2.5364	1999	2001	██████████
TOROK TJ, 1997, JAMA-J AM MED ASSOC, V278, P389	1997	2.5771	2000	2001	██████████

↙

Figura 45. Uma lista resumida de referências com explosões de citação.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Detecção de explosão e visualização podem ser aplicadas a outros tipos de nós. Para o tipo de nó do autor, serão mostrados os autores que rapidamente aumentaram o número de publicações. Da mesma forma, as instituições identificarão as universidades que estão particularmente ativas nas áreas de investigação. Para palavras-chave, mostrará tópicos de crescimento rápido.

O procedimento geral é o mesmo para diferentes tipos de nós. Aqui ilustramos o procedimento com um exemplo de detecção de explosão de palavras-chave em publicações da Drexel University, no intervalo 2000 e 2014.

1. Selecione o tipo de nó: ***Keyword***
2. Gere uma rede: **2000-2014; Slicelenght: 1; Top N = 100; GO**
o (N = 392, E = 3033)
3. Execute a função de detecção de explosão: ***Citation/Frequency Burst***
4. Visualize as entidades, ou seja, nós, com explosões: **View> Citation Burst History**

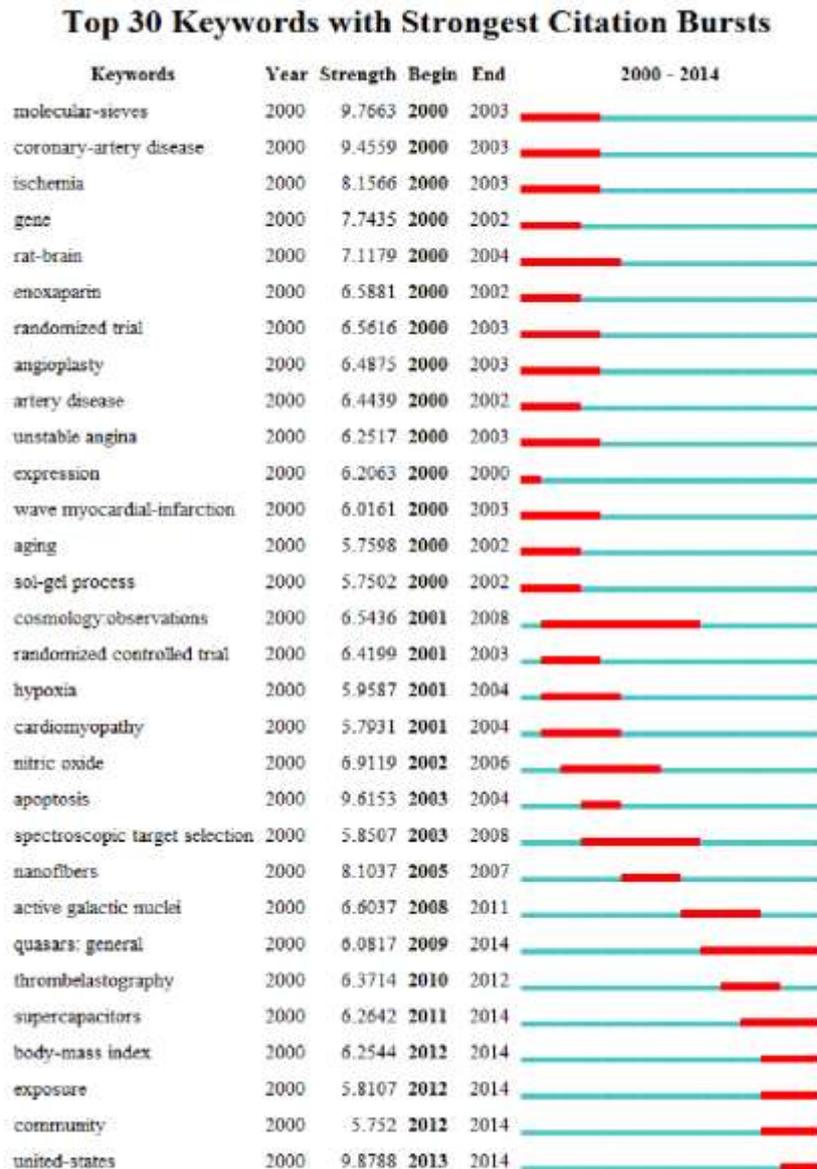


Figura 46. Uma visualização do histórico das explosões de palavras-chave em publicações da Drexel University (2000-2014). Por exemplo, cosmologia: observações em Astronomia tem o mais longo período de explosão de 2001 até 2008.

5.1.6.2 Detecção de explosão – Controlos Adicionais

Se o número de itens de explosão é muito excessivo ou reduzido, podem-se ajustar vários parâmetros para o algoritmo de detecção de explosão.

Ilustraremos os passos para detectar explosões de autores, ou seja, autores que publicaram a um ritmo muito rápido, com o conjunto de dados sobre “terrorismo”, que acompanha o pacote do *CiteSpace*. O procedimento é o mesmo para outros tipos de nó.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Primeiro, selecione *Author* no painel *Node Types* e desmarque outros tipos de nó.

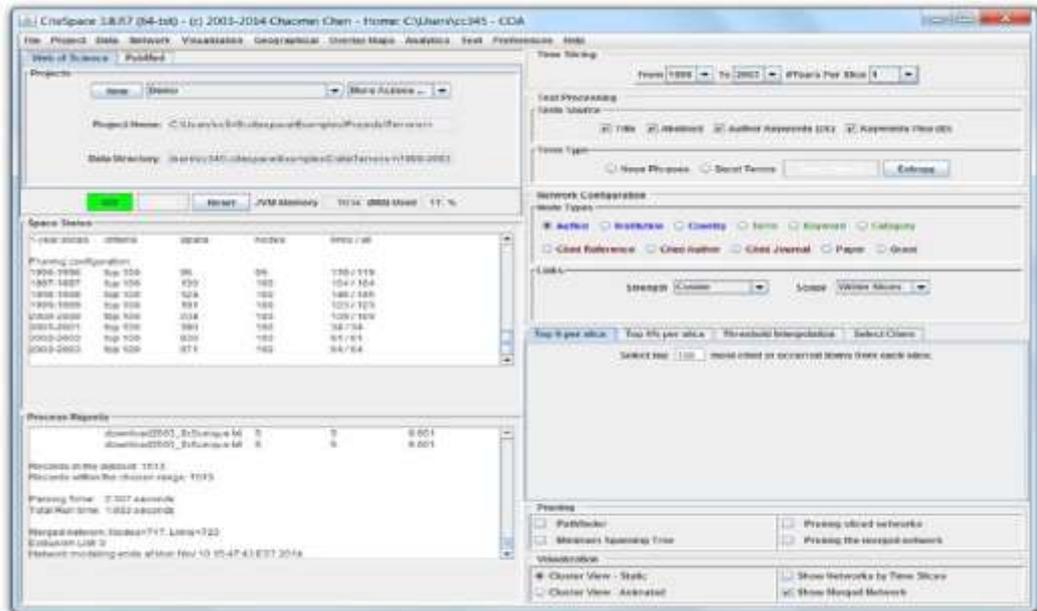


Figura 47. Selecione *Author* como *Node Type* e desmarque outros tipos.

Espreze até que a janela de visualização apareça. O processo de *layout* pode continuar a rodar por um tempo. Considerando que nosso foco está na detecção de explosão, pode-se parar o processo de *layout* a qualquer momento que desejar clicando no botão **Stop** (que é o único com um quadrado amarelo sobre um fundo vermelho). Em seguida, clique no botão *Citation Burst*.

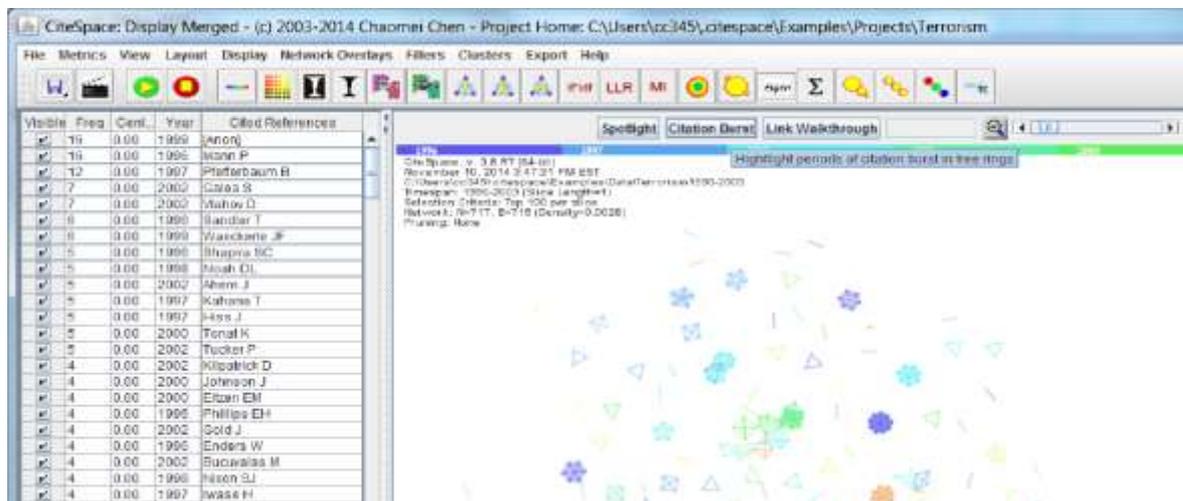


Figura 48. Uma vez que o processo de *layout* é completado, ou seja, quando se vê o fundo tornar-se branco, clique no botão *Citation Burst*. Também, é possível forçar a parada do processo de *layout* clicando no botão *Stop*.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Para visualizar a distribuição dos itens com explosões, selecione o menu *View* e então o item de menu *Citation Burst History*. Embora seja chamado de *Citation Burst History*, abrange os padrões de explosão de outros tipos de eventos associados em conformidade com diferentes tipos de nós. Por exemplo, se o tipo de nó é *Author*, então uma explosão será da frequência com que o autor publica artigos. Se o tipo de nó é um *Cited Reference*, então uma explosão será sobre a frequência com que a referência é citada.



Figura 49. Selecione a *Citation Burst History* abaixo do menu *View* para visualizar a duração e intensidade dos itens de explosão.

Uma caixa de diálogo irá solicitar para que seja digitado o tipo de nó. Neste exemplo, foi introduzida a palavra “*Authors*” e esta palavra será usada como cabeçalho na janela seguinte.

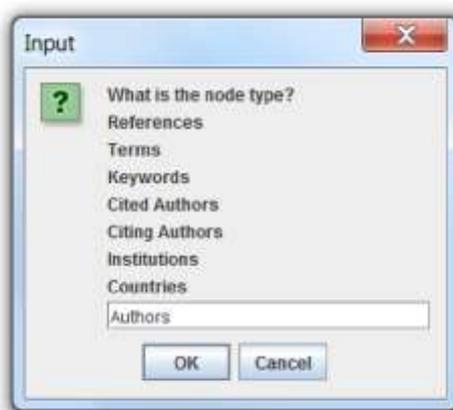


Figura 50. Digite a palavra “*Authors*” para gerar um título apropriado da tela.

Em seguida, se terá a oportunidade de selecionar o número de itens a serem incluídos na janela de visualização.

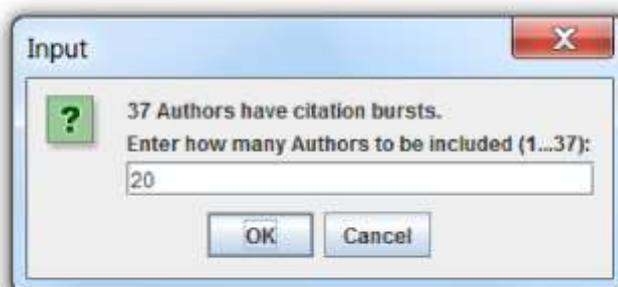


Figura 51. Selecione o número de itens a serem incluídos na visualização. O número deve estar dentro da faixa mostrada na caixa de diálogo.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

A figura 52 mostra a janela de visualização resultante. Pode-se classificar a lista de duas formas: pela força de uma explosão ou pelo início do ano da explosão.

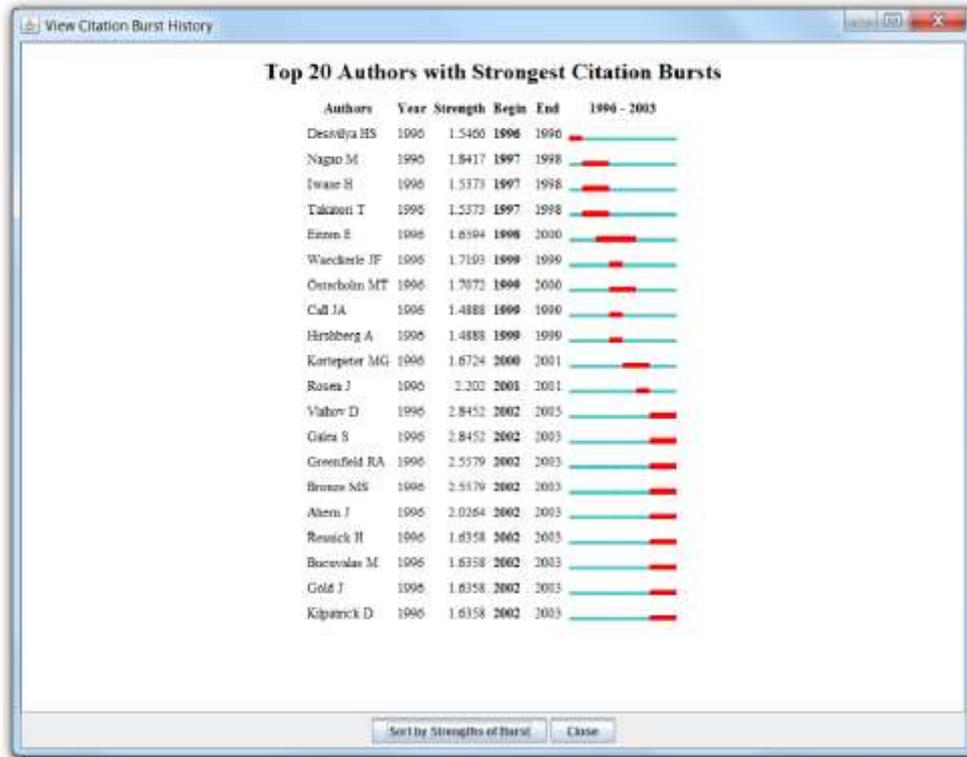


Figura 52. Top 20 dos *Authors* com explosões mais fortes, isto é, os que eram mais ativos nos artigos publicados de acordo com o conjunto de dados (*dataset*).

Se a detecção de explosão inicial identifica apenas um pequeno número de itens, se podem ajustar os parâmetros fornecidos no painel de *Burst Detection* para aumentar ou diminuir o número total de itens de explosão. Por exemplo, a seguinte combinação de parâmetros aumenta o número de autores com padrões de explosão de 37 a 85.

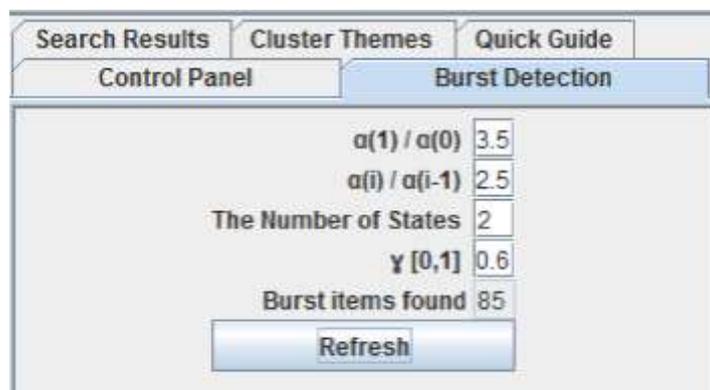


Figura 53. Quatro parâmetros adicionais no painel de *Burst Detection* podem ser ajustados para controlar o algoritmo de detecção de explosão.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Depois de clicar no botão *Refresh*, o *CiteSpace* irá recalculer a explosão de todos os itens. Então se podem usar as etapas descritas acima para exibir os novos resultados.

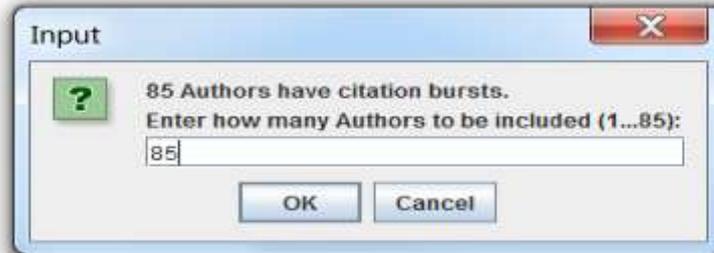


Figura 54. A detecção de explosão recalculada encontrou 85 autores, um aumento substancial dos 37 autores encontrados pela configuração padrão.

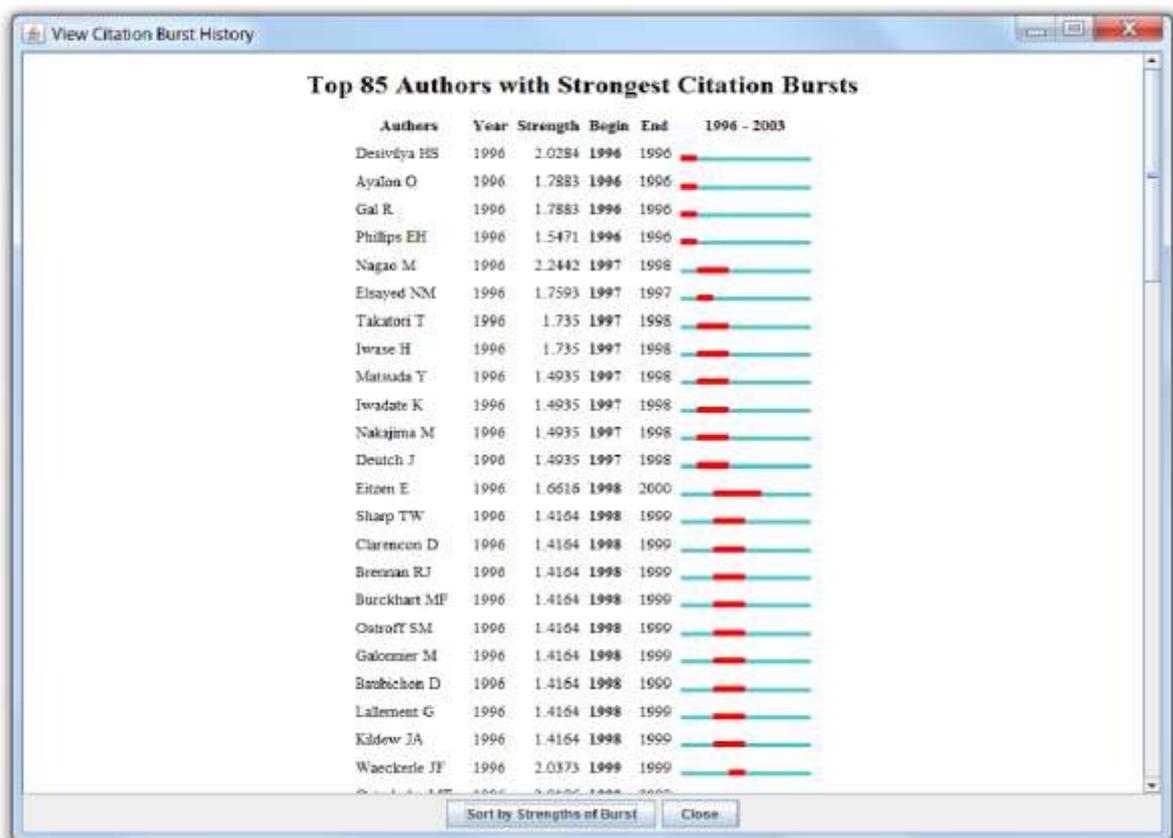


Figura 55. A lista mostra os novos resultados de um processo de detecção de explosão reconfigurado.

5.1.7 O que é importante sobre cada área? Qual /onde estão os artigos-chave para uma determinada área?

Os *Clusters* rotulados podem nos dizer o contexto em que eles são mais citados porque os termos dos rótulos são extraídos de títulos, palavras-chave ou resumos dos artigos que citam o trabalho.

Para explorar esses *clusters* com mais detalhes, deve-se usar o *Cluster Explorer*:

Clusters ► Cluster Explorer

A aparência inicial do *Cluster Explorer* mostra quatro janelas: 1) *Clusters*, 2) *Citing Articles*, 3) *Cited References*, e 4) *Representative Sentences*. Janelas 2-3 estão em branco até que seja selecionado um *cluster* na janela de *Clusters*, marcando a caixa de seleção na frente de cada linha de informações do *cluster*.

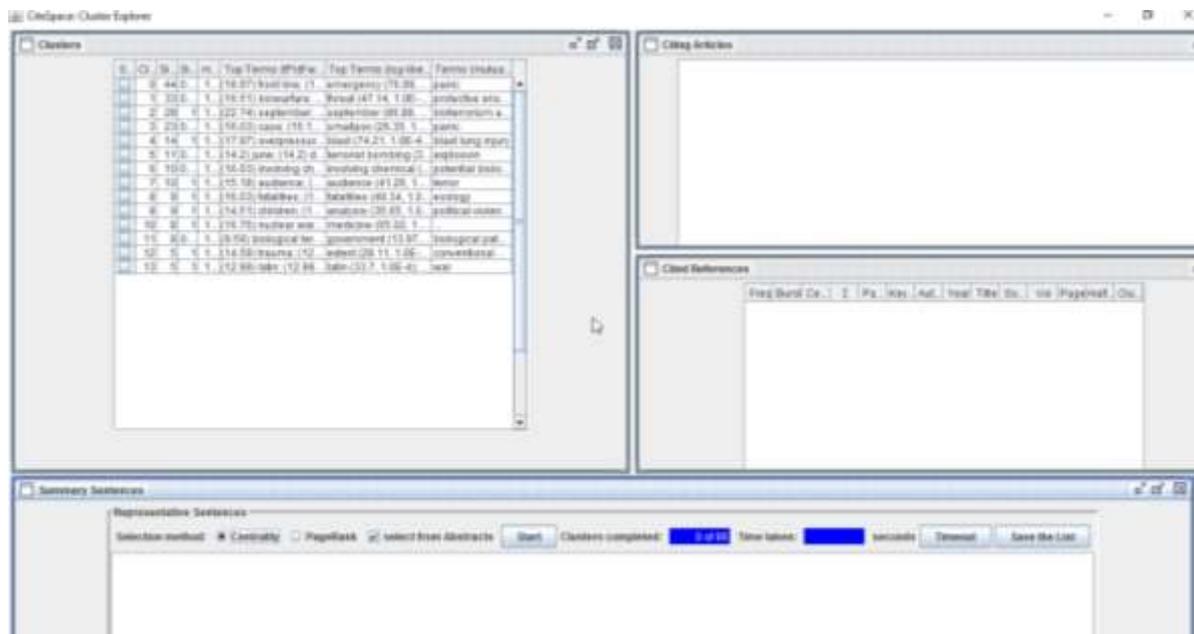


Figura 56. A aparência inicial do *Cluster Explorer*.

A figura 57 mostra uma tela depois que o *Cluster* # 0 foi selecionado na caixa de seleção que está na primeira coluna da janela *Clusters*. Como se pode ver, após a seleção do *cluster* são exibidas as informações na janela *Citing Articles* e *Cited References*. Na janela *Citing Articles*, cada entrada é um artigo que cita os membros do *cluster*. O número na frente de cada entrada mostra a parte das referências citadas por este artigo em particular, de todas as referências no total. Por exemplo, Bak, SJ (2000) tem uma cobertura de 0,09, ou seja, 9% do total de 52 referências neste *cluster* (pode-se encontrar a quantidade de referências que fazem parte do *cluster* “44”, na terceira coluna denominada *size* da janela *Clusters*).

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>

A frase **emergency** foi destacada em amarelo na janela *Citing Articles*. Nota-se que a frase é também o rótulo deste *cluster* na visualização. Além disso, a frase também aparece na sétima coluna da janela *Clusters - Top Terms (log-likelihood ratio)*. Para obter detalhes técnicos, consulte (C. Chen et al., 2010).

A janela *Cited References* mostra as referências que são integrantes deste *cluster*. Cada referência é listada com o número de citações, explosões, se for o caso, sua pontuação de centralidade, juntamente com o nome do primeiro autor, ano de publicação, fonte (isto é, periódico ou conferência), número do volume e número da página.

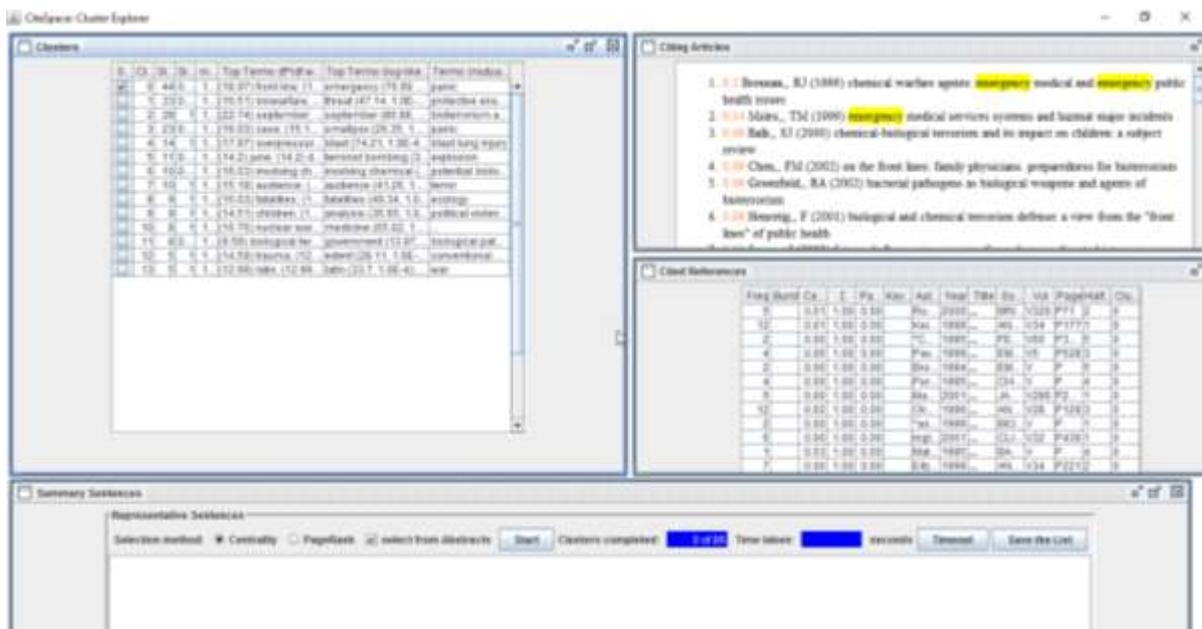


Figura 57. *Cluster Explorer*: *Cluster # 0* está selecionado na caixa de seleção.

Na janela *Summary Sentences*, se clicar no botão *Start*, o *CiteSpace* irá extrair as frases mais representativas dos resumos dos artigos que citam cada *cluster*. A sentença é considerada representativa se ele é ou uma frase com alto grau de centralidade ou uma frase com uma alta pontuação de *PageRank*.

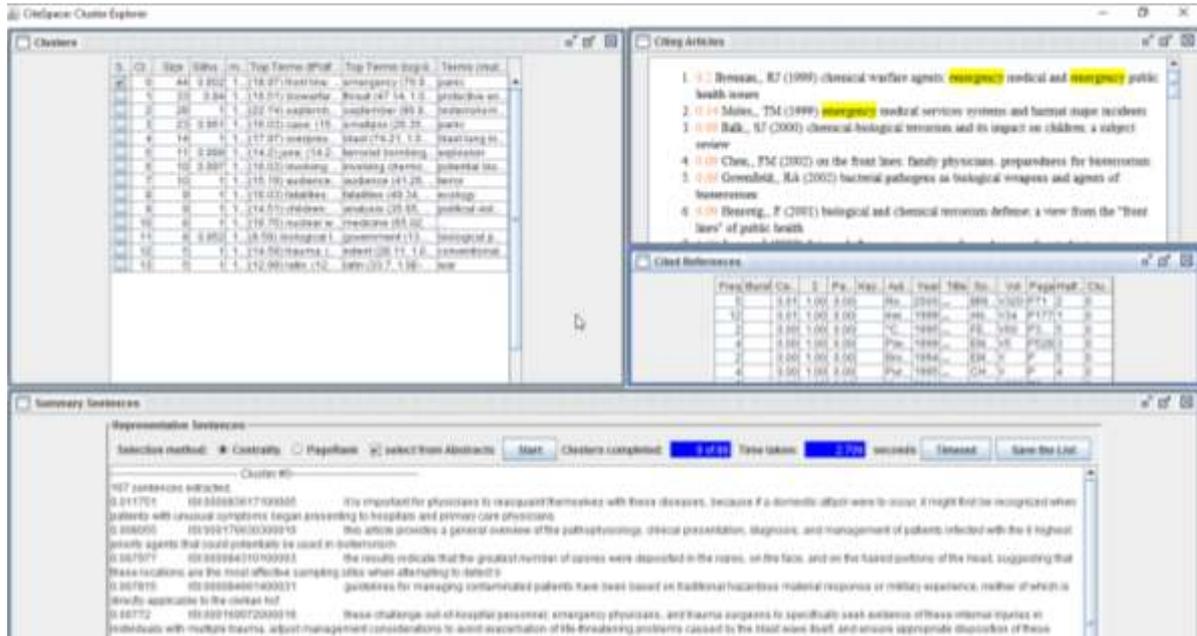


Figura 58. Frases representativas são exibidas ao clicar no botão *Start* na janela de *Summary Sentences*.

5.1.8 Visão da Linha do Tempo

Pode-se alternar para a exibição da linha de tempo da rede, escolhendo o botão de opção *Timeline* no painel *Layout* do lado direito (como apontado pela seta vermelha na figura 59). Em uma visão de linha de tempo, cada *cluster* é organizado em uma linha de tempo horizontal. A direção de tempo aponta para a direita.

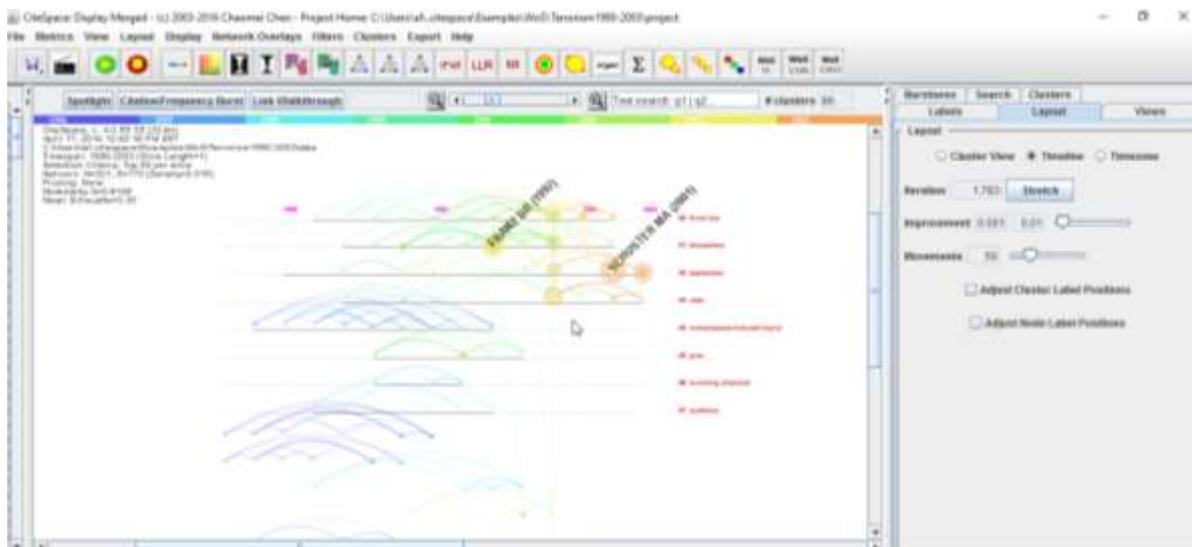


Figura 59. A visão da linha de tempo (*timeline*) da rede.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Foram vistos alguns dos movimentos básicos. O *CiteSpace* tem muitas outras características. Vamos introduzir outros recursos em níveis mais avançados.

5.2 Tente com seu próprio conjunto de dados (dataset)

5.2.1 Coleta de dados

5.2.1.1 Como construir meus próprios dados utilizando a Web of Science

A principal fonte de dados para o *CiteSpace* é a *Web of Science* (WoS).

Mais importante ainda, o *dataset* deve incluir as referências citadas, a fim de maximizar o potencial do *CiteSpace*.

A *Web of Science* tem várias maneiras de procurar registros bibliográficos. O mais básico é chamado de pesquisa básica, que inclui tópico, autor, e vários outros campos pesquisáveis. O exemplo a seguir mostra uma pesquisa no campo tópico (*topic*) usando o termo "*CiteSpace*" entre o período de tempo de 2004 e 2014.

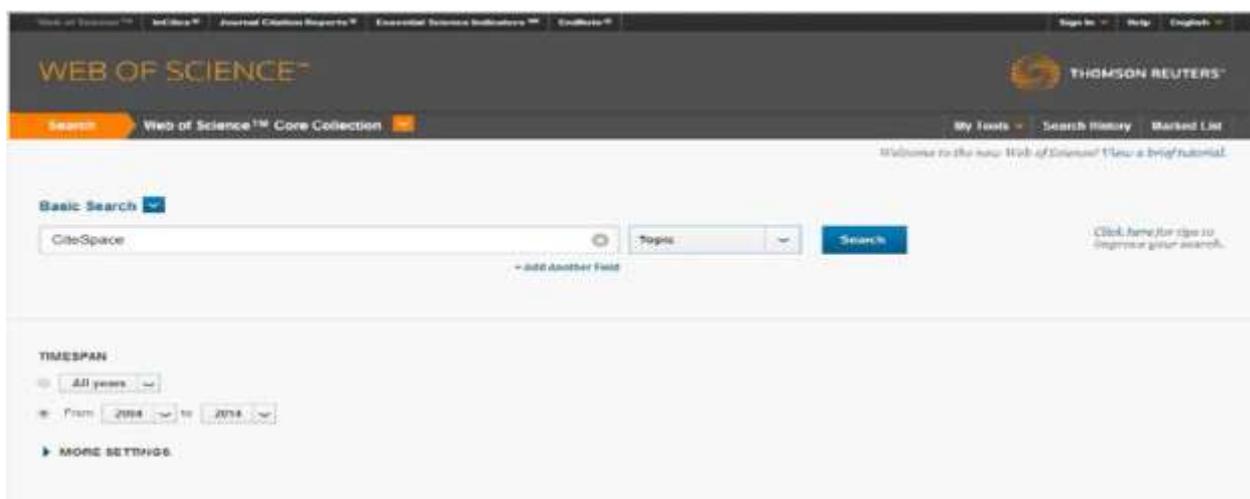


Figura 60. Pesquisa por tópicos na *Web of Science*.

A busca de tópicos encontrou 71 resultados. Estes são exibidos inicialmente na ordem cronológica da data de publicação do mais recente para o mais antigo. Pode-se alternar para uma ordem diferente, por exemplo, pelo número de citações, do mais alto ao mais baixo, para que se possa rapidamente restringir a um pequeno subconjunto dos registros altamente citados.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 61. Classificar os resultados por número de citações - maior para o menor.

Será visualizada a lista de resultados ordenados por número de citações - maior para o menor. O registro com o maior número de citações é o artigo 2006 JASIST sobre o *CiteSpace II*, com 313 citações. A pesquisa de tópicos encontrou 71 registros. Pode-se fazer o *download* destes 71 registros, no entanto, não seria representativo. Se clicarem no *link Criar relatório de citações*, poderá expandir os 71 registros para cerca de 498 registros que citaram o conjunto de 71 registros. Referimo-nos a esta forma de obter mais registros potencialmente relevantes com a expansão da citação. Uma vez que a única coisa que sabemos é que cada registro no conjunto expandido pelo menos cita um dos 71 registros originais, pode vir a ser um registro menos relevante, devido às variações de como os autores citam. Vamos ver se podemos fazer melhor do que encontrar 498 registros relacionados por indexação de citação.

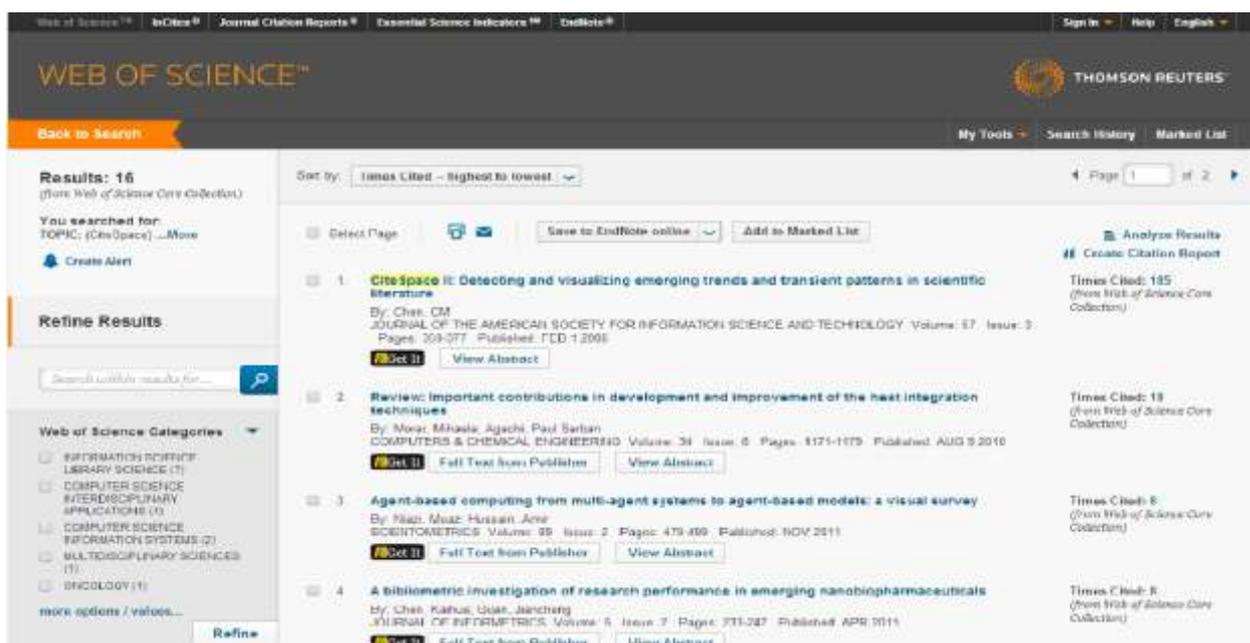


Figura 62. Os resultados agora são ordenados por *Times Cited - the highest to the lowest* (Número de citações – maior para o menor).

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Também, pode-se perceber que o 2004 PNAS e o artigo 2010 JASIST sobre *CiteSpace* não aparecem na lista de resultados, embora eles certamente versam sobre o *CiteSpace* e suas citações iria colocá-los na lista também. Assim, este exemplo mostra que se deve ter cuidado ao utilizar a busca por tópicos na WoS para construir o seu próprio *dataset*.

Na coluna à direita da lista de resultados aparece o número de citações de todas as bases de dados para cada registro, como um *link* 'clícável'. No caso do artigo 2004 PNAS o número de citações é de 151. Se clicarem nele, trará a lista de 151 registros que citaram o artigo 2004 PNAS. O artigo 2006 JASIST deverá estar na lista. Se a lista for ordenada por Número de citações – do maior para menor, então veremos o 2006 JASIST no topo com 313 citações em todas as bases de dados.

Searching for intellectual turning points: Progressive knowledge domain visualization
By: Chen, CM (Chen, CM)

PROCEEDINGS OF THE NATIONAL ACADEMY OF SCIENCES OF THE UNITED STATES OF AMERICA
Volume: 101 Pages: 1203-1210 Supplement: 1
DOI: 10.1073/pnas.0307513100
Published: APR 6 2004
View Journal Information

Abstract
This article introduces a previously unreported method progressively visualizing the evolution of a knowledge domain's citation network. The method first derives a sequence of citation networks from a series of equal length time interval slices. These time-registered networks are merged and visualized in a panoramic view in such a way that intellectually significant articles can be identified based on their visually salient features. The method is applied to a citation study of the superconducting field in theoretical physics. The study focuses on the search of articles that triggered true superconducting revolutions. Visually salient nodes in the panoramic view are identified, and the nature of their intellectual contributions is isolated by tracing networks in the field. The analysis has demonstrated that a search for intellectual turning points can be narrowed down to visually salient nodes in the visualized network. The method permits a promising way to simplify otherwise cognitively demanding tasks in a search for landmarks, pivots, and hubs.

Keywords
KeyWords Plus: AUTHOR COCITATION ANALYSIS; CO-CITATION NETWORKS; DECOMPOSITION; GROWTH

Author information
Beginn Address: Chen, CM (paper author)
• Drexel Univ, Coll Inform Sci & Technol, 3511 Chestnut St, Philadelphia, PA 19104 USA
Addresses:
• [1] Drexel Univ, Coll Inform Sci & Technol, Philadelphia, PA 19104 USA
E-mail Addresses: ckmchen@drexel.edu
• Author Identifiers:

Citation Network
104 Times Cited
34 Cited References
View Related Records
View Citation Map
Create Citation Alert
(Alerts from Web of Science™ Core Collection)

All Times Cited Counts
113 in All Databases
104 in Web of Science Core Collection
23 in BIOSIS Citation Index
8 in Chinese Science Citation Database
0 in Data Citation Index
1 in ScELO Citation Index

Most Recent Citation
Munafin, Ruzmil. Exploring the modeling and simulation knowledge base through partial citation analysis. SCIENTOMETRICS, MAR 2014
View All

Figura 63. O artigo 2004 PNAS é citado 313 vezes, mas o tópico da busca não será capaz de encontrá-lo porque o termo *CiteSpace* não aparece em seu título, resumo ou palavras-chave.

Agora, se for clicado sobre o *link* no número 313 e posteriormente clicar no *link* Relatório de citações na direita, se terá acesso a todos os registros que citam os 313 artigos, ou seja, seria a expansão da citação que queremos.

Back to Search My Tools Search History Marked List

Citing Articles: 66 (from Web of Science Core Collection)

For Searching for intellectual turning points: Progressive knowledge domain visualization

Times Cited Counts
113 in All Databases
104 in Web of Science Core Collection
23 in BIOSIS Citation Index
8 in Chinese Science Citation Database
0 data sets in Data Citation Index
0 publications in Data Citation Index
1 in ScELO Citation Index
View Additional Times Cited Counts

Sort by: Times Cited - highest to lowest Page 1 of 7

Select Page Save to EndNote online Add to Marked List Analyze Results Create Citation Report

1 CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature
By: Chen, CM
JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY Volume: 57 Issue: 3
Pages: 359-377 Published: FEB 1 2006
View Abstract

2 Informetrics at the beginning of the 21st century - A review
By: Bar-Ilan, Judit
JOURNAL OF INFORMETRICS Volume: 2 Issue: 1 Pages: 1-62 Published: 2008
Full Text from Publisher View Abstract

Times Cited: 313 (from Web of Science Core Collection)
Times Cited: 88 (from Web of Science Core Collection)

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

no futuro. Por exemplo, pode-se facilmente ocultar um arquivo do *CiteSpace* adicionando um prefixo aos nomes de alguns arquivos que se deseja que o *CiteSpace* pule.

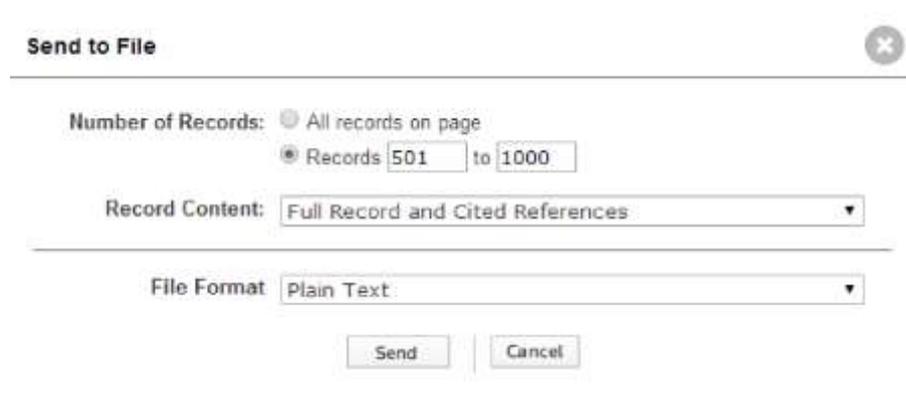


Figura 67. Baixar registros 501-1000 em Texto sem formatação. A *Web of Science* permite baixar o máximo de 500 registros de cada vez. As vezes é necessário repetir o passo várias vezes.

5.2.2 Trabalhando com um projeto no CiteSpace

Um projeto *CiteSpace* é projetado para facilitar sua análise. Cada projeto está associado a um conjunto de dados (*dataset*). Pode-se analisar o *dataset* de muitas maneiras, selecionando uma variedade de parâmetros e propriedades do projeto. O *CiteSpace* gera vários tipos de arquivos intermediários que é possível inspecioná-los em detalhe se desejar. Pode-se tratar com a maioria desses arquivos intermediários diretamente.

5.2.2.1 Criar um projeto CiteSpace

É necessário criar duas pastas separadas para um novo projeto. Uma pasta contém arquivos de dados que foram baixados. Referimo-nos a ela como a pasta de dados. A outra pasta é a pasta do projeto, que será usada para armazenar vários arquivos intermediários.

5.2.2.2 Editar um projeto existente

Pode-se editar as propriedades de um projeto existente. Para escolher esta função, abra o menu que mostra "*More Actions*" ao lado do projeto atual.

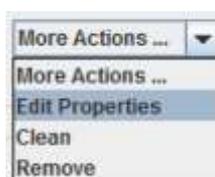


Figura 68. Editar propriedade de um projeto existente.

Podem-se editar várias propriedades de um projeto existente com base em suas necessidades.

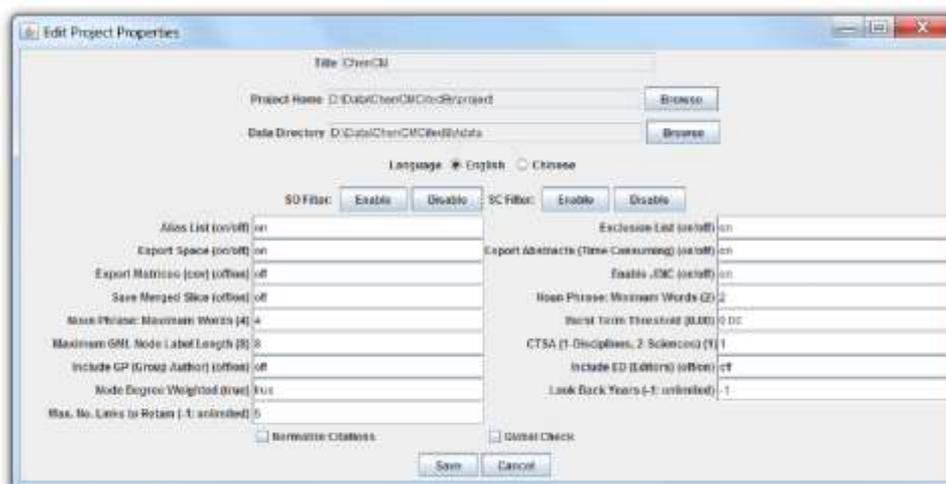


Figura 69. Propriedades de um projeto existente.

Se quiser manter os registros de um conjunto específico de periódicos em seu conjunto de dados, pode-se ativar a função *SO Filter* (clicando no botão *Enable*). Primeiro precisa-se criar uma lista dos nomes dos periódicos cujos registros se deseja manter e salvar a lista em um arquivo ASCII conforme as instruções abaixo.



Figura 70. Instruções sobre como criar um arquivo *SO Filter*.

Podem-se filtrar registros semelhantes com base em seus campos *SC*, ou seja, as suas categorias de assunto.

Alias List: on/off

Esta propriedade é usada para ativar ou desativar o recurso de mesclar diferentes variantes da mesma entidade em um único nó.

ExclusionList: on/off

Esta propriedade é usada para ativar ou desativar o *CiteSpace* para excluir uma lista de itens para aparecer nas visualizações.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Look Back Years

Esta propriedade controla o intervalo máximo de uma citação, considerando a diferença entre a data de publicação do artigo que está citando e a data de publicação do artigo citado. Defina esta propriedade como -1 se você não quiser qualquer limite. Por exemplo, um valor de 5 nesta propriedade significa que as citações referenciadas por mais de 5 anos serão ignoradas.

Esta propriedade é um método simples de redução de *link*.

Max. No. Links to Retain

Esta propriedade controla o número máximo de *links* para manter em cada nó na rede. Defina esta propriedade como -1 se não quiser qualquer limite.

Por exemplo, um valor de 5 nesta propriedade significa que até 5 *links* mais fortes de ligação a um nó será permitido. Se o nó tem mais de 5 conexões vizinhas, então eles serão eliminados, ou seja ignorados.

Esta propriedade é um método simples de redução de *link*.

5.2.2.3 *Limpar um Projeto*

Esta função tentará excluir arquivos de dados intermediários, por exemplo, arquivos de extração de palavras-chave, arquivos de gráfico no formato graphml, arquivos de *clusters*, e arquivos com a palavra *citespace* como prefixo de seus nomes, que registam como você configura seu projeto.

O *CiteSpace* solicitará confirmação para exclusão de alguns tipos de arquivos para evitar de que se apague arquivos que se possa precisar.

5.2.2.4 *Remover um projeto*

Esta função removerá o projeto atual do *CiteSpace*, mas deixará as pastas e arquivos nessas pastas intactas para que se possa restaurá-los por meio da criação de um novo projeto e apontando para as pastas existentes.

5.2.3 Fontes de Dados em Chinês

Uma aplicação de utilidade Java que pode converter dados no formato CSSCI para o formato WoS está disponível para baixar no seguinte *link*:

[http://cluster.ischool.drexel.edu/~cchen/citespace/utilities/CSSCIREC\(new\).jar](http://cluster.ischool.drexel.edu/~cchen/citespace/utilities/CSSCIREC(new).jar)

Armazene arquivos de dados baixados da CSSCI em uma pasta. Antes de usar o conversor de formato, certifique-se de que os arquivos de dados de entrada com a codificação ANSI são salvos em arquivos com a codificação UTF-8 (Use qualquer editor de texto e, em seguida,

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Salvar Como para arquivos com a codificação UTF-8). Em seguida, aplique o conversor para a pasta de dados.



Figure 71. Salve o arquivo baixado CSSCI codificado em ANSI para o código UTF-8 antes de usar a conversão de formato.

A fim de utilizar arquivos de dados com codificação de Chinês, use **Preferences ► Chinese Encoding**.

Para mais discussões em chinês, consulte o seguinte *link*:

<http://blog.sciencenet.cn/blog-496649-427780.html>

5.2.4 Como lidar com resultados da pesquisa contendo tópicos irrelevantes

Pode-se perceber que não importa o quão cuidadosamente se formulou a consulta de busca na *Web of Science* ou quaisquer outras fontes, é sempre possível que os seus resultados de pesquisa contenham temas irrelevantes.

Recomendamos que se utilize a seguinte estratégia. Em vez de refinar sua consulta incessantemente, se pega o *dataset* que pode incluir tópicos irrelevantes e deixe o *CiteSpace* diferenciar os vários tópicos. Pode-se determinar se um tópico é de fato irrelevante somente depois que se tiver uma oportunidade de examinar a visualização dos resultados.

Na maioria dos casos, se torna simples detectar tópicos irrelevantes porque eles iriam acabar em um *cluster* isolado. Pode ser que um tópico aparentemente irrelevante torna-se relevante após uma análise mais profunda.

Por exemplo, o *dataset* foi coletado por um tópico de busca 'hacker*', destinado a capturar os tópicos relevantes para os *hackers*, o comportamento de *hacker*, e tópicos associados.

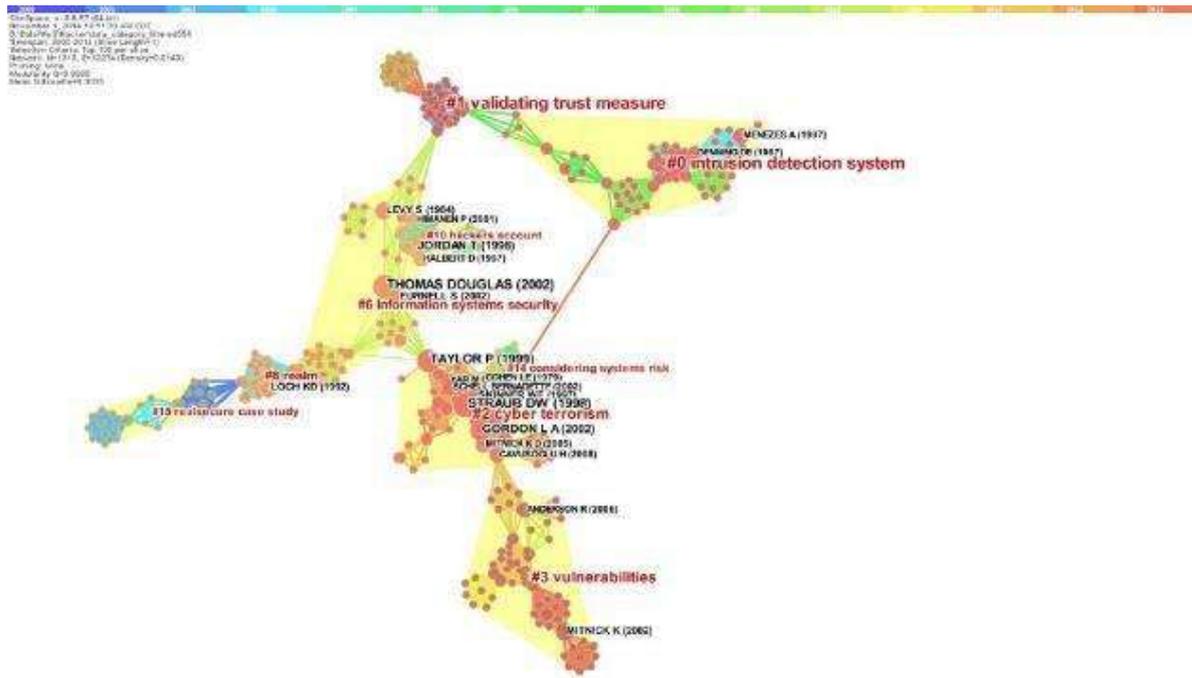


Figura 72. O maior componente conectado numa rede de referências de co-citações.

O maior componente mostrado acima evidentemente contém os temas relevantes, tais como sistema de detecção de intrusão, validação de medida de confiança, *cyber* terrorismo e vulnerabilidades. Por outro lado, a visualização também revela um segundo grande componente interessante.

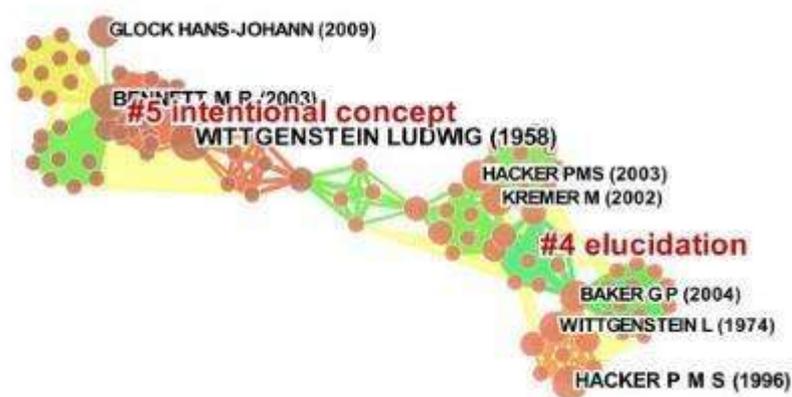


Figura 73. O segundo maior componente da rede de "hacker".

O segundo maior componente não é sobre os tópicos de 'hacker' que desejávamos. Em vez disso, esses itens foram incluídos no resultado da pesquisa original porque o termo *Hacker* é,

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

como se constata, o nome de um prolífico autor - Hacker, P. M. S., um filósofo que publicou uma série de artigos em 1996 e 2003 e um livro sobre o ceticismo, regras e linguagens. Os dois *clusters* contidos no segundo maior componente são essencialmente sobre tópicos irrelevantes para os *hackers* no contexto da segurança de computador.

Este exemplo mostra que é uma boa ideia usar uma pesquisa mais ampla que uma mais restrita e adiar a diferenciação até o processo analítico visual. A relevância de um tópico poderia ser mais fácil para detectar num estágio de processamento posterior.

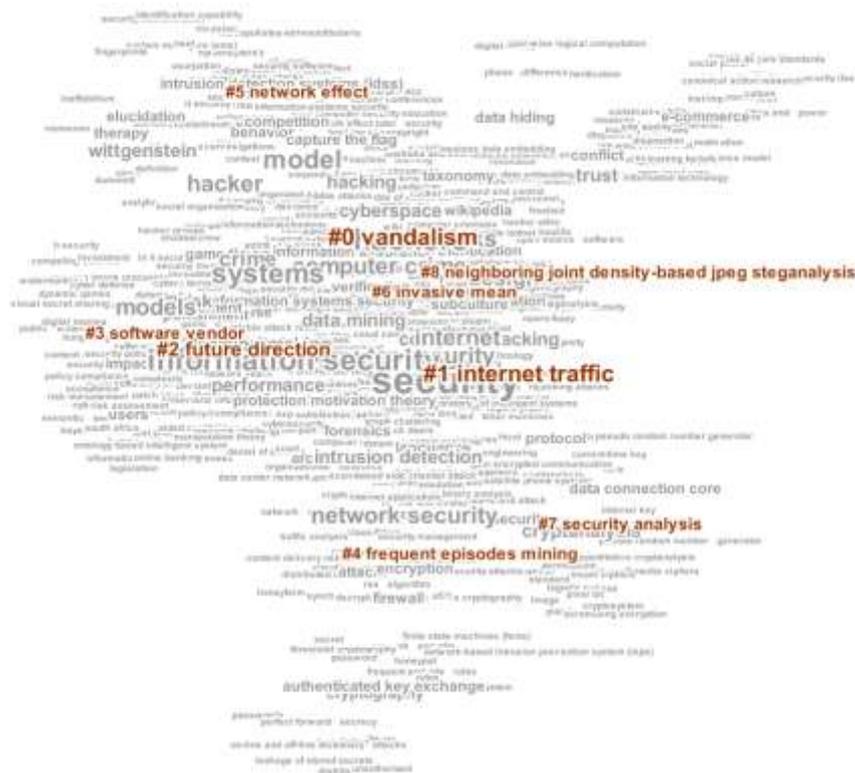


Figura 74. "Clusters" de uma rede de co-ocorrências de palavras-chave. Ajuste *Node size* = 0 e *Transparency* = 0.

6 Configurar a execução do CiteSpace

A chave do *CiteSpace* é o processo de construção da rede. Pode-se configurar o processo através de uma série de parâmetros. Sua configuração irá afetar os resultados do processo.

6.1 *Time Slicing (intervalo de tempo)*

Dado um conjunto de dados (dataset) de registros bibliográficos, precisa-se escolher o período de tempo desejado para que o *CiteSpace* analise, assim, qualquer registro que estiver fora do período será ignorado. Por exemplo, o seu *dataset* contém registros a partir dos anos de 1800 até 2014, pode-se escolher para se concentrar nos últimos 10 anos, ou em um período intermediário. Também, pode-se incluir o *dataset* inteiro se for desejado.



Figura 75. Configurando o *Time Slicing*.

É possível particionar o tempo do intervalo de diversas maneiras, definindo o valor de *#Years Per Slice*. Por *default*, o *CiteSpace* divide o intervalo de tempo em intervalo de 1 ano. Normalmente, se utiliza o intervalo de 1 ano e o número de redes será o mesmo que o número de anos do intervalo. Alternativamente, você poderia usar intervalos de k -anos para que cada intervalo represente os dados de k anos. Também, é possível fazer um único intervalo para que se lide apenas com uma rede.

6.2 *Processamento de texto (Text Processing)*

Cada registro bibliográfico contém quatro campos textuais. Estes campos oferecem texto não-estruturado que pode ser processado e analisado como parte de um processo analítico visual.

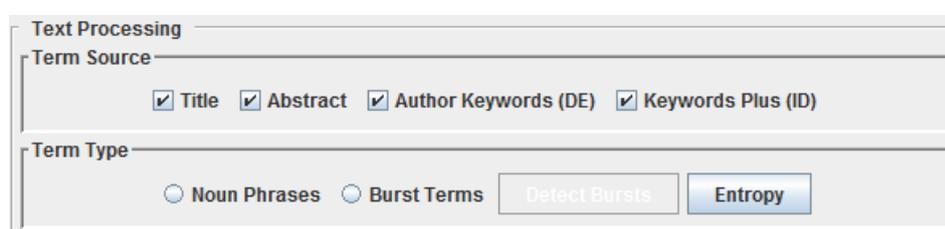


Figura 76. Configurações de *Text Processing*.

Pode-se ignorar o resto desta seção se estiver apenas interessado em criar uma rede de citação de documentos, por exemplo, redes de referências citadas, ou outros tipos de nós que não utiliza termos.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

6.3 Configurando as redes

O *CiteSpace* pode gerar vários tipos de redes. O tipo de nó *default* são as referências citadas. Neste caso, os *links* são de co-citação. As redes são feitas de referências co-citadas.

O *CiteSpace* permite escolher um único tipo de nó ou vários tipos de nós simultâneos. Por exemplo, é possível selecionar *Author*, *Cited References*, e *Category* para formar redes de três tipos de nós e seis tipos de *links*, ou seja, *Author-Author* (colaboração), *Reference-Reference* (co-citação), *Category-Category* (co-ocorrência), *Author-Reference* (referência citada pelo autor), *Author-Category* (autor-publica-em-categoria), e *Category-Reference* (documento-em-categoria-cita-referência).

As redes de co-citação de documentos são construídas sobre os métodos pioneiros de Henry Small (Small, 1973), porém estendido de um intervalo único para análise de redes de intervalo múltiplo, por exemplo, uma rede de série temporal para detectar transições críticas no decorrer do tempo, de forma mais efetiva.

As redes de co-citação de Autor são originadas de (White & Griffith, 1981).



Figura 77. *Network Configuration* (Configuração da rede).

Grande parte do projeto do *CiteSpace* tem sido dedicado a análise de co-citações de documentos, devido às preferências que os padrões de citação de referências fornecem particularmente, revelando *insights* sobre a estrutura e a dinâmica dos paradigmas científicos.

6.3.1 Acoplamento bibliográfico

Pode-se escolher *Paper* como tipo de nó, a similaridade entre os artigos será calculada com base em seus acoplamentos bibliográficos (Kessler, 1963).

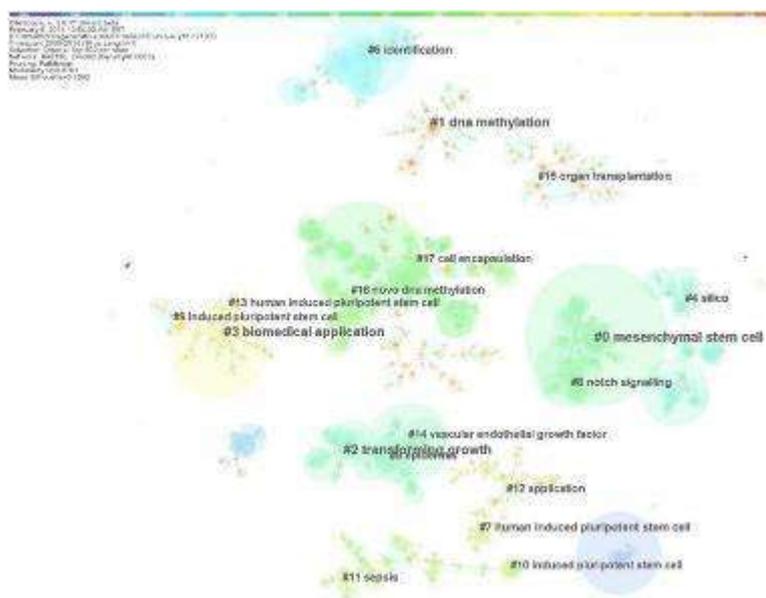


Figura 78. Um exemplo de uma rede de artigos que citam baseada em suas pontuações de acoplamento bibliográfico.

Fonte: Chen, Dubin, and Kim (2014) *Emerging Trends and New Developments in Regenerative Medicine: A Scientometric Update (2000-2014)*. *Expert Opinion On Biological Therapy*. (In Press)

6.4 *Cr terios de Sele o de N *

O *CiteSpace* fornece v rias maneiras para amostragem de registros para formar as redes finais. Esses cr terios s o conhecidos como cr terio de sele o de n .

A forma mais simples e recomendada   a primeira aba **Top N per slice**. Se for inserido um valor de 50, o *CiteSpace* selecionar  os 50 mais citados ou ocorr ncia de itens de cada intervalo, para construir a rede, dependendo do tipo de n  que foi selecionado no passo anterior. Se forem selecionados v rios tipos de n , ent o os n s ser o classificados pelo n mero de vezes que eles aparecem nos registros de cada intervalo.

O segundo m todo de sele o   **Top N% per slice**. Por exemplo, podem-se selecionar os *top* 15% dos itens mais citados ou os itens mais frequentes por intervalo. Tamb m, pode-se selecionar o conjunto de dados (*dataset*) completo, especificando *top* 100% (contanto que se eleve o limite do valor superior alto o suficiente, digamos, 10.000 por intervalo).

O terceiro m todo   **Threshold Interpolation** (interpola o limiar). Ele seleciona ambos os n s e *links*.   complexo. Recomenda-se que se explorem outros cr terios de sele o antes deste.

O quarto deve ser utilizado juntamente com um dos tr s m todos acima – **Select Citers**.   poss vel selecionar os registros com base numa distribui o de cita es. Pode-se especificar um intervalo de distribui o de cita o, por exemplo, um intervalo de [5, max] ir  incluir os

registros que tenham cinco ou mais citações. Após a escolha, é necessário escolher qual dos três métodos de seleção que será utilizado, ou seja, *Top N*, *Top N%*, ou *Threshold Interpolation*.

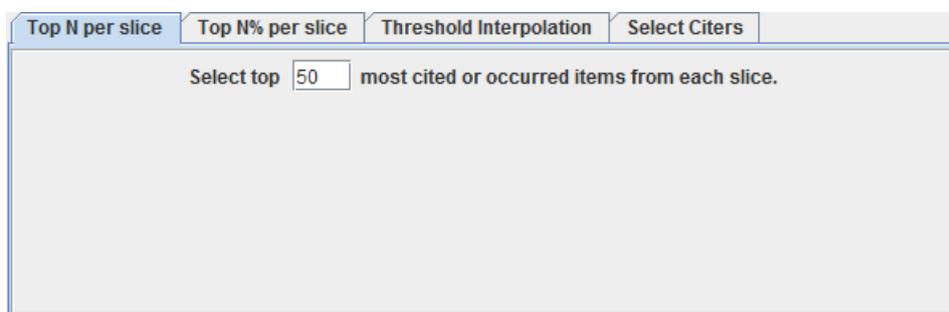


Figura 79. Critérios de seleção de Nó.

6.4.1 Eu tenho a rede certa?

Obviamente, o tamanho de uma rede visualizada influencia a clareza e a complexidade dos padrões que é possível aprender com a visualização. A estrutura de uma rede é determinada pelo número de nós selecionados para cada intervalo de tempo. É pouco provável que se possa saber com antecedência se um Top N de 100 irá gerar uma rede mais desejável que um Top N de 50.

Aqui estão algumas sugestões:

Em primeiro lugar, comece com um **Top N de 50** e gere uma visualização de rede. Em seguida, verifique a modularidade da rede, o número de *clusters*, e as pontuações da silhueta média. Não se aprenderá muito da rede se houver apenas um par de *clusters*. Não haverá um panorama mais amplo, se existir centenas de *clusters*. Um bom intervalo de números de *clusters* seria entre 7~10 *clusters* principais com 10 ou mais membros e cada um dos *clusters* com alto valores de silhueta (p. ex. > 0,70).

Em seguida, é possível experimentar um **Top N de 100** para cada intervalo. Se o seu computador é suficientemente potente, pode-se, sem dúvida, tentar um **Top N de 1.000** por intervalo ou até mesmo superiores.

Deve-se iniciar o processo a partir de uma rede pequena (embora se for incluídos muitos intervalos, até mesmo um Top N de 50, pode se acumular em uma rede de grande porte), e depois, com base na sua avaliação inicial da rede, ampliar a rede.

Por fim, note-se que a rede de maior dimensão não é necessariamente a mais informativa. Primeiro estabeleça com clareza as perguntas que se deseja responder.

6.5 *Pruning (Poda), ou Link Reduction (Redução de Link)*

As redes bibliográficas podem ser muito densas com muitos *links*. O processo para remover sistematicamente os *links* excessivos é chamado *network pruning* (poda da rede) ou *link reduction* (redução de *link*).

O *CiteSpace* disponibiliza duas formas para esta finalidade: *Pathfinder* e *Minimum Spanning Tree*. A comparação dos prós e contras dos dois métodos é detalhada em uma publicação de 2003 (C. Chen & Morris, 2003). Em suma, *Pathfinder* é teoricamente a melhor escolha, mas ela tem um custo superior.

Recomenda-se que se inicie com as redes sem qualquer poda, porque às vezes a poda, pode reduzir as características dos agrupamentos naturais.

Está sendo tratado uma série temporal de redes, ou seja, redes particionadas, e a rede integrada. Quando for selecionar *Pathfinder* ou *Minimum Spanning Tree*, será necessário tomar uma decisão sobre se deseja aplicar o algoritmo de poda em todas as partições de redes individuais (*pruning sliced networks*) ou somente as redes integradas (*pruning the merged networks*), ou ambas. Uma vez que a rede integrada é resultado de tudo que é feito com as redes particionadas, a poda da rede particionada só irá conduzir a uma rede integrada com *links* reduzidos. Se for selecionado os dois, será recebida uma rede integrada com menores números de *links*.



Figure 80. *Pruning ou link reduction. (Poda ou redução de link).*

6.6 *Visualização*

O *CiteSpace* por *default* só irá mostrar a fusão da rede integrada. Se desejar, pode-se ativar a opção para ver as redes com todos os intervalos de tempo. Se houver 20 intervalos de tempo, o *CiteSpace* abrirá 20 janelas extras para cada rede de intervalo de tempo – seja criterioso ao fazer suas escolhas!

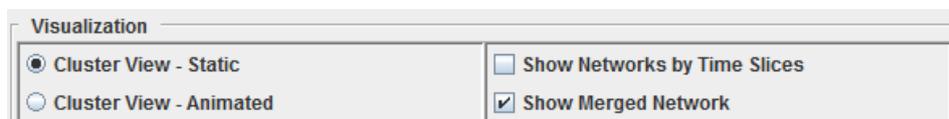


Figura 81. Opções de visualização.

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>

7 Interagindo com o CiteSpace

A maioria dos controles da aparência visual encontra-se no menu *Display*.

7.1 Como Mostrar ou Ocultar a Intensidade dos Links

Para ativar/desativar a exibição da intensidade do *link*, use a tela **Display ► Link Strength Show/Hide**.

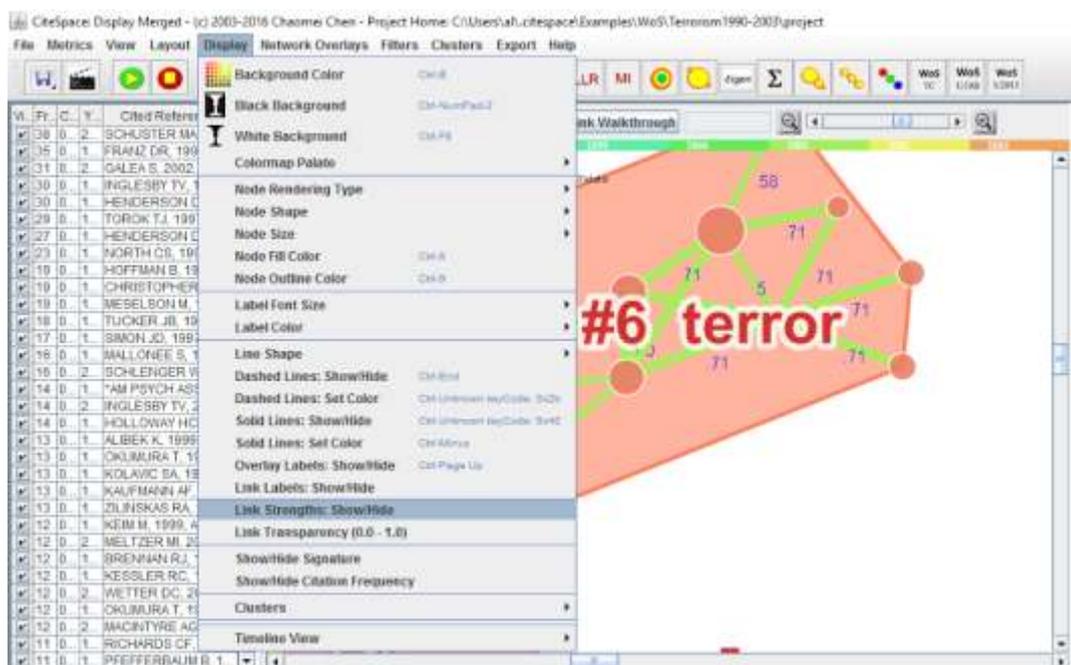


Figura 82. Mostrando a intensidade de cada *link* na tela.

7.2 Adicionando um Rótulo Persistente ao Nó

Além dos rótulos controlados pelos controles deslizantes de citação ou frequência, pode-se adicionar um rótulo para qualquer nó que se deseja. Clique com o botão direito do mouse sobre o nó a ser rotulado e escolha *Label the Node*.

Para limpar o rótulo, clique com o botão direito do mouse novamente sobre o nó e escolha *Clear the Label*.

Da mesma forma, pode-se "*bookmark*" (favoritos) um nó. Um "*bookmark*" será mostrado

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>

como uma estrela vermelha no centro do nó, como é mostrada na figura 83 no artigo de Schuster 2001.

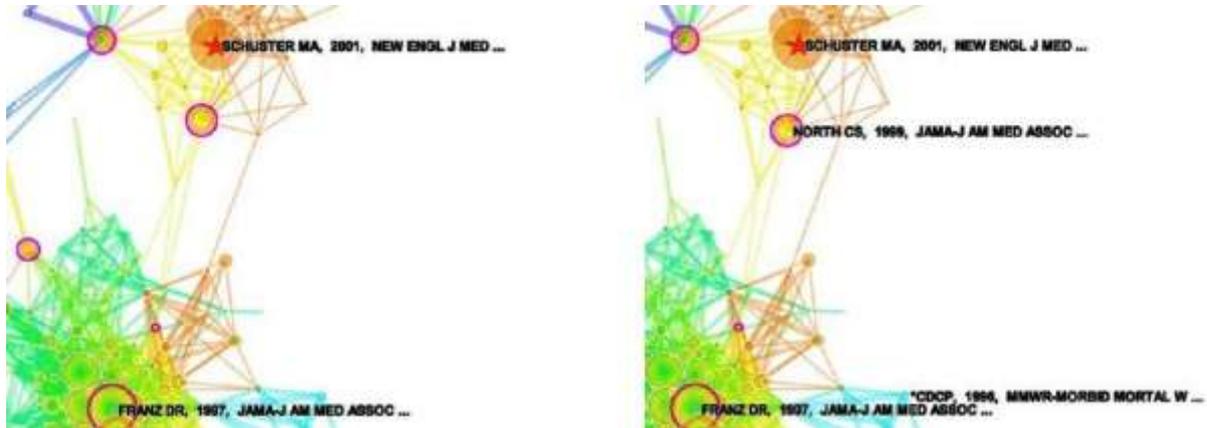
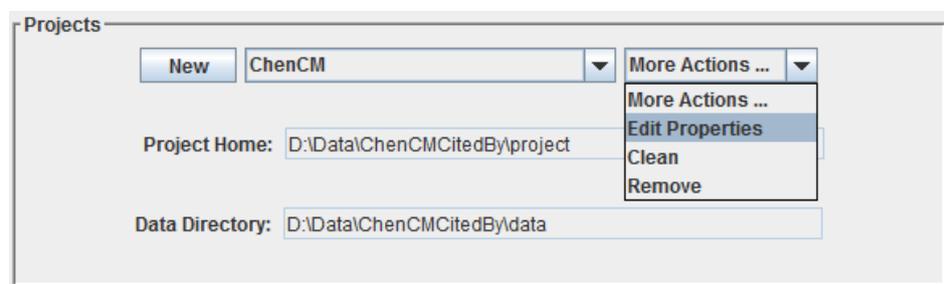


Figura 83. Para adicionar um rótulo persistente de um nó, clique com o botão direito do mouse no nó e escolha *Label the Node*.

7.3 Usando Pseudônimos para Fundir Nós

Se for notado que alguns nós da rede são variações da mesma entidade, podem-se usar pseudônimos para fundi-los, assim eles irão aparecer como um único nó. Por exemplo, uma rede de co-citação de autor da figura 86, CHEN CM e CHEN C são duas publicações do mesmo autor, portanto, elas devem ser fundidas em CHEN CM.

Para usar a função pseudônimo, primeiro edite as propriedades do projeto atual e, digite *on* na opção *Alias List (on/off)* e click no botão *Salve* (salvar).



A Figura 84. Editar as propriedades do projeto atual.

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>

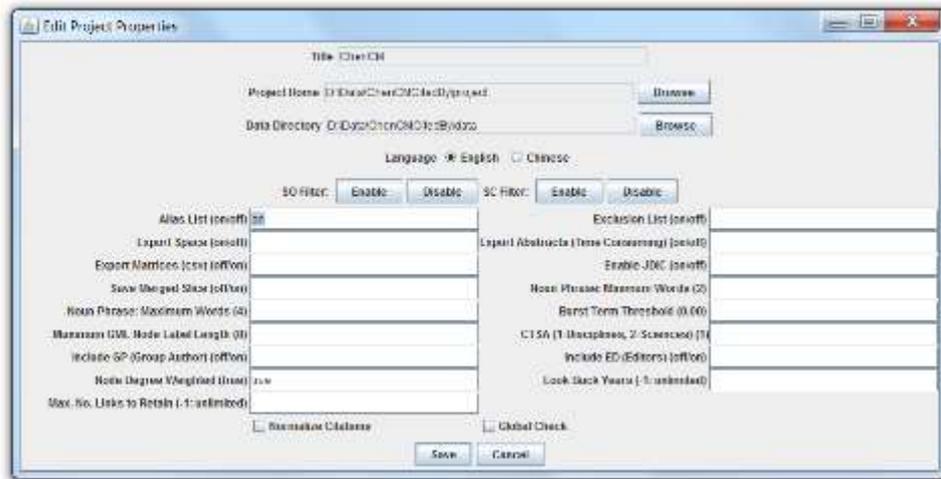


Figura 85. Certifique-se de que o *Alias List (on/off)* está ligado. Digite *on* na opção *Alias List (on/off)* e click no botão *Salve* (salvar).

Clique com o botão direito do mouse no nó CHEN CM e marque-o com o pseudônimo primário, clicando com na opção “*Add to the Alias List (Primary)*”. Clique com o botão direito do mouse no nó CHEN C e marque-o com o pseudônimo secundário, clicando na opção “*Add to the Alias List (Secondary)*”. O *CiteSpace* lembrará que se precisa reexecutar os processos para visualizar as mudanças.

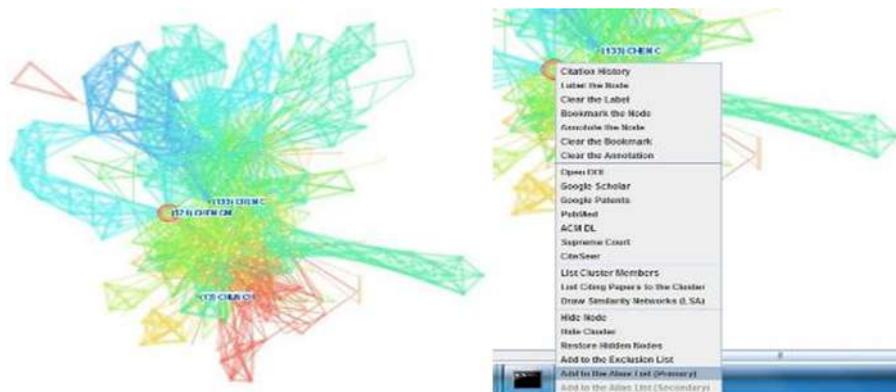


Figura 86. Clique com o botão direito do mouse no nó (574) CHEN CM e selecione “*Add to the Alias List (Primary)*” e selecione “*Add to the Alias List (Secondary)*” para o nó “(133) CHEN C.”

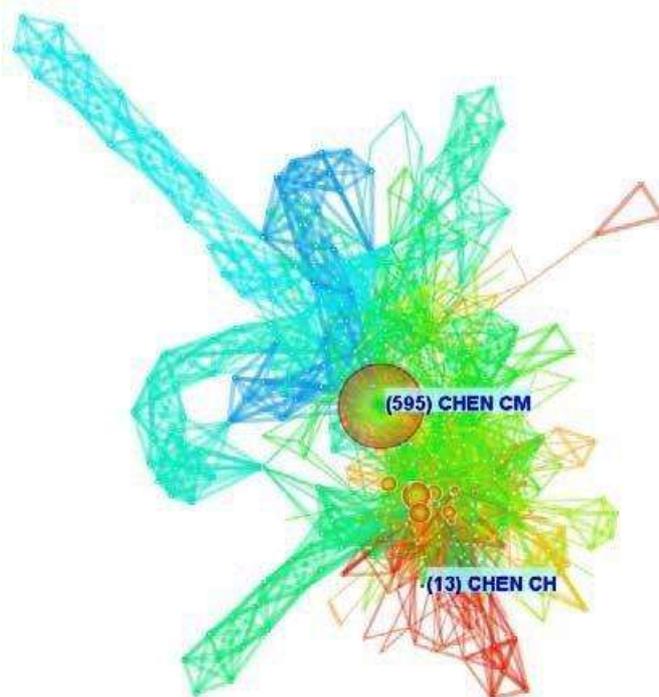


Figura 87. A rede visualizada após o CHEN CM e CHEN C serem fundidos.

Além de fundir nós de forma interativa por meio da interface gráfica do usuário, é possível editar o arquivo *citespace.alias* diretamente. O arquivo *citespace.alias* precisa ser localizado na pasta do projeto a ser executado. Pode-se utilizar um editor de texto para criar e editar um arquivo com nome *citespace.alias*.

O conteúdo do arquivo é formatado de acordo com as regras a seguir:

1. Cada um deles contém um par de nós de referências separados pelo caractere '#'. O nó referências podem ser referências citadas, autores citados, ou instituições.
2. A forma primária do pseudônimo, ou seja, o nó que se deseja manter deve aparecer em primeiro lugar, e, seguido pelo separador '#'. O pseudônimo secundário deve aparecer depois do separador '#', isto é, o nó que se deseja fundir no nó do pseudônimo primário.
3. Incluir o maior número de linhas (ou seja, aos pares) que se deseja.

Salve o arquivo e volte para o *CiteSpace*. Certifique-se de que o campo *Enable Alias* nas configurações do projeto esteja *on*. Em seguida, execute o projeto com o botão GO!

A maneira mais fácil de aprender sobre os vários formatos detalhados que se deseja usar é tentar com alguns exemplos por meio do modo interativo. Em seguida, abra o arquivo *citespace.alias* na pasta *Project* que foi gerado pelo *CiteSpace* com um editor de texto.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 88. Editando o conteúdo do arquivo *citespace.alias*.

7.4 Como Excluir um Nó da Rede

Pode-se excluir um nó da rede clicando com o botão direito do mouse no nó e selecionando "Add to the Exclusion List". Re-execute clicando no botão GO! Para recalcular o modelo da rede.

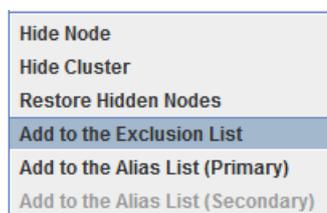


Figura 89. Excluir um nó do processo de modelagem da rede e também a visualização do processo.

A lista de exclusão é salva no arquivo *citespace.exclusion* na pasta do projeto. Pode-se editar diretamente esse arquivo de texto, por exemplo, adicionando novas entradas ou removendo as entradas existentes. Se desejar remover a lista de exclusão de uma só vez, basta renomear ou excluir o arquivo *citespace.exclusion*.

7.5 Como utilizar o Controle Deslizante de Visualização Fisheye

O controle deslizante *fisheye* está disponível para visualização da linha do tempo, assim podem-se visualizar os últimos anos exibidos numa tela maior que a dos anos anteriores. Conforme a figura 90, a maioria das publicações está lotada nos últimos anos na visualização linha de tempo original, porque algumas referências são datadas, a partir dos anos de 1625, provavelmente pelos filósofos.

Deslizando o controle deslizante *fisheye* de 0 a 6 vai ajudar a espalhar a tela cheia, assim pode-se visualizar cada *cluster* de atividades, de forma mais detalhada.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

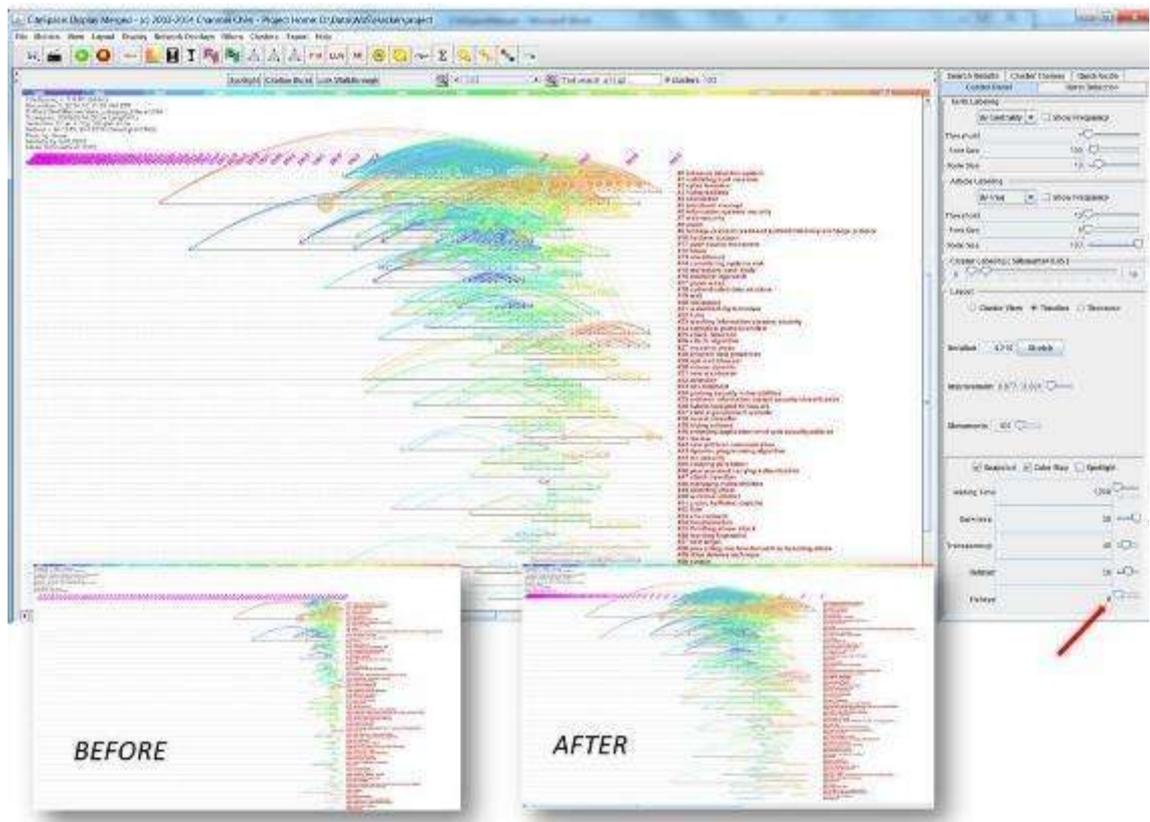


Figura 90. Use o controle deslizante *Fisheye* para ajustar o *layout* da linha de tempo para visualizar os últimos anos com mais detalhes.

7.6 Como Configurar Quando Calcular Automaticamente as Pontuações de Centralidade (*Centrality Scores*)

Na figura 91 mostra se o tamanho da rede é maior que 350 nós, o *CiteSpace* desativará o cálculo automático das pontuações de centralidade intermediária (*betweenness centrality*), para evitar um atraso desnecessário.

```
Centrality(240): The calculation of centrality is deferred due to the size of the network (653>350). Use CiteSpace->Preferences to reset the parameter.  
Centrality(345): The network exceeds the centrality turn-off point (653>350). Use CiteSpace->Preferences to reset the parameter.
```

Figura 91. *CiteSpace* desativou o cálculo automático de pontuações de centralidade intermediária (*betweenness centrality*).

Se o cálculo automático foi desativado, podem-se visualizar todos os valores na coluna de *centrality* preenchidos com 0s.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Visible	Freq	Cent...	Year	Cited References
<input checked="" type="checkbox"/>	38	0.00	2001	SCHUSTER MA, 2001, ...
<input checked="" type="checkbox"/>	35	0.00	1997	FRANZ DR, 1997, JAMA...
<input checked="" type="checkbox"/>	31	0.00	2002	GALEA S, 2002, NEW E...
<input checked="" type="checkbox"/>	30	0.00	1999	INGLESBY TV, 1999, JA...
<input checked="" type="checkbox"/>	30	0.00	1999	HENDERSON DA, 1999...
<input checked="" type="checkbox"/>	29	0.00	1997	TOROK TJ, 1997, JAMA...
<input checked="" type="checkbox"/>	27	0.00	1999	HENDERSON DA, 1999...
<input checked="" type="checkbox"/>	23	0.00	1994	*AM PSYCH ASS, 1994, ...
<input checked="" type="checkbox"/>	23	0.00	1999	NORTH CS, 1999, JAMA...
<input checked="" type="checkbox"/>	21	0.00	1994	MESELSON M, 1994, S...

Figura 92. Os valores da terceira coluna *Centrality* são todos 0s porque as pontuações de centralidade não foram calculadas automaticamente, neste exemplo devido aos 653 nós da rede é maior que o ponto de corte de 350 nó.

Para calcular manualmente a centralidade, use o Menu **Metrics ► Compute Centrality**.



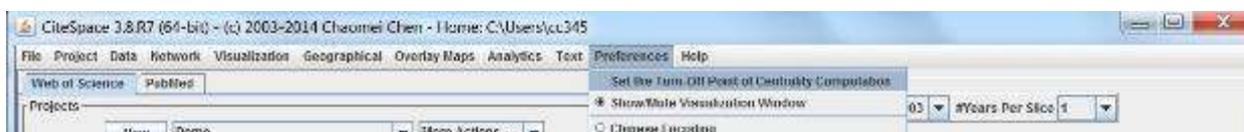
Figura 93. Iniciar manualmente o cálculo da centralidade.

Uma vez que o cálculo é concluído, podem-se visualizar os valores de centralidade diferente de zero na coluna *Centrality*.

Visible	Freq	Cent...	Year	Cited References
<input checked="" type="checkbox"/>	38	0.00	2001	SCHUSTER MA, 2001, ...
<input checked="" type="checkbox"/>	35	0.07	1997	FRANZ DR, 1997, JAMA...
<input checked="" type="checkbox"/>	31	0.05	2002	GALEA S, 2002, NEW E...
<input checked="" type="checkbox"/>	30	0.05	1999	INGLESBY TV, 1999, JA...
<input checked="" type="checkbox"/>	30	0.07	1999	HENDERSON DA, 1999...
<input checked="" type="checkbox"/>	29	0.05	1997	TOROK TJ, 1997, JAMA...
<input checked="" type="checkbox"/>	27	0.05	1999	HENDERSON DA, 1999...
<input checked="" type="checkbox"/>	23	0.11	1994	*AM PSYCH ASS, 1994, ...
<input checked="" type="checkbox"/>	23	0.06	1999	NORTH CS, 1999, JAMA...

Figura 94. A coluna *Centrality* tem valores diferentes de zero.

Pode-se alterar o limite padrão para desabilitar o cálculo automático da centralidade. Use o menu **Preferences ► Set the Turn-Off Point of Centrality Computation** e digite o número que deseja na caixa de diálogo. Por exemplo, se digitar 1000, o *CiteSpace* utilizará esse novo valor na próxima vez para determinar se calcula ou não automaticamente as pontuações de centralidade.



Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Figura 95. Defina o valor limite.

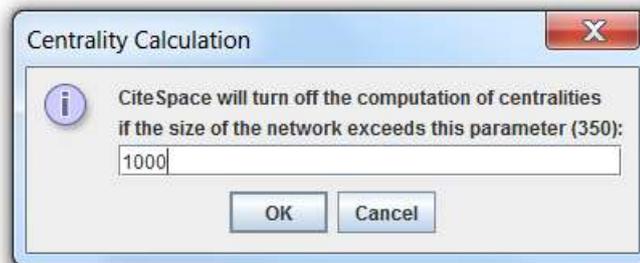


Figura 96. Se for definido o valor para 1000, o *CiteSpace* automaticamente calculará as pontuações de centralidade para redes com valores menores do que 1.000 nós.

7.7 Como Salvar a Visualização Como um Arquivo PNG

Pode-se salvar a visualização para um arquivo PNG de 300-dpi na pasta do projeto no seu computador. Clique no segundo ícone na barra de menus.



Figura 97. Clique no segundo ícone para salvar a visualização atual para um arquivo PNG de 300-dpi.

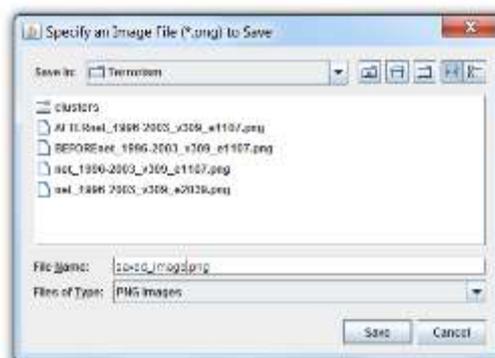


Figura 98. A pasta padrão é a pasta do projeto atual. O nome de arquivo padrão contém informações sobre a rede, ou seja, o número de nós é de 309 e o número de arestas é 2039.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Name	Date modified	Type	Size
saved_image	11/10/2014 3:43 PM	PNG Image	930 KB

Figura 99. O novo arquivo de imagem PNG é salvo em seu computador.

7.8 Filtros: Correspondência dos Registros na Base Pubmed

Se o tópico analisado está relacionado com a área médica, existe uma boa chance se encontrar pelo menos alguns dos registros na base de dados *PubMed*. O *CiteSpace* pode construir uma ponte entre a rede visualizada e os registros correspondentes na *PubMed*, fornecendo-lhe *links* diretos para a exibição desses registros no *PubMed*. No *PubMed* pode-se explorar um pouco mais, por exemplo, os artigos similares e outras informações.

Os passos básicos são os seguintes:

1. Usando registros bibliográficos (com as referências citadas) para gerar uma visualização de rede no *CiteSpace*.
2. Dividir a rede em *clusters* e rotular estes *clusters* como de costume.
3. Selecione o Menu *Filters* e escolha **Match records with PubMed**, e aguarde a conclusão do processo (como o *CiteSpace* cumpre com o protocolo NCBI, ele irá dizer-lhe com antecedência quanto tempo será necessário esperar).
4. Clique com o botão direito do mouse em um nó ou de um *cluster* do seu interesse, escolha **List Cluster Member** ou **List Citers to the Cluster**.



Figura 100. Escolha a função **Match Records with PubMed** do menu **Filters**.

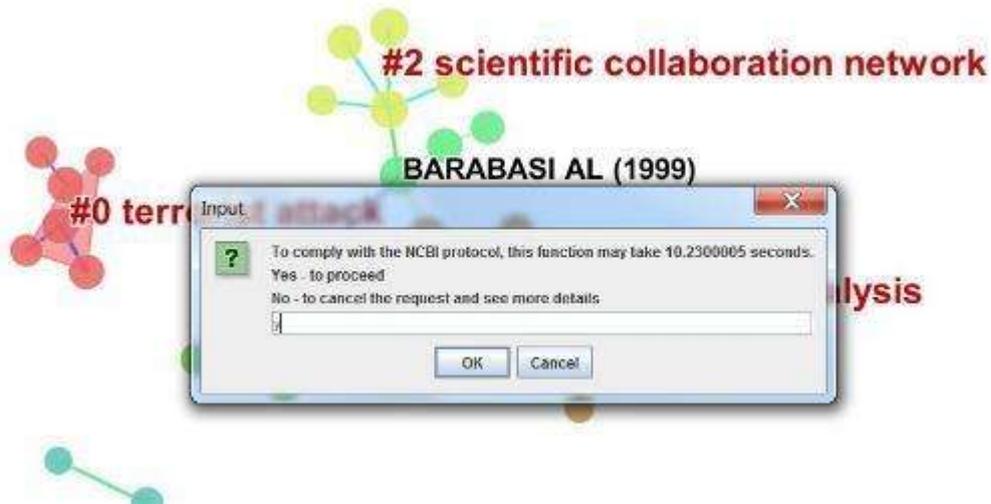


Figura 101. Para cumprir com o protocolo NCBI, o *CiteSpace* estima o tempo para concluir o processo.

```
PubMed(338): Searching for 31 records in PubMed ...
PubMed(383): 10% completed.
PubMed(383): 20% completed.
PubMed(383): 30% completed.
PubMed(383): 40% completed.
PubMed(383): 50% completed.
PubMed(383): 60% completed.
PubMed(383): 70% completed.
PubMed(383): 80% completed.
PubMed(383): 90% completed.
PubMed(389): Found 12 matches of the 31 records.
```

Figura 102. O andamento do processo é relatado na linha da janela do *prompt* de comando. Como se pode ver aqui, dentre os 31 registros no exemplo, foram encontrados apenas 12 registros no *PubMed*. Portanto, deve-se ter em mente que este será útil apenas se o tópico não estiver completamente fora do escopo do *PubMed*.

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>

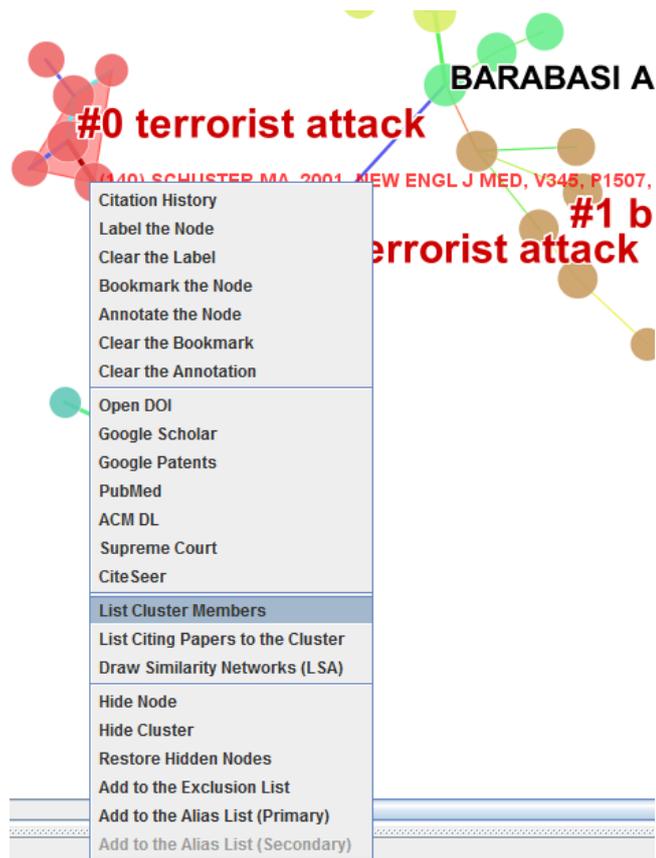


Figura 103. Clique com o botão direito do mouse em um nó de um Cluster #0 ataque terrorista (Schulster 2001) e selecione a opção **List Cluster Members**.

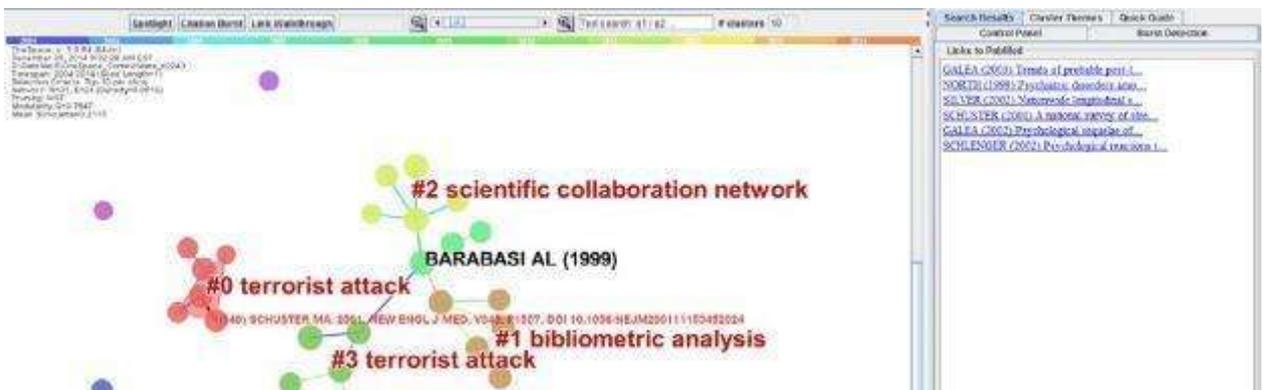


Figura 104. Os **links** dos membros do **cluster (#0)** que foram encontrados no **PubMed** serão mostrados no painel **Links to PubMed** sob a aba **Search Results**.

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 105. Clique em um *link* no painel **Links to PubMed** que levará para a página *PubMed* em uma nova janela do navegador.

- Citation History
- Label the Node
- Clear the Label
- Bookmark the Node
- Annotate the Node
- Clear the Bookmark
- Clear the Annotation
- Open DOI
- Google Scholar
- Google Patents
- PubMed
- ACM DL
- Supreme Court
- CiteSeer
- List Cluster Members
- List Citing Papers to the Cluster
- Draw Similarity Networks (LSA)
- Hide Node
- Hide Cluster
- Restore Hidden Nodes
- Add to the Exclusion List
- Add to the Alias List (Primary)
- Add to the Alias List (Secondary)

Figura 106. **List Citing Papers to the Cluster** abrirá uma janela *Pop Up* que resume os artigos que citam o *cluster*.

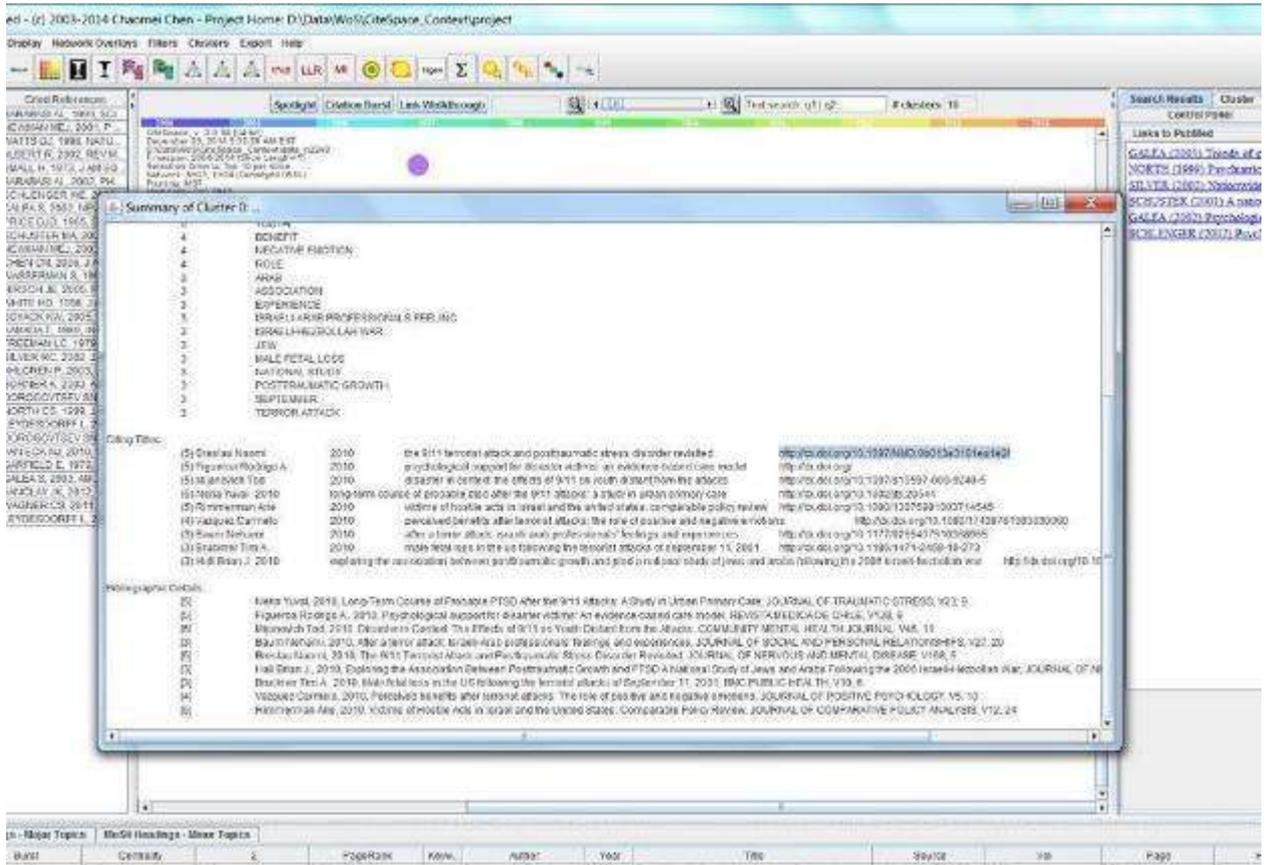


Figura 107. A janela *Summary of Cluster* atual (com base onde foi clicado com o botão direito), incluindo palavras-chave, artigos que citam e os *links* dos seus DOI. Pode-se copiar a URL de um *link* DOI e colá-lo em um navegador para obter o texto completo se houver uma assinatura adequada.

8 FUNÇÕES ADICIONAIS

O menu principal fornece acesso às funções adicionais.

8.1 Menu: Data (dados)

Data ► Import/Export

O *CiteSpace* fornece algumas funções de utilitários para facilitar a importação e exportação de dados.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

8.1.1 Banco de Dados incorporado no *CiteSpace*

O *CiteSpace* fornece uma interface de usuário para uma base de dados MySQL no *localhost* (servidor local). A interface do usuário fornece diversas funções para importação e exportação de registros em conexão com o banco de dados.

Antes de se usar este grupo de funções, é necessário configurar seu MySQL como se segue.

No seu computador, localize a sua própria pasta de usuário e encontre a pasta *.citespace*. Crie um arquivo de texto *mysql.ini* com o nome de pares de valores separados por tabulação, conforme o conteúdo a seguir:

```
host    localhost
user    user_id
pass    password
```

onde *user_id* e *password* são seu id e senha para o seu próprio *login* do MySQL.

Certifique-se de que o seu servidor MySQL está ativado antes de utilizar esta função no <i>CiteSpace</i> .
--

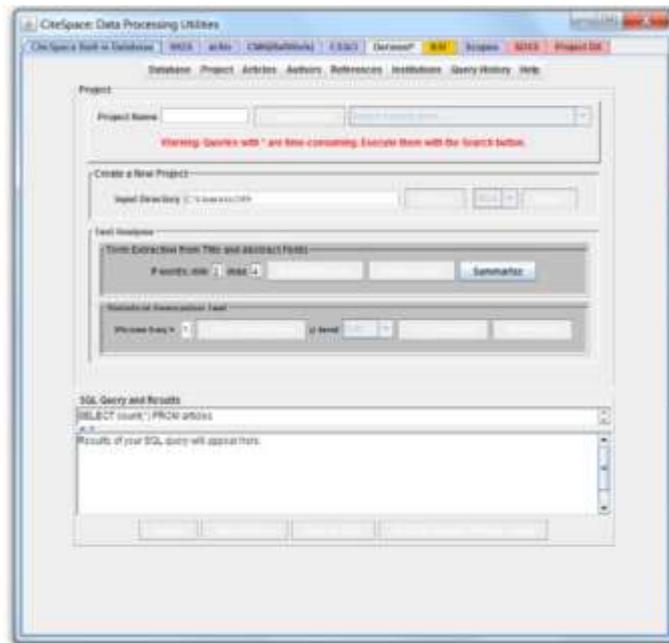


Figura 108. Utilitários de Processamento de Dados.

Depois de conectar ao banco de dados, irão aparecer os projetos existentes, isto é, projetos que foram carregados para o banco de dados. Note que os projetos aqui são armazenados no banco de dados e que eles são diferentes dos projetos que apareceram na interface principal do *CiteSpace*, que são baseados em arquivo, ou seja, os arquivos baixados da *Web of Science*. Podem-se importar os arquivos baixados para o banco de dados, editá-los e exportá-los no formato de arquivos da *Web of Science*.

Uma vez que o banco de dados é um MySQL em *localhost*, pode-se acessar o banco de dados diretamente com o seu próprio *login* do MySQL. Podem-se usar este banco de dados para processar seus dados antes de se aplicar funções de visualização sobre eles.

8.1.1.1 Estrutura da Base de Dados

O nome do banco de dados é “*wos*”. Ele contém as tabelas a seguir:

TABELA *articles*

id(int), uid, project, author, title, abstract, source, j9, volume, issue, bp, ep, page, dt, doi, year(int), month(int), date(int), citations(int), editor, tagged(boolean)

TABELA *authors*

id(int), lastname, firstname, initials, project, uid, pos

TABELA *refs*

id(int), bibcode, ref, doi, author, year, source, volume, page, citer_uid, project

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

TABELA keywords

id(int), keyword, uid, year, project, type

TABELA phrases

id(int), phrase, isTitlePhrase(boolelan), project, uid, year(int), month(int), date(int), freq(int)

TABELA verbs

id(int), project, uid, verb, freq

TABELA bursts

id(int), project, term, weight(double), start(int), end(int)

TABELA institutions

id(int), name, country, uid, year(int), project



Figura 109. Um gráfico de um projeto no banco de dados incorporado no *CiteSpace*.

Articles ► Most Cited Articles

Pode-se consultar o banco de dados com algumas funções integradas em um conjunto de dados (*dataset*) carregados. Por exemplo, podem-se encontrar os artigos mais citados no projeto atual. A consulta SQL é exibida juntamente com os resultados. Isso irá ajudar a se familiarizar com a estrutura interna do banco de dados.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

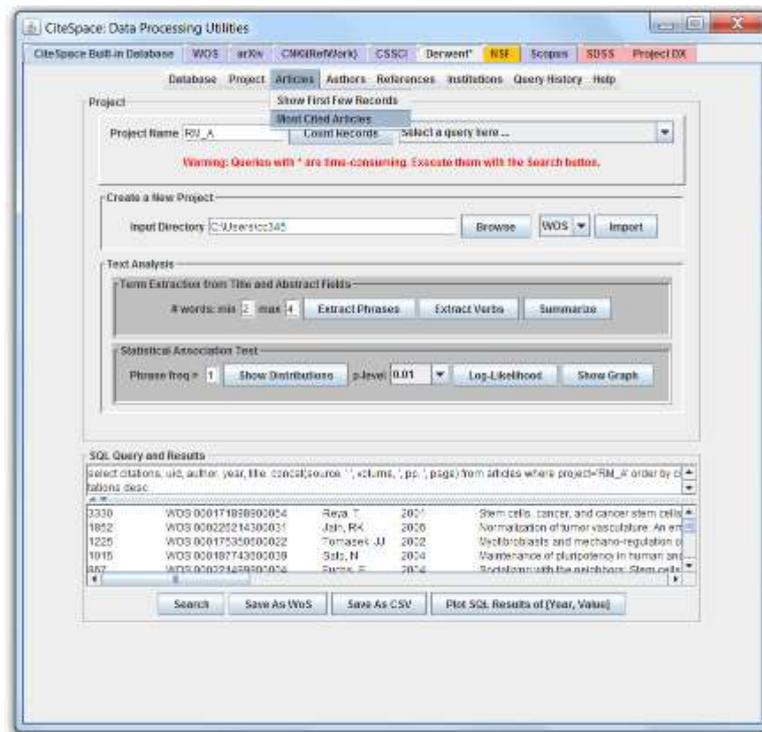


Figura 110. Usando uma função pré-definida para encontrar os artigos mais citados utilizando uma consulta SQL.

8.1.2 Funções de Utilitário para o Formato *Web of Science*

8.1.2.1 Como remover registros duplicados

Podem-se juntar vários conjuntos de dados que foram “baixados”, juntando os arquivos baixados para a mesma pasta de dados. Se alguns arquivos tiverem os mesmos nomes, precisará ser renomeados primeiramente, para resolver os conflitos antes de movê-los em conjunto. A maneira mais simples é adicionar um sufixo para os nomes dos arquivos. Por exemplo, se há dois conjuntos de dados e cada um contém um arquivo chamado `download_500.txt`, pode-se renomeá-lo para `download_500_part1.txt` e `download_500_part2.txt`.

É necessário certificar-se de que os arquivos integrados não tenham registros duplicados. O *CiteSpace* tem uma função utilitária para isso. Especifique a pasta de entrada e a pasta para salvar uma cópia do conjunto de dados após a remoção dos duplicados, então pressione o botão "Remove duplicates (WoS)". Note que o formato dos arquivos de entrada deve estar no formato *Plain Text* da *Web of Science*.

Chen, C. (2014) o *CiteSpace Manual*. <http://cluster.cis.drexel.edu/~cchen/citespace/>

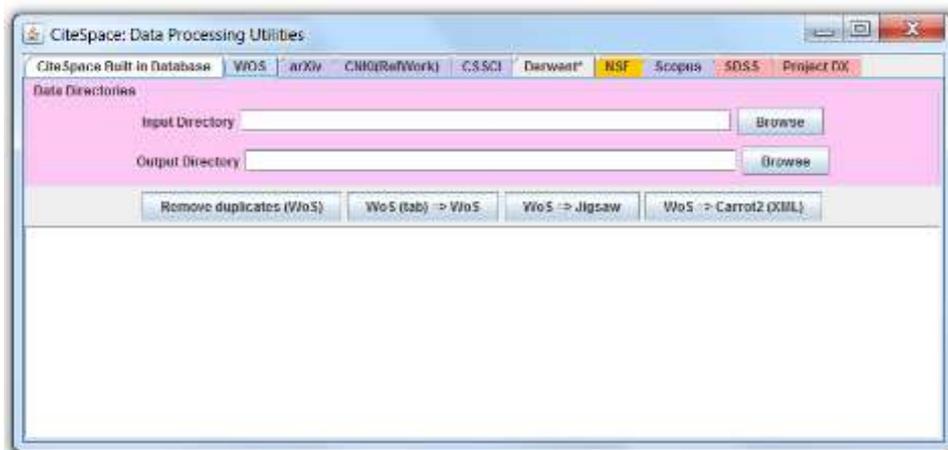


Figura 111. Funções utilitárias para trabalhar registros bibliográficos no formato da *Web of Science*.

8.1.2.2 Converter o Tab Delimited WoS Format

Pode-se converter o formato delimitado por tabulação *WoS* para o formato *Plain Text* (ou seja, cada campo é marcado por um código de duas letras, como AU, TI, e AB) usando outra função utilitária “WoS (tab) WoS.”

8.1.2.3 Converter o WoS Format for Jigsaw

Pode-se converter arquivos no formato *WoS* para o formato que possa ser processado por *Jigsaw* – uma aplicação visual analítica, a qual está também disponível em versão livre. (Stasko, Gorg, & Liu, 2008).

8.1.2.4 Converter o WoS Format for Carrot 2

Pode-se converter arquivos no formato *WoS* para o formato que possa ser processado por *Carrot2* – uma ferramenta de código aberto de busca e visualização de texto ("*Carrot2: Open source framework for building search clustering engines*," 2012). Os arquivos convertidos são documentos XML.

8.1.3 Base de Dados Scopus

O formato preferido para os arquivos da base de dados Scopus é o formato RIS. Na Scopus, salve os resultados da busca para um ou mais arquivos de dados em formato RIS, por

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

exemplo, *my_scopus_search_results.ris*, para uma pasta em seu computador. Especifique a pasta de dados e uma pasta onde se deseja ter os arquivos convertidos. E inicie o processo de conversão.

O *CiteSpace* mostrará quantos registros em cada arquivo RIS foram convertidos e, mais especificamente, quantas referências citadas no total foram encontradas nos arquivos de dados e como alguns deles foram convertidos com sucesso. A taxa 95,0% considera-se bem-sucedida, levando-se em conta todas as irregularidades das referências citadas.

Se, por algum motivo, se possuir o arquivo de dados da Scopus em formato CSV ao invés do formato RIS, precisa-se fazer uma rápida conversão de CSV para um formato delimitado por tabulações, utilizando a planilha Excel antes de se utilizar a função pré-definida, que permite converter um arquivo delimitado por tabulações para o formato da *Web of Science* no *CiteSpace*.

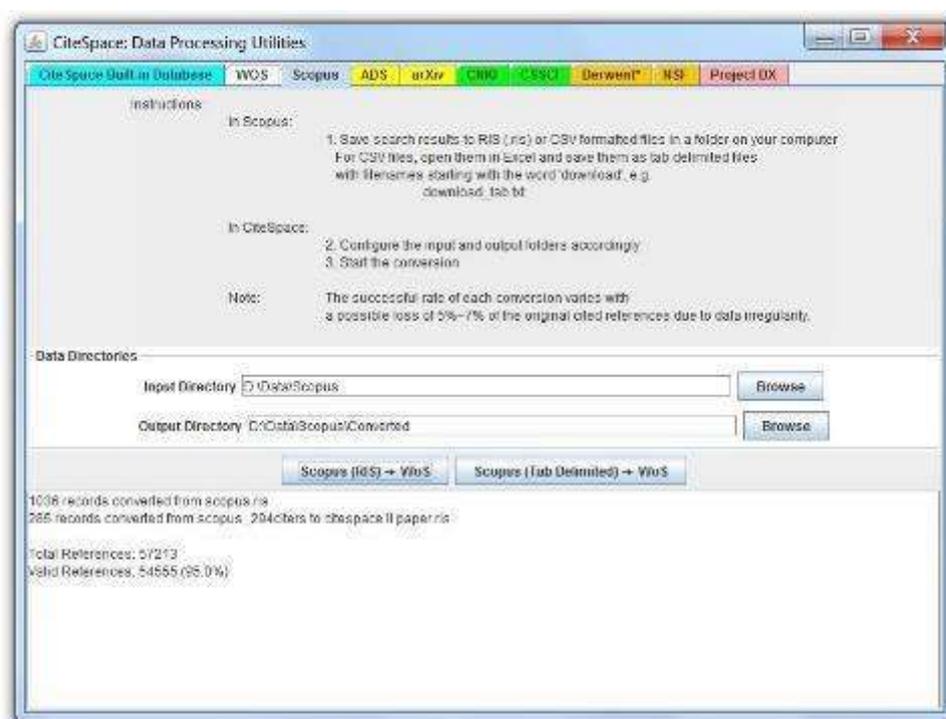


Figura 112. Convertendo arquivos de dados da Scopus para o formato WoS do *CiteSpace*.

Name	Date modified	Type	Size
download_converted_1	12/29/2014 11:22 ...	TXT File	4,658 KB
download_converted_2	12/29/2014 11:22 ...	TXT File	1,176 KB

Figura 113. Dados da Scopus convertidos são salvos na pasta de saída designada.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

8.1.4 Base de Dados PubMed

O *CiteSpace* permite que se recupere registros bibliográficos da *PubMed*. Por exemplo, para recuperar os registros sobre a hipertensão arterial com base na *MeSH headings* pode-se usar a consulta "hipertension [mh]" entre 2008 e 2010. Pode-se especificar o número máximo de registros que se deseja recuperar cada ano. Para fins meramente ilustrativos, pode-se limitar o número máximo de 25 por ano. Registros recuperados serão salvos em uma pasta especial *\$your_username\PubMed\SearchResults*.

Uma vez que a recuperação de dados for concluída, precisa-se mudar para a *Web of Science* e analisar os dados da mesma forma que se fez com um conjunto de dados da *Web of Science*.

Em virtude de os registros *PubMed* não incluírem informações sobre referências citadas, não é possível realizar análise de citações, ou seja, não se pode escolher os tipos de nó como referências citadas, autores citados, ou revistas citadas. No entanto, podem-se realizar outras análises, como as redes de colaboração de autores, termos, palavras-chave e categorias.

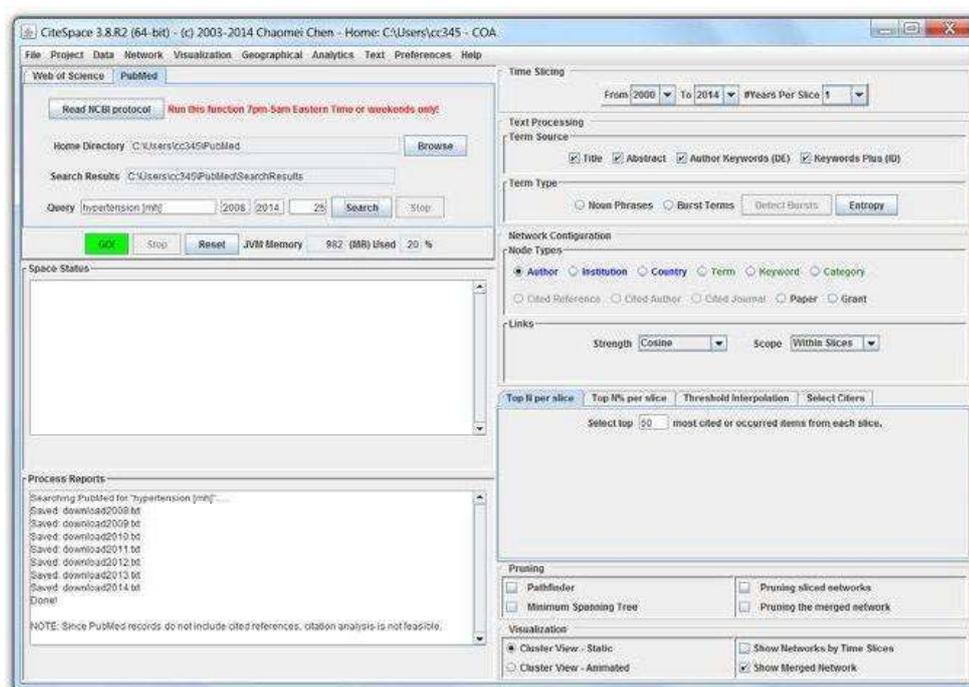


Figura 114. Recuperar registros bibliográficos da PubMed.

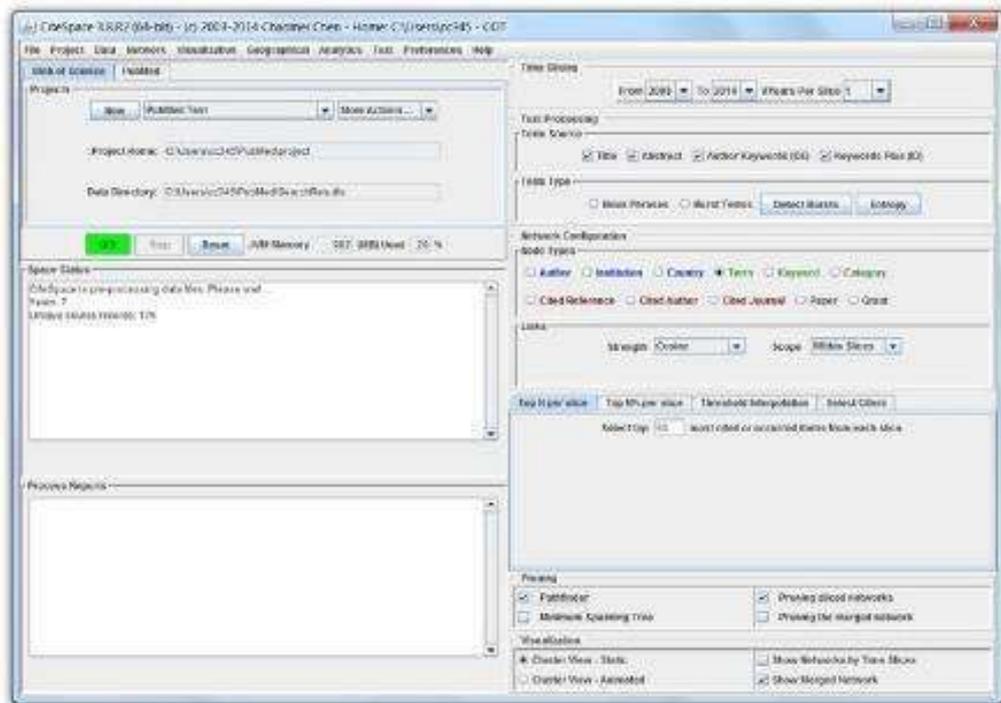


Figura 115. Analisando os registros da PubMed...

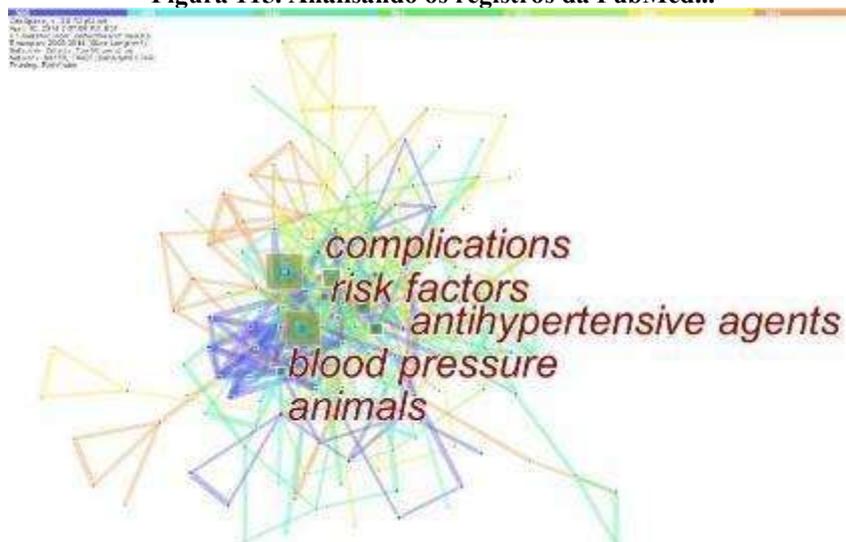


Figura 116. Uma rede de co-ocorrências de substantivos sobre a hipertensão.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

8.2 Menu: Network

8.2.1 Lote de Exportação para arquivos Pajek.net

8.3 Menu: Geographical

8.3.1 Gerar mapas no *Google Earth*

A localização geográfica dos autores em seus registros de publicações pode ser mapeada para um mapa geoespacial em KML. Pode-se usar o *Google Earth* como a interface para explorar as localidades dos autores e *links* para seus colaboradores. Também se podem ir para os artigos originais diretamente no *Google Earth*.

Para gerar o arquivo do mapa, precisa-se especificar uma pasta de dados que contenha registros bibliográficos no formato da *Web of Science* (texto simples), que é o mesmo formato de projetos do *CiteSpace*. Desta vez, precisa-se apenas da pasta de dados. A nova pasta será criada automaticamente sob a pasta de dados chamada kml. O arquivo gerado KML será salvo na pasta kml quando o processo de geocodificação for concluído.

O gerador do mapa *Google Earth* do *CiteSpace* precisa saber o período que se está interessado, similar a configuração de intervalo de tempo na interface principal do *CiteSpace*. Navegue até a pasta de dados de seus dados e clique no botão “*Make Map*”. Isso pode levar algum tempo para que o processo seja concluído.



Figura 117. O Gerador *Google Earth* KML.

Uma vez que o mapa é gerado, aparecerá uma mensagem informando onde está o arquivo do mapa, no formato kmz, ou seja, um arquivo KML compactado.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 118. O mapa foi gerado.

Se houver alguns erros relatados pelo gerador, pode-se verificar o arquivo de *log* de erro - *geocoding_log_tab.txt* - fazer correções necessárias nos dados e em seguida repetir o processo. O mapa é armazenado no arquivo *master-medium.kmz* se for usada a escala padrão de médias.

Name	Date modified	Type	Size
geocoding_log_tab	4/12/2014 10:55 PM	TXT File	47 KB
locations-2006	4/12/2014 10:53 PM	Microsoft Excel Co...	6 KB
locations-2007	4/12/2014 10:54 PM	Microsoft Excel Co...	8 KB
locations-2008	4/12/2014 10:54 PM	Microsoft Excel Co...	11 KB
locations-2009	4/12/2014 10:54 PM	Microsoft Excel Co...	16 KB
locations-2010	4/12/2014 10:55 PM	Microsoft Excel Co...	14 KB
locations-2011	4/12/2014 10:55 PM	Microsoft Excel Co...	21 KB
locations-2012	4/12/2014 10:55 PM	Microsoft Excel Co...	17 KB
locations-2013	4/12/2014 10:55 PM	Microsoft Excel Co...	15 KB
locations-2014	4/12/2014 10:55 PM	Microsoft Excel Co...	5 KB
master-medium	4/12/2014 10:55 PM	KMZ File	59 KB

Figura 119. Os arquivos gerados na pasta KML.

Se *Google Earth* está instalado em seu computador, é só dar dois cliques sobre o arquivo kmz.



Figura 120. Rede de colaboração do autor é mostrada no *Google Earth*.

Sob os *Places*, será visto uma lista de anos como camadas. Pode-se selecionar ou desmarcar essas camadas, marcando ou desmarcando a caixa de seleção na frente delas para que se possam controlar quais anos de dados que se deseja visualizar. Trabalhos em co-autoria nos anos mais recentes estão ligados por linhas em vermelho, enquanto que as colaborações mais antigas são mostradas em linhas verde ou azul.

Pode-se marcar uma camada de um ano para uma localização, em seguida, uma lista de artigos publicados pelos autores nesta localização. Cada artigo da lista é selecionável. Ela levará para o texto integral por meio dos *links* do DOI. Precisa-se ter a assinatura para acessar os documentos na base de dados.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

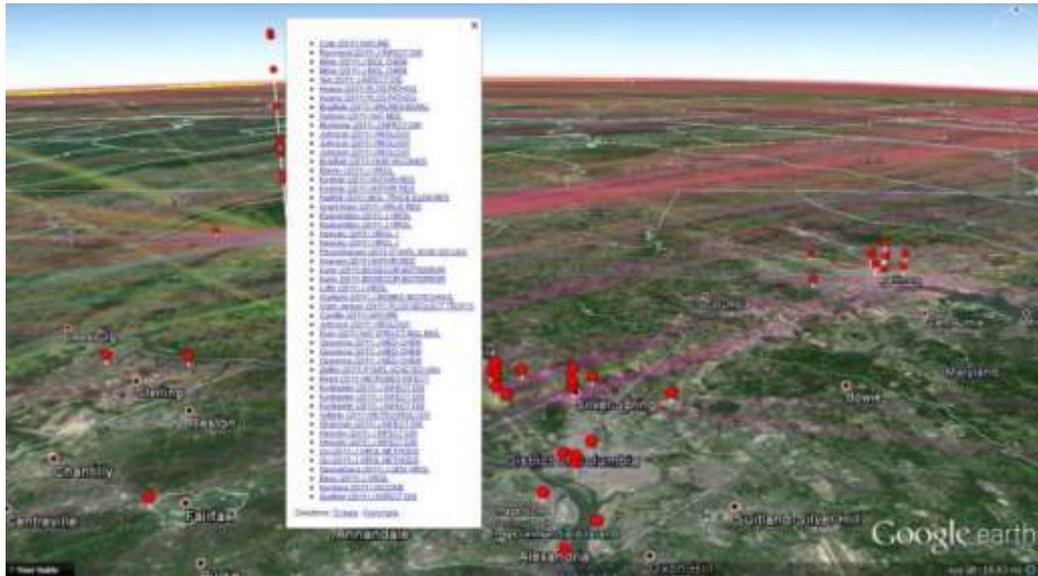


Figura 121. Abrindo a lista dos lugares à esquerda e localizando uma cidade de interesse: Frederick, Maryland, EUA - sobre Ebola.

Clique em qualquer um dos artigos da lista para explorar o seu conteúdo. Aqui é um exemplo do que será visualizado depois de clicar em um *link* para o artigo 2010 JASIST no *Google Earth*.



Figura 122. Clicando sobre o *link* Chen (2010) levará a página do editor do artigo.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

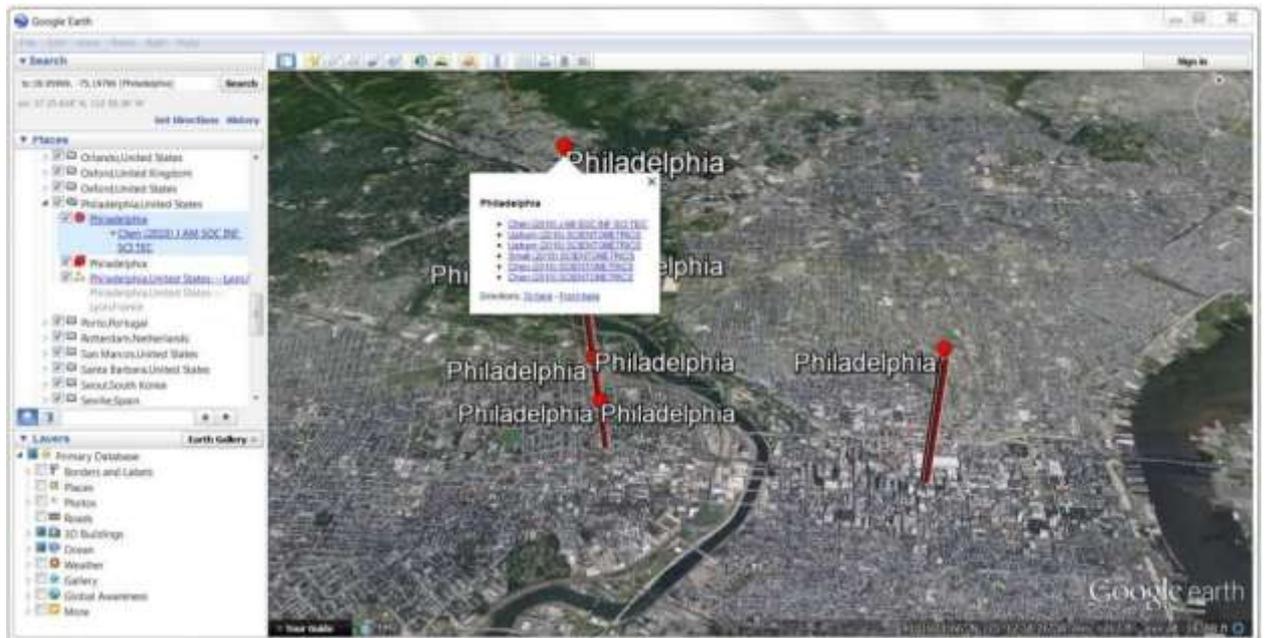


Figura 123. Aqui é a vista aérea do centro da cidade de Philadelphia. A barra vermelha do lado esquerdo é o campus principal da Drexel.

8.4 Menu: Overlay Maps

Dois mapas sobrepostos são introduzidos em (C. Chen & Leydesdorff, 2014).

Esta função está disponível para uso não comercial e educacional. Para uso comercial, favor entrar em contato direto com chaomei.chen@drexel.edu para futuros detalhes.

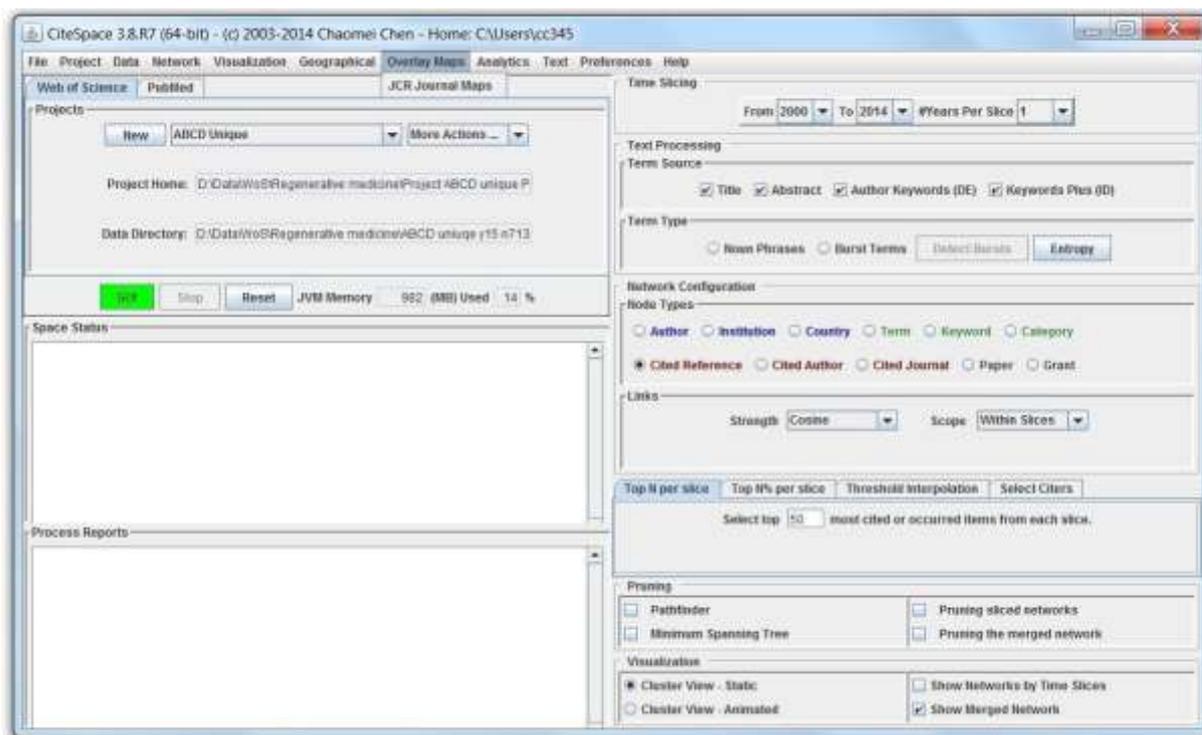


Figura 124. A função *dual-map overlay* para mapas está acessível a partir do menu *Overlay Maps*.

8.4.1 Adicionar uma Sobreposição (*Overlay*)

A versão atual permite adicionar até 12 sobreposições (*overlays*). Cada sobreposição é representada por um conjunto distinto de registros bibliográficos em formato WoS. Iniciar o processo clicando no botão *Add*.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 125. Iniciar o processo de adição de um *overlay* clicando no botão *Add*.

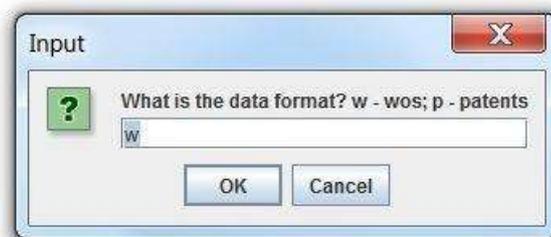


Figura 126. Selecione 'w' para registros bibliográficos no formato WoS. A opção 'p' não está disponível na versão atual.

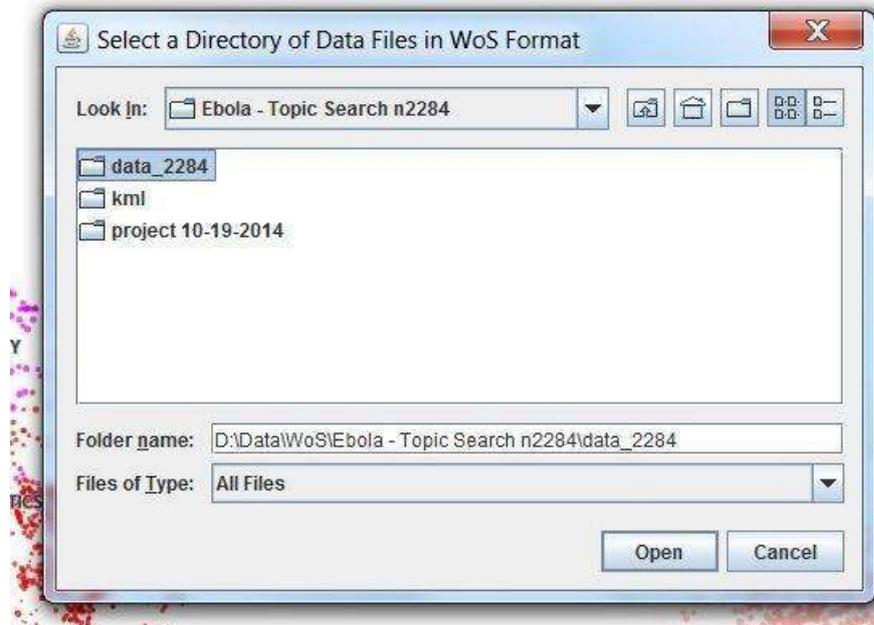


Figura 127. Seleccione o diretório de dados de um conjunto de dados da mesma forma como foi criada um projeto *CiteSpace*.



Figura 128. Seleccione uma cor para o novo *overlay* e espere a tela atualizar ...

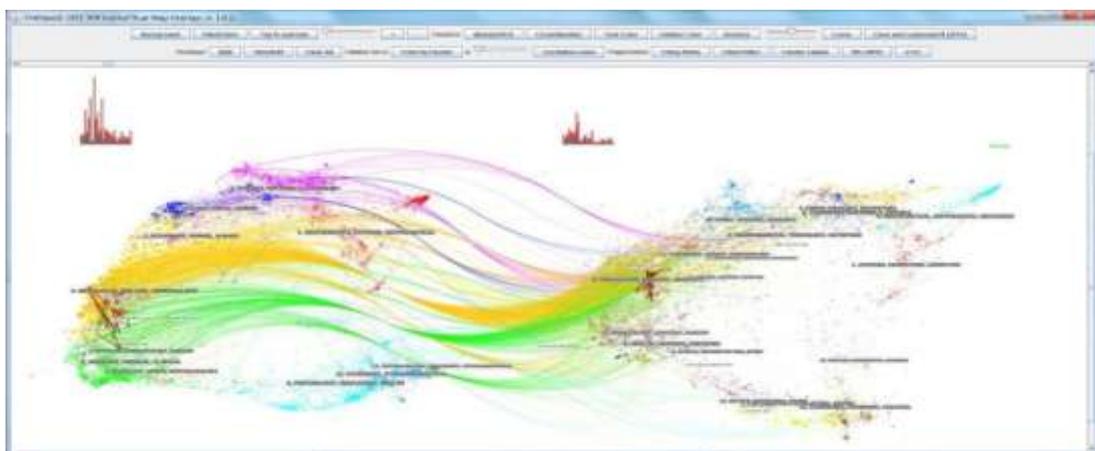


Figura 129. O resultado da adição de um conjunto de registros bibliográficos no formato WoS.

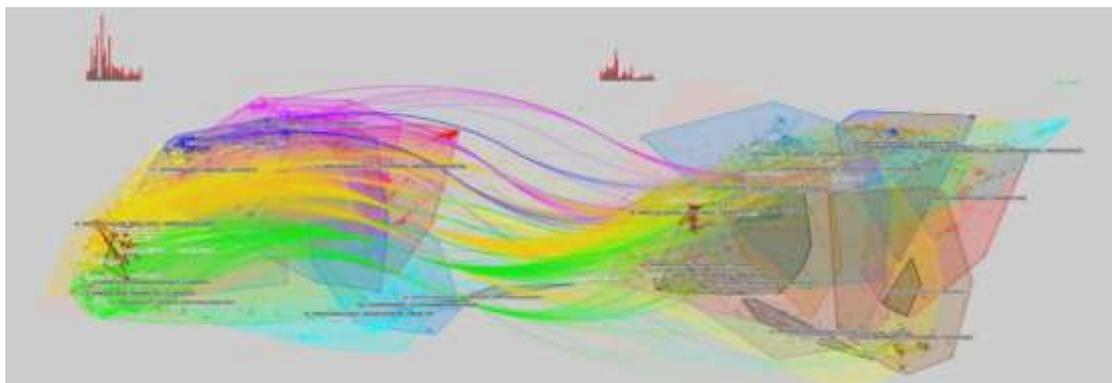


Figura 130. O botão *Border* controla a tela da borda de um *cluster*, e permite preencher essas áreas em cores.

8.4.2 Outras Leituras e Termos de Uso

Detalhes Técnicos e alguns exemplos são fornecidos em (C. Chen & Leydesdorff, 2014).

Esta função está disponível para uso não comercial e educacional. Para uso comercial favor entrar em contato comigo para futuros detalhes.

Se for utilizar a função *dual-map* em suas publicações, deve-se citar o seguinte:

Chen, C., Leydesdorff, L. (2014) *Patterns of connections and movements in dual-map overlays: A new method of publication portfolio analysis*. *Journal of the American Society for Information Science and Technology*, 65(2), 334-351.

8.5 *Menu: Text*

8.5.1 Árvores conceituais e árvores de predicado

Árvores conceituais e árvores de predicado no *CiteSpace* são geradas a partir de três tipos de documentos textos não estruturados: 1) cortar e colar o texto para uma janela de entrada, 2) a partir de arquivos de texto completo, e 3) a partir de uma pasta de arquivos no formato WoS, incluindo os arquivos de dados que foi “baixado” diretamente a partir do WoS e arquivos intermediários salvos na pasta do projeto depois de ter executado o algoritmo de agrupamento para a atual rede.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

O exemplo a seguir mostra como gerar árvores de conceito e árvore de predicado dos registros que citam o maior *cluster* no projeto de demonstração (ou seja, a pesquisa terrorismo). Primeiro, selecione o conjunto Projeto Demo como o projeto atual. A seguir escolha no menu Text ► Build Concept/Predicate Trees (from WoS files).

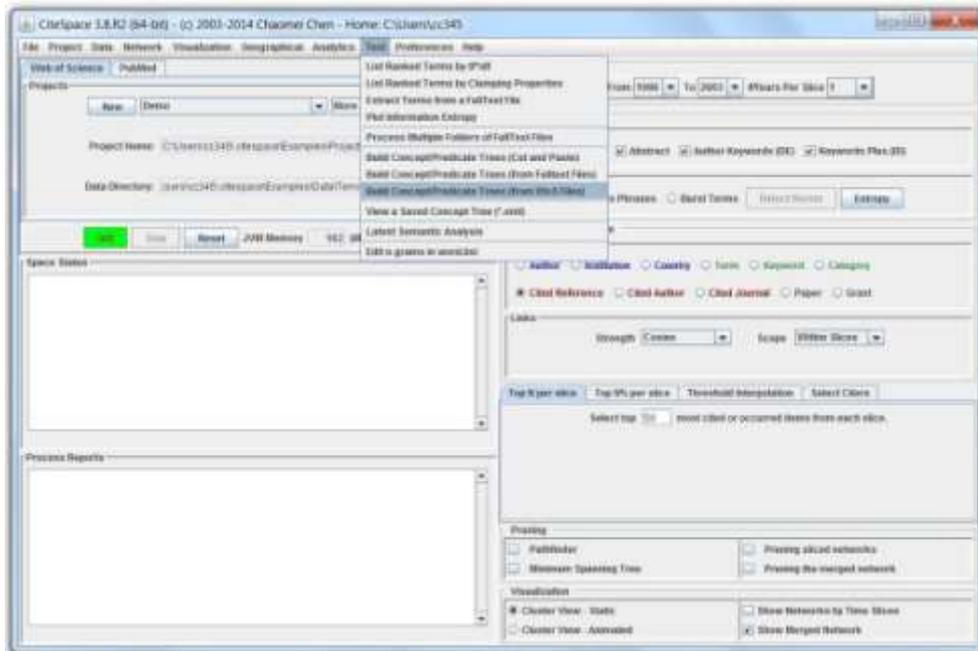


Figura 131. Gerar árvores conceituais e árvores de predicado.

É necessário selecionar o arquivo que representa as citações para o maior *cluster* do Projeto Demo. O *CiteSpace* mostrará uma lista de pastas e arquivos. Selecione a pasta *clusters*, então 0.txt, que corresponde ao cluster #0, o maior *cluster*.

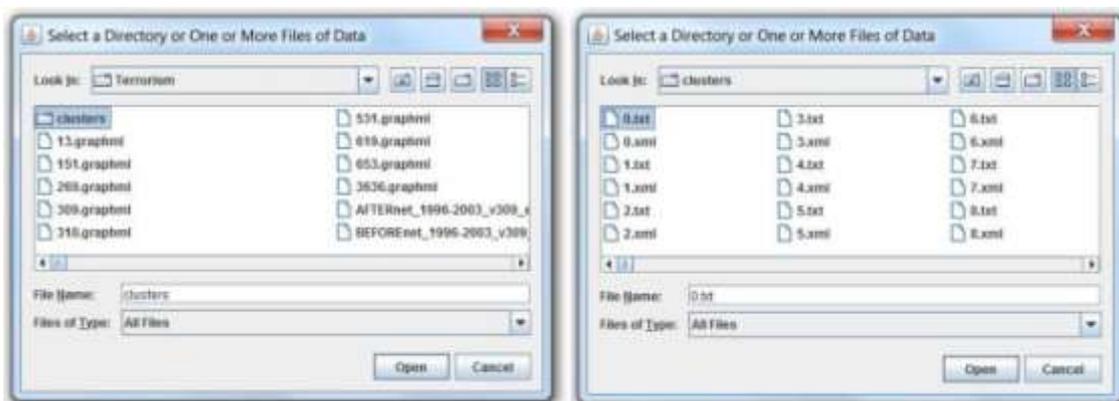


Figura 132. Selecione a pasta *clusters* do Projeto Demo e o maior *cluster* #0.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

A janela da árvore conceitual tem três painéis. A janela da árvore mostra uma visualização da árvore conceitual. A janela de contexto mostra as sentenças que contêm um conceito, ou seja, o nó na árvore conceitual. O exemplo a seguir mostra quando se move o cursor do *mouse* sobre o nó bioterrorismo na janela *Tree*. Diversas frases que contêm o termo *bioterrorism* são mostradas como nós filhos do conceito, por exemplo, *threat (of) bioterrorism*, *weapons and agent (of) bioterrorism*.

Os nós próximos do topo da árvore são os principais conceitos e principais interesses do *cluster*. Assim, se sabe que o maior *cluster* no Projeto Demo trata sobre o bioterrorismo, Estados Unidos, ataque biológico, e resposta eficaz. Esses conceitos, em conjunto, fornecem um sentido bastante focado da natureza do *cluster*.

Para deslocar a árvore visualizada, mantenha pressionado o botão esquerdo do *mouse* e movimente na direção desejada.

Para fazer um *zoom* da árvore visualizada, mantenha pressionado o botão direito do *mouse* e mova-o para cima para aumentar (*zoom in*) ou para baixo para diminuir (*zoom out*).

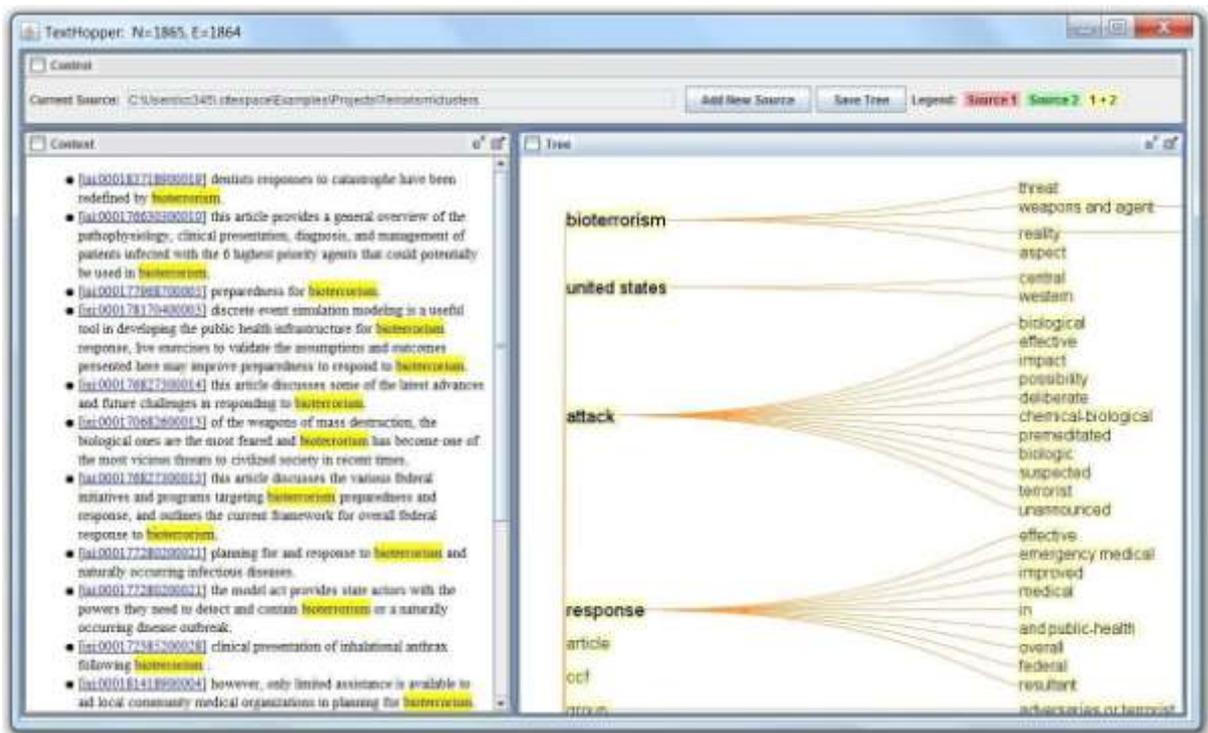


Figura 133. O conceito de árvore de *cluster* #0 - *bioterrorism* no Projeto Demo.

Na janela de controle, pode-se adicionar uma nova fonte para o conceito de árvore existente. Aqui vamos adicionar o segundo maior *cluster* para que se possa visualizar o que esses dois maiores *clusters* têm em comum e onde exatamente eles diferem. Lembre-se de que o segundo maior *cluster* é identificado como PTSD - *post traumatic stress disorder*.

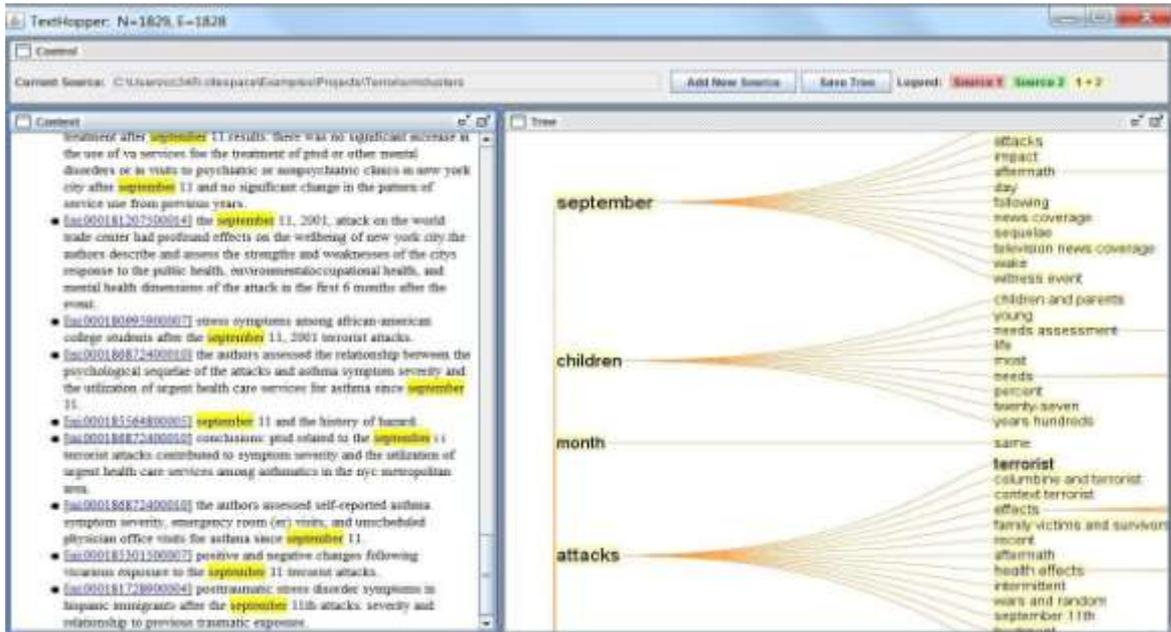


Figura 134. O cluster PTSD. Conceitos-chave: *September, children, same month, e terrorist attacks.*

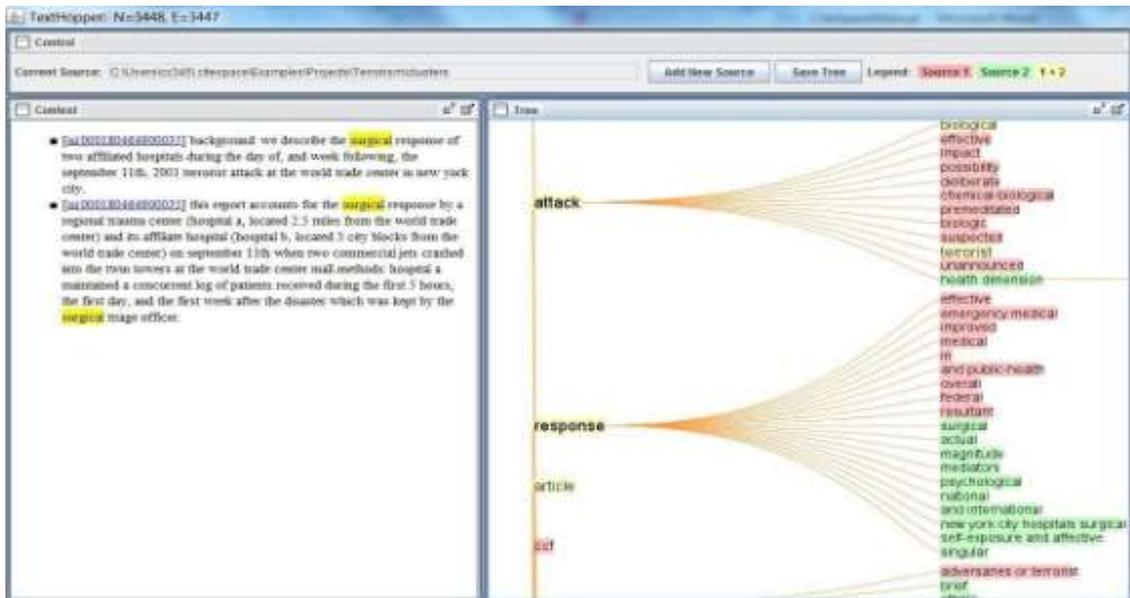


Figura 135. A árvore conceitual de duas fontes. O cluster bioterrorismo está em vermelho. O cluster PTSD está em verde. A sobreposição entre os dois está em amarelo.

8.5.2 Lista de termos por Propriedades de Agrupamento

No menu *Text*, pode-se encontrar diversas funções que lidam com o texto.

Por exemplo, **Text ► List Ranked Terms by Clumping Properties** pode classificar termos pelas suas propriedades de *Clumping* (*serial-clustering*), ou seja, como eles tendem a aparecer no texto (Bookstein, Klein, & Raita, 1998). No Projeto Demo, os termos mais proeminentes incluem ataques terroristas, *World Trade Center*, destruição em massa e terrorismo biológico.

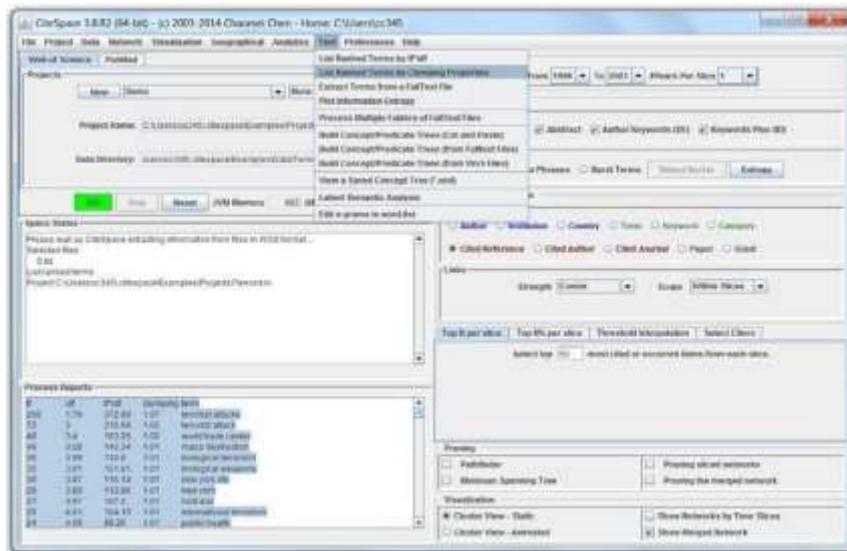


Figura 136. Lista de termos classificados utilizando a propriedade de *clumping*.

8.5.3 Análise Semântica Latente

O *CiteSpace* fornece uma função de análise semântica latente um pouco subdesenvolvida sob **Text ► Latent Semantic Analysis**. A análise semântica latente é baseada numa decomposição de valor singular do termo pelo documento matriz. Trata-se de um método de redução dimensional (Deerwester, Dumais, Landauer, Furnas, & Harshman, 1990).

Use o botão *Browse* para localizar pelo menos duas fontes de dados, por exemplo, pastas de arquivos de texto em texto simples completo ou o formato WoS. Após selecionar cada fonte de dados, adicione-o à lista por meio do botão "Add to the List", em seguida, pressione o botão "Analyze". Aguarde até terminar...

Uma vez concluído, cinco palavras mais representativas em cada dimensão são mostradas na *interface*.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>



Figura 137. Indexação Semântica Latente.

Três visualizações inadequadas do espaço semântico latente são fornecidas para as três mais importantes dimensões do espaço semântico latente. Cada visualização mostra uma mistura de termos e documentos. É possível aumentar e diminuir o *zoom*, mudar o tamanho da fonte dos rótulos, e o comprimento do rótulo. É isto. Esta função tem estado lá há anos, mas ela não foi desenvolvida de forma ativa.

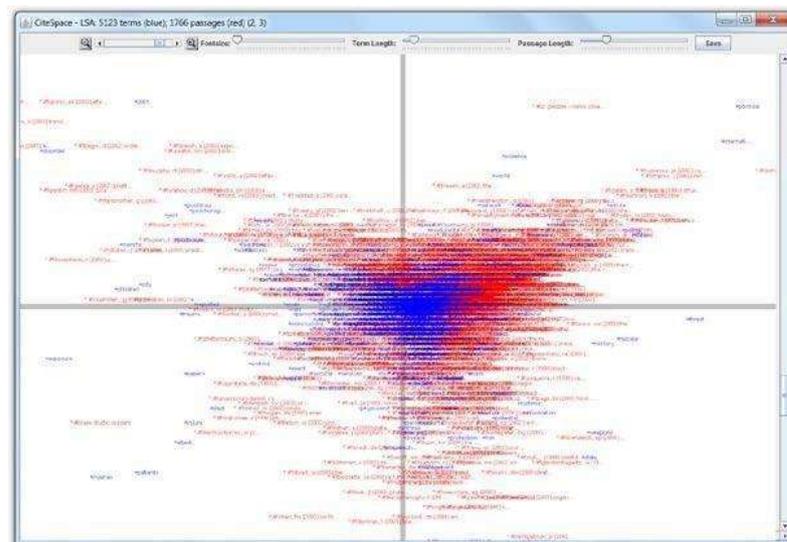


Figura 138. Uma visualização do espaço semântico latente, a 2ª e a 3ª dimensões.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

9 Exemplos Seleccionados

Aqui estão alguns bons exemplos observados na Internet. Estes exemplos foram criados por usuários do *CiteSpace*.

A visualização seguinte é do *blog.sciencenet.cn* de Jie Li.

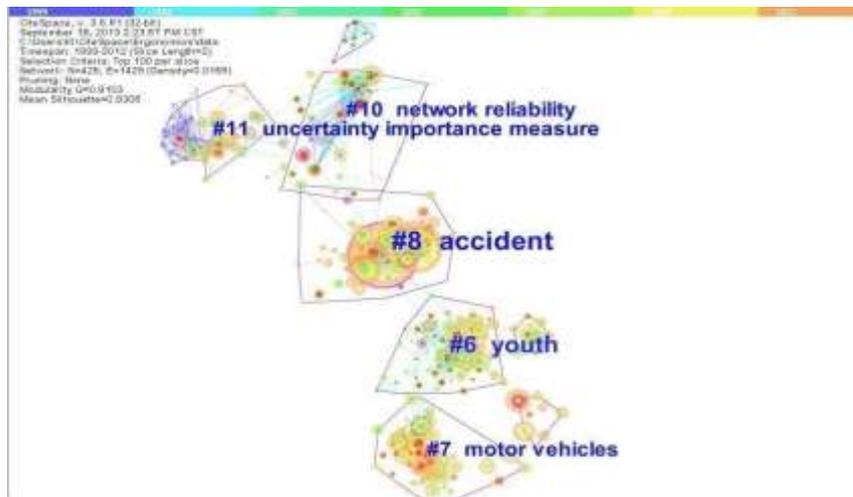


Figura 139. <http://blog.sciencenet.cn/blog-554179-729837.html>
<http://image.sciencenet.cn/album/201310/03/2133022yo2o5g7ppygpp5i.jpg>

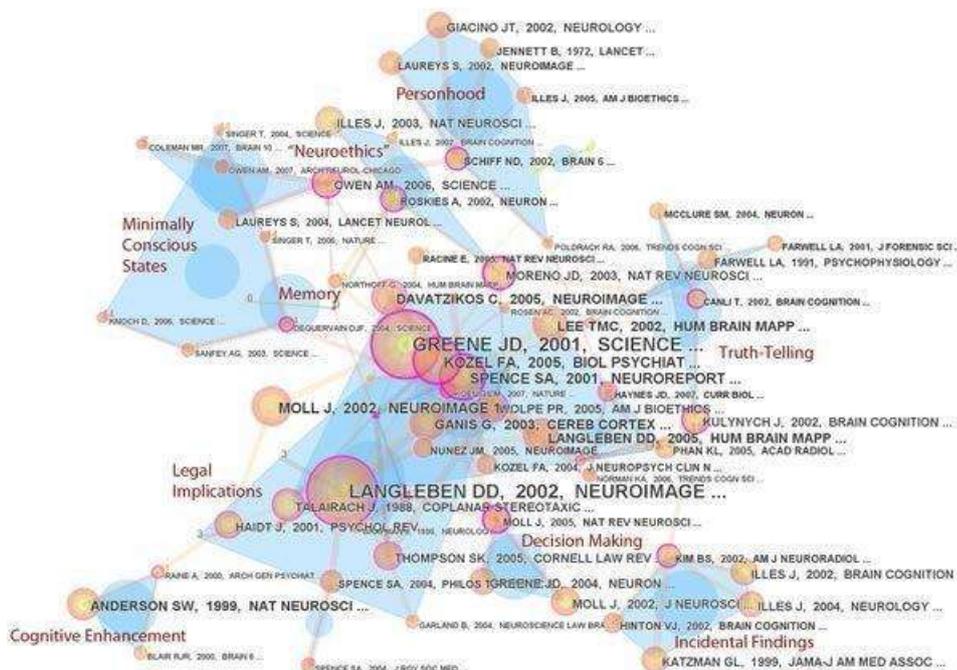


Figura 140. Um artigo publicado na PlosOne, Figura 3.
<http://www.plosone.org/article/info%3Adoi%2F10.1371%2Fjournal.pone.0018537>

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

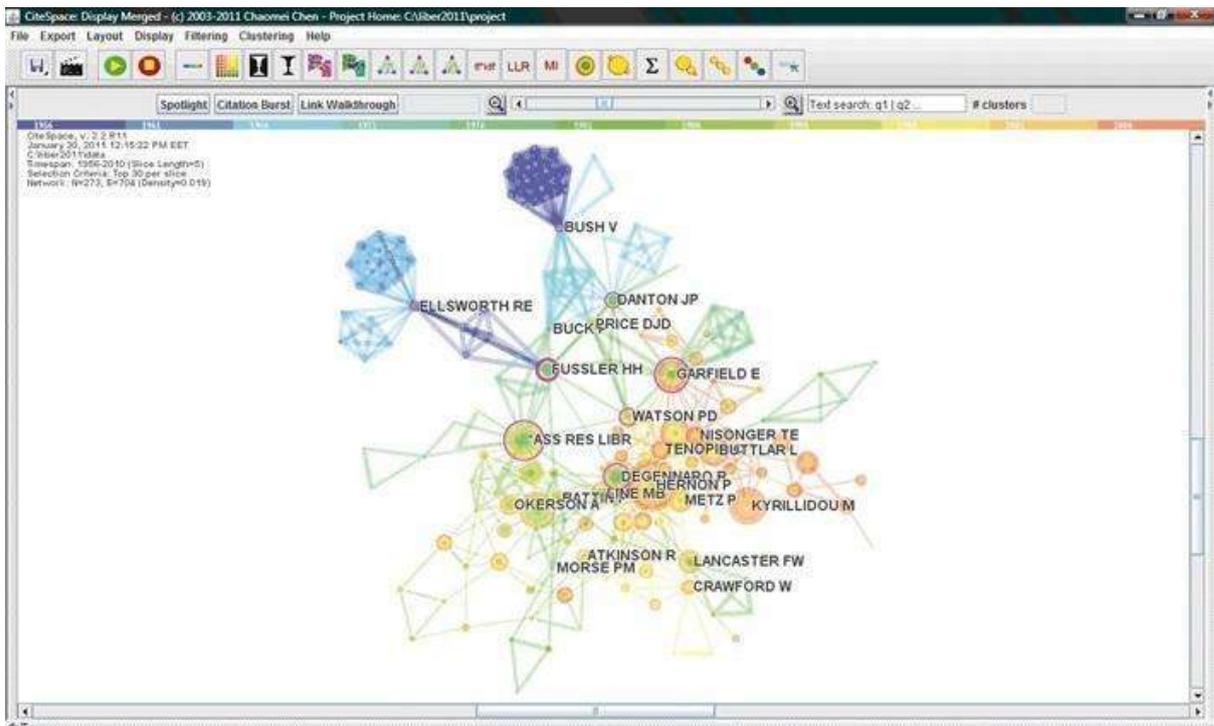


Figura 141. Um artigo publicado na *Liber Quarterly*, Figura 4.

<https://liber.library.uu.nl/index.php/lq/article/view/URN%3ANBN%3ANL%3AUI%3A10-1-113638/8398>

10 MÉTRICAS E INDICADORES

10.1 *Informações Teóricas*

10.1.1 Entropia Informacional

O *CiteSpace* processa a entropia informacional baseada em frases substantivas extraídas de registros para representar a diversidade de um conjunto de dados.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

10.2 Estrutural

10.2.1 Centralidade de Intermediação

A centralidade de intermediação de um nó em uma rede mede o grau em que o nó é parte dos caminhos que conectam um par arbitrário de nós na rede (Brandes, 2001; C. M. Chen, 2006; Freeman, 1977).

10.2.2 Modularidade

A modularidade de uma rede mede a extensão em que uma rede pode ser decomposta em vários componentes, ou módulos. Esta métrica fornece uma referência da clareza geral de uma determinada decomposição da rede (C. Chen et al., 2010).

A modularidade muda a taxa induzida por um conjunto de informações recebidas sendo considerado um sinal de uma perturbação para um complexo sistema adaptativo (C. Chen, 2012).

10.2.3 Silhouette (silhueta)

O valor de uma silhueta de um *cluster* mede a qualidade de configuração de *cluster*. O seu valor varia entre -1 e 1. O valor mais alto representa uma solução perfeita. No entanto, para garantir uma boa interpretação no *CiteSpace* é recomendado equilibrar as pontuações da modularidade e da silhueta simultaneamente (C. Chen et al., 2010).

10.3 Temporal

10.3.1 *Burstness* (explosão)

O *burstness* da frequência de uma entidade ao longo do tempo indica uma duração específica na qual uma mudança abrupta da frequência acontece (Kleinberg, 2002). No *CiteSpace*, explosão de citação e explosão de ocorrência ambas são suportadas.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

10.4 Combinado

10.4.1 Sigma

Este indicador mede a força combinada de propriedades estruturais e temporais de um nó, isto é, a sua centralidade de intermediação e a explosão de citação (C. Chen et al., 2009).

10.5 Rotulagens do Cluster

10.5.1 Frequência de Termo por frequência Invertida de Documento

10.5.2 Log Relação Probabilidade

10.5.3 Informação mútua

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

11 REFERÊNCIAS

- Bookstein, A., Klein, S. T., & Raita, T. (1998). Clumping properties of content-bearing words. *Journal of the American Society for Information Science*, 49(2), 102-114.
- Brandes, U. (2001). A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2), 163-177.
- Carrot2: Open source framework for building search clustering engines. (2012). from <http://project.carrot2.org/>
- Chen, C. (2004). Searching for intellectual turning points: Progressive Knowledge Domain Visualization. *Proc. Natl. Acad. Sci. USA*, 101(Suppl.), 5303-5310.
- Chen, C. (2012). Predictive effects of structural variation on citation counts. *Journal of the American Society for Information Science and Technology*, 63(3), 431-449. doi: 10.1002/asi.21694
- Chen, C., Chen, Y., Horowitz, M., Hou, H., Liu, Z., & Pellegrino, D. (2009). Towards an explanatory and computational theory of scientific discovery. *Journal of Informetrics*, 3(3), 191-209.
- Chen, C., Hu, Z., Liu, S., & Tseng, H. (2012). Emerging trends in regenerative medicine: A scientometric analysis in CiteSpace. *Expert Opinions on Biological Therapy*, 12(5), 593-608.
- Chen, C., Ibekwe-SanJuan, F., & Hou, J. (2010). The structure and dynamics of co-citation clusters: A multiple-perspective co-citation analysis. *Journal of the American Society for Information Science and Technology*, 61(7), 1386-1409. doi: 10.1002/asi.21309
- Chen, C., & Leydesdorff, L. (2014). Patterns of connections and movements in dual-map overlays: A new method of publication portfolio analysis. *Journal of the Association for Information Science and Technology*, 65(2), 334-351. doi: 10.1002/asi.22968
- Chen, C., & Morris, S. (2003, October 19-24, 2003). *Visualizing evolving networks: Minimum spanning trees versus Pathfinder networks*. Paper presented at the IEEE Symposium on Information Visualization, Seattle, Washington.
- Chen, C. M. (2006). CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. *Journal of the American Society for Information Science and Technology*, 57(3), 359-377. doi: 10.1002/asi.20317
- Deerwester, S., Dumais, S. T., Landauer, T. K., Furnas, G. W., & Harshman, R. A. (1990). Indexing by Latent Semantic Analysis. *Journal of the American Society for Information Science*, 41(6), 391-407.
- Freeman, L. C. (1977). A set of measuring centrality based on betweenness. *Sociometry*, 40, 35-41.
- Kessler, M. M. (1963). Bibliographic coupling between scientific papers. *American Documentation*, 14, 10-25.
- Kleinberg, J. (2002). *Bursty and hierarchical structure in streams*. Paper presented at the Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Edmonton, Alberta, Canada. <http://www.cs.cornell.edu/home/kleinber/bhs.pdf>
- Small, H. (1973). Co-citation in the scientific literature: A new measure of the relationship between two documents. *Journal of the American Society for Information Science*, 24, 265-269.

Chen, C. (2014) o *CiteSpace* Manual. <http://cluster.cis.drexel.edu/~cchen/citespace/>

Stasko, J., Gorg, C., & Liu, Z. (2008). Jigsaw: Supporting investigative analysis through interactive visualization. *Information Visualization*, 7(2), 118-132.

White, H. D., & Griffith, B. C. (1981). AUTHOR COCITATION - A LITERATURE MEASURE OF INTELLECTUAL STRUCTURE. *Journal of the American Society for Information Science*, 32(3), 163-171.