

Análise da co-citação entre *webs* de instituições de ensino superior politécnico

VÍCTOR F. HERRERO-SOLANA
SUSANA MARTINS

PALAVRAS-CHAVE

WEBMETRIA

ANÁLISE DE CO-CITAÇÃO

MAPAS DO CONHECIMENTO

ESCALONAMENTO MULTIDIMENSIONAL

ENSINO SUPERIOR POLITÉCNICO

RESUMO

Este trabalho pretende representar a co-citação existente entre as *webs* de algumas instituições de ensino superior politécnico, portuguesas e estrangeiras, determinar a sua relação, representatividade, distribuição e distância existente entre si, utilizando, para isso, a representação gráfica. Para se atingir o objectivo final foi utilizada a técnica do escalonamento multidimensional (MDS) tendo sido criado um mapa bidimensional representativo dos resultados obtidos.

ABSTRACT

This work intends to represent the existing co-citation between webs of some polytechnical higher education institutions, Portuguese and foreign, to determine their relation, representation, distribution and existing distance among them, using, for this purpose, the graphical representation. To achieve their aim the authors used the technique of the multidimensional scaling (MDS) having designed a representative bidimensional map of the obtained results.

É indubitável a importância que a Internet e o seu crescimento espectacular têm assumido. Os serviços oferecidos, tanto comerciais como não comerciais, constituem, actualmente, uma parte substancial da vida de muitos cidadãos. O aumento das capacidades da rede trouxe um desenvolvimento das aplicações informáticas associadas à World Wide Web, de tal forma que a que hoje conhecemos pouco se parece com a que surgiu há mais de dez anos.

Outro dos efeitos deste crescimento foi a criação de uma linha de investigação que estuda a dimensão quantitativa da rede. Durante muito tempo, o desenvolvimento desta linha esteve ligado a questões específicas de fluxo de informação na rede, estatísticas de visitas a sítios Web e questões semelhantes. Não obstante, nos últimos tempos começaram-se a usar técnicas próprias da bibliometria como são o desenho e a construção de mapas. A construção de mapas e outras representações gráficas, com o fim de poder visualizar estruturas de relações sumamente complexas, é uma tarefa científica que se vem realizando desde há duas décadas. Estas investigações floresceram particularmente no âmbito dos mapas da ciência, permitindo-nos ver as relações entre os investigadores de determinado campo temático. Existe uma infinidade de trabalhos científicos que apresentam mapas da ciência, mas são particularmente conhecidos os realizados por WHITE e MCCAIN da Universidade de Drexel, tanto para representar relações de autores (MCCAIN 1990; WHITE 1998), como de revistas (MCCAIN 1991). Posteriormente, outros autores tornaram estes primeiros mapas mais sofisticados, com outras metodologias e dotando-os de muita mais informação e interactividade (CHEN 1998, 1999, 2001; NOYONS 1998; DING et. al. 2000; MOYA-ANEGÓN & HERRERO-SOLANA 2002).

Será legítimo partir do princípio que a contagem de *backlinks* a páginas de determinadas universidades reflecte a investigação desenvolvida por essas universidades? De acordo com THELWALL (2001), a resposta será negativa sempre que a contagem for feita de modo indiscriminado uma vez que as páginas *web* académicas servem uma grande variedade de objectivos. No entanto, é assumido que as instituições de ensino superior se apresentam como uma área de investigação promissora devido à maturidade da *web* académica sendo que a análise de *backlinks* poderá ser usada, ao nível macro, para comparar grupos de instituições.

No nosso caso, o mapa que desejamos construir, à diferença dos anteriores, trabalha exclusivamente com informação extraída da Internet, mas parecer-se-á com os anteriores no sentido que vamos utilizar na sua construção técnicas

contrastadas, como as que encontramos nos trabalhos já citados. A aplicação destas técnicas informétricas ao âmbito da Web deu lugar à *webometrics* ou *webmetria*, que podemos definir como a ciência que compreende a investigação de todas as comunicações em rede, usando para tal técnicas de medição de carácter quantitativo (ALMIND & INGWERSEN 1997). No entanto, nos últimos anos parece que a forma mais aceite é a *cybermetrics* ou *cibermetria* já que deu, inclusive, lugar à criação de uma revista científica específica do tema: *Cybermetrics* (<http://www.cindoc.csic.es/cybermetrics>). Duas das principais técnicas usadas na *webmetria* são a análise de citações de sítios Web, também denominada de “citações” (ROUSSEAU 1997), e a análise do factor de impacto Web (INGWERSEN 1998). Podemos encontrar uma revisão bastante exaustiva das investigações realizadas neste campo no trabalho de JUDIT BAR-ILAN (2001).

Para a análise das hiperligações entre universidades, THELWALL propõe a experimentação de três modelos:

MODELO	METODOLOGIA
<i>WEB PAGE / FILE RANGE</i>	O NÚMERO DAS PÁGINAS WEB HIPERLIGADAS AO SÍTIO WEB ALVO DEVE SER CONTABILIZADO. QUALQUER ALVO URL SÓ É CONTADO UMA VEZ E AS SUBSEQUENTES HIPERLIGAÇÕES PARA O MESMO URL DEVERÃO SER IGNORADAS. UMA PÁGINA WEB NESTE CONTEXTO É IDENTIFICADA COM A SUA URL. O CERNE DESTA DEFINIÇÃO ENCONTRA-SE NO URL E NÃO NA PÁGINA <i>PER SE</i> . NÃO É ASSUMIDO QUE A PÁGINA ALVO EXISTA NA REALIDADE AQUANDO DA REALIZAÇÃO DO ESTUDO. A RAZÃO PARA ESTA SITUAÇÃO É A SEGUINTE: A INTENÇÃO DE HIPERLIGAR É ENCARADA COMO MAIS IMPORTANTE DO QUE SABER SE EXISTE UM <i>TYP0</i> NA URL OU SE O ALVO DESAPARECEU.
<i>WEB DIRECTORY RANGE</i>	O NÚMERO DE DIRECTÓRIOS WEB HIPERLIGADOS AO SÍTIO WEB ALVO DEVERÁ SER CONTABILIZADO. QUALQUER DIRECTÓRIO ALVO SÓ DEVERÁ SER CONTABILIZADO UMA VEZ.
<i>WEB DOMAIN RANGE</i>	O NÚMERO DE DIFERENTES DOMÍNIOS WEB HIPERLIGADOS AO SÍTIO WEB ALVO DEVERÁ SER CONTABILIZADO. QUALQUER DOMÍNIO ALVO SÓ DEVE SER CONTABILIZADO UMA VEZ.

A contagem efectuada com base no directório URL apresenta-se como a melhor abordagem para a análise das hiperligações entre universidades em detrimento dos restantes modelos.

Se a contagem das hiperligações for executada de modo cuidado, esta poderá

ser correlacionada com a actividade científica institucional.

Segundo THELWALL (2002), o estudo quantitativo dos casos isolados e dos dados estatísticos, em particular a contagem de hiperligações entre pares de universidades, sugere que tanto os modelos de directório como o modelo de domínio são mais fiáveis do que o modelo de página Web, do ponto de vista da contabilização de hiperligações entre sítios Web, sendo o segundo, o mais fiável. Ainda de acordo com THELWALL, de modo a se desenvolver uma métrica significativa para a análise de sítios Web de universidades, deverá ser construída uma teoria da criação da hiperligação. Esta decorrerá dos resultados de vários tipos de actividades sendo que a investigação nessas áreas poderia ser espelhada por vários tópicos:

A contagem fiável das hiperligações de páginas Web depende de motores de busca. Vários resultados mostram que o AltaVista é a fonte mais fiável de dados relativamente às *webs* académicas.

O desenvolvimento e teste de metrias foi levado a cabo por INGWERSEN no que diz respeito ao FIW mas, uma versão do FIW externo relativo foi criada especialmente para sítios Web académicos usando, para tal, o número de equivalente a tempo integral do pessoal académico da instituição em vez da contagem de páginas como denominador de cálculo.

Identificar as razões inerentes à criação de hiperligações é um aspecto a considerar uma vez que os sítios Web académicos contêm informação e material criado com os mais variados objectivos.

A capacidade de lidar com anomalias é um último aspecto a considerar e deve ser tida em consideração. A publicação *online* tem como principal característica a liberdade que, neste contexto, se assume como a ausência de controlo de qualidade. Esta liberdade interfere directamente com a quantidade de hiperligações e páginas existentes. Assim, o espaço Web é altamente volúvel sendo alterado a cada minuto. As contagens de hiperligações sofrem as consequências desta situação.

O princípio da indexação de citações sempre incorporou a ideia da visualização da informação de forma gráfica (ver *historiograph* de Garfield). Como foi referido anteriormente, recentemente surgiu um novo ramo da LIS. Esse ramo dedica-se à visualização de domínios de conhecimento. Nesta área foram sendo desenvolvidas ferramentas de modo a automatizar o processo de criação de gráficos, explorar

as diferentes abordagens de visualização da informação e a navegar entre quantidades maciças de informação científica. As “paisagens” de citações permitem visualizar as diferentes especialidades científicas, navegar entre frentes de investigação e detectar tendências e áreas inovadoras.

A teoria das redes complexas encara a citação bibliométrica e as redes de co-autoria como exemplos de redes sociais aplicando-se os mesmos princípios para as redes baseadas em páginas Web hiperligadas. A análise de hiperligações compreende, agora, ligações entre unidades académicas de diferentes níveis de agregação (países, instituições de um só país, departamentos e mesmo de investigadores). A estrutura topológica das redes de hiperligações numa macroanálise, pode revelar o que é de facto visível da instituição ou campo científico, relativamente à rede.

No caso das universidades, a importância da geografia e a capacidade de usar a tecnologia Web são dois aspectos que estão patentes nas redes de hiperligações. Já uma análise a uma escala mais reduzida, revela que as estruturas de colaboração estão patentes nas hiperligações. A colheita de sítios Web entre grupos específicos em certos campos científicos poderá detectar uma “Webosfera” complexa, reflectindo padrões de comunicação formal e informal.

De acordo com SMITH (1999), os instrumentos base para se proceder à realização de estudos webométricos têm sido os motores de busca, pois permitem trabalhar com grandes volumes de informação. Estas ferramentas agilizam a quantificação e avaliação dos fluxos de intercâmbio de dados e de informação na Web e permitem, ainda, uma contabilização do número total de páginas num espaço Web² e as hiperligações existentes para tais espaços.

A webometria possibilita, entre outros, o estudo e análise das citações entre páginas Web (através das hiperligações existentes). O número de hiperligações que levam a um determinado sítio Web é usado como factor de hierarquização no momento de exibir os sítios Web recuperados por alguns motores de busca. Por isso acabam por ser encarados como um indicador de elevada importância de um sítio Web.

Tal como defende SMITH, é evidente que os diferentes motores de busca possibilitam a realização de estudos nacionais semelhantes àqueles realizados através das análises de citações em publicações impressas. Para indicar a citação entre *homepages* foi utilizado, pela primeira vez por AGUILLO em 1996, o seguinte neologismo: *sitation*, em português, “sitação”. Relativamente às “sitações”, é de notar que o indicador WIF (Web impact factor) / FIW

(factor de impacto na Web) assume cada vez maior preponderância e importância, suscitando um grande interesse nos cientistas e estudiosos da webometria.

Como foi anteriormente referido, foi INGWERSEN (1998) quem sugeriu, pela primeira vez, a criação de uma medição análoga à do IF (impact factor) / FI (factor de impacto) utilizado nas publicações periódicas. A utilização de uma medição semelhante iria possibilitar conhecer o grau de interesse de sítios e domínios Web.

Segundo este autor, o FIW implica a soma do número de hiperligações (internas e externas) contidos nas páginas Web, referentes a um determinado domínio, independentemente da sua índole, dividido pelo número de páginas encontradas nesse domínio em determinado momento. Ou seja, a contagem consiste no número de páginas “hiperligadas” e não no número de ligações. Assim, e de acordo com THELWALL, o FIW é o número de páginas que levam a determinado domínio, de país ou institucional, sobre o número de páginas nesse sítio Web ou área.

É, então, possível determinar que este indicador serve para medir e comparar a atracção e influência que podem alcançar diferentes espaços e sítios na Web, possibilitando aos estudiosos conhecer o grau de reconhecimento relativo que ostentam vários domínios em determinado período temporal.

Ainda de acordo com INGWERSEN, comparativamente às citações científicas de revistas, instituições ou indivíduos (que podem ser estáveis ou não), o número de ligações que remetem a um objecto em particular, dentro da Web, pode diminuir ou mesmo desaparecer. Tal situação deve-se à eventual reestruturação de páginas disponíveis na rede e que, em algum momento mudaram de localização ou estão indisponíveis. Sem qualquer dúvida que esta situação ensombra e inviabiliza o cálculo retrospectivo do FIW.

É hoje assumido por todos que a Internet é uma fonte variada e rica de recursos, que se encontram disponíveis a todo o tipo de utilizador. No entanto, existem algumas dificuldades para a realização de um estudo webométrico, de entre as quais se destaca a natureza dinâmica da Internet, como referido anteriormente, bem como a enorme massa de dados não padronizada que esta contém, a par de algumas inconsistências nos serviços que esta presta e nos motores de busca disponíveis.

Como é facilmente comprovado, há documentos que desaparecem ao mesmo tempo que novas páginas relevantes são constantemente adicionadas e que não são assumidas de imediato pelos motores de busca, o que dificulta o processo de análise dessas páginas na rede.

É ainda possível identificar outra dificuldade, referida por LYNCH (1997) e que deriva da estrutura instável das próprias páginas, nomeadamente as que trabalham com bases de dados interactivas e que, por esse motivo, se encontram em constante mutação.

METODOLOGIA

O presente trabalho pretende estabelecer a relação entre os sítios Web de algumas instituições de ensino superior politécnico portuguesas, espanholas, brasileiras, mexicanas, argentinas e dos Estados Unidos. Para tal, foram analisadas as “co-citações” dos sítios Web dessas instituições. Este estudo baseou-se na aplicação do MDS – multidimensional scalling, i. e., escalonamento multidimensional e na posterior criação de um mapa tridimensional representativo dos resultados obtidos.

Para uma positiva prossecução do trabalho, foi necessário proceder à escolha das instituições que iriam constituir a amostra, tendo a escolha sido baseada nos seguintes critérios:

- a origem das instituições,
- a sua tipologia (pertença ao ensino superior politécnico),
- a sua “dimensão” (preponderância assumida no ciberespaço).

Uma vez que os sistemas de ensino variam de país para país, nalguns casos apenas foi encontrada uma instituição correspondente aos requisitos de selecção supracitados. Por esse motivo, não foi possível utilizar o mesmo número de instituições para cada país.

Inicialmente a tabela continha as instituições seleccionadas para o estudo que perfaziam um total de 32 (Anexo 1).

Uma vez escolhida a amostra alvo do estudo, procedeu-se à contagem e posterior análise estatística das co-citações. No entanto, aquando da recolha de dados (contagem) verificou-se que teriam que ser eliminadas duas das 32 instituições seleccionadas uma vez que não serviam o propósito do estudo. Procedeu-se, então, à eliminação das seguintes instituições:

– Espanha: Centro Politécnico Superior (www.cpc.unizar.es) – o endereço URL deixou de existir;

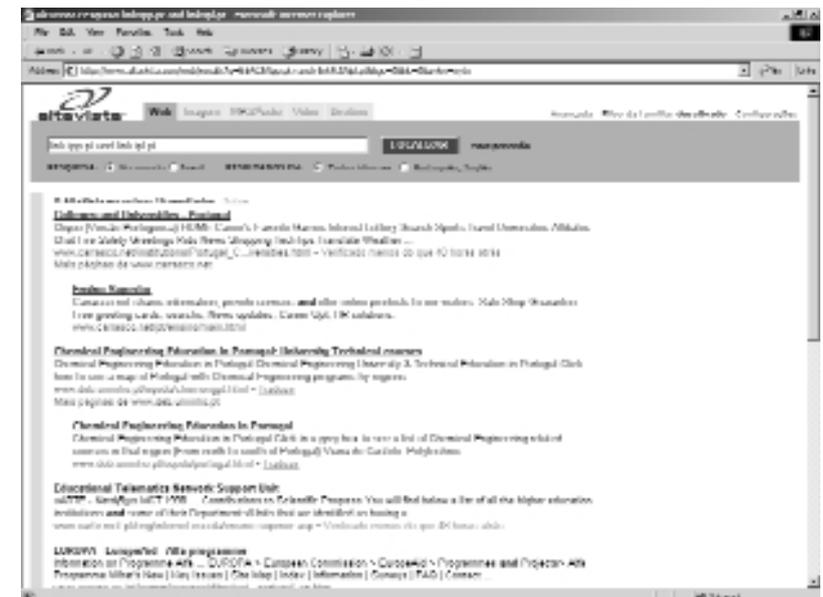
– México: Instituto Técnico Nacional (www.itn.mx) – o endereço URL deixou de existir;

Nesta fase, apenas constituem objecto de estudo, um total de 30 instituições.

Para a recolha de dados foi utilizado o motor de busca Altavista, através da busca por campos *link* e *host*; os valores utilizados foram os domínios das instituições objecto de estudo. De notar que o campo *link* quando associado a um valor de domínio, proporciona-nos uma lista de todas as páginas existentes na Internet que fazem referência a esse domínio em concreto. Já o campo *host* indica-nos o número de páginas sob determinado domínio. Todos os valores obtidos mediante esta fórmula permitem-nos representar o tamanho ou área da bolha na representação gráfica. Na realidade, os valores obtidos representam a visibilidade que o domínio em causa tem na Internet.

Para medir as co-citações foi realizada uma busca booleana com a seguinte fórmula de pesquisa:

Link:dominio1 and Link:dominio2

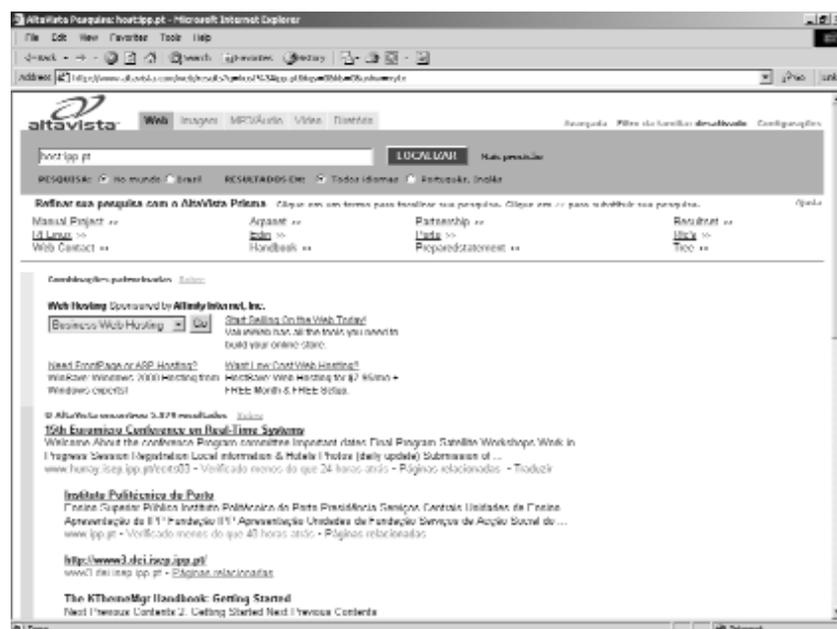


Ao efectuar-se uma busca com a seguinte expressão “**Link:ipp.pt and Link:ipl.pt**” o sistema pesquisa e devolve todas as páginas Web nas quais o Instituto Politécnico do Porto e o Instituto Politécnico de Lisboa apareçam co-citados – neste caso, 16. Aquando da realização das pesquisas, foram seleccionadas as seguintes opções: “Pesquisa: no mundo” e “Resultados: todos os idiomas”.

Ao realizar-se esta operação para todas as instituições seleccionadas, obtivemos dados para criar uma matriz de 30 x 30 valores absolutos denominada de “Matriz de Co-citação” (anexo 2).

Posteriormente foi utilizado o *software* SPSS 11.5 for Windows para conseguirmos obter o MDS – escalonamento multidimensional (anexo 3). Como resultado desta operação foram obtidos os valores **X** e **Y** que irão ser utilizados na fase final para a construção da representação gráfica

Para essa operação foi utilizado o *software* Microsoft Excel. Os valores **X** e **Y** foram introduzidos numa folha de cálculo Excel de modo a se que pudesse elaborar um gráfico de bolhas – deste modo conseguimos uma visualização imediata dos resultados obtidos. No entanto, antes de se passar à fase final, foi necessário acrescentarmos à tabela dos valores **X** e **Y** os valores relativos ao tamanho da bolha. Para esse efeito foram utilizados os valores obtidos através da pesquisa por campo host do Altavista.



HOST:IPP.PT

Ao efectuar-se uma busca com a expressão “**Host:ipp.pt**”, o sistema pesquisou e posteriormente devolveu todas as páginas Web sob este domínio – neste caso, 5879. Aquando da realização das pesquisas, foram seleccionadas as seguintes opções: “Pesquisa: no mundo” e “Resultados: todos os idiomas”.

Nesta altura era já possível aceder aos valores que nos permitiram construir o gráfico de bolhas em Excel (anexo 4).

Eis o resultado da pesquisa efectuada:

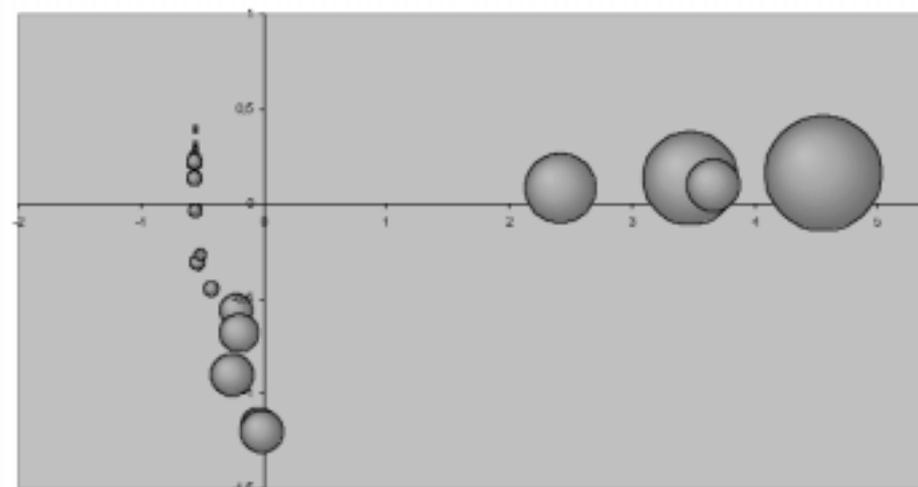


GRÁFICO MDS
“CO-CITAÇÃO ENTRE INSTITUIÇÕES DO ENSINO SUPERIOR”

É possível constatar que os resultados não são de fácil leitura, pois estamos perante instituições cuja visibilidade na Internet é bastante reduzida e outras cuja visibilidade é bastante elevada. Por esse facto, decidimos eliminar da representação gráfica todas as instituições cujo valor de representatividade (**T**) fosse inferior a 2000.

Nesse sentido, eliminámos as seguintes instituições:

- Instituto Tecnológico de Saltillo www.its.mx
- Instituto Politécnico de Viana do Castelo www.ipv.pt
- Instituto Politécnico da Universidade Católica de Minas www.ipuc.pucminas.br
- Instituto Politécnico Superior “General San Martins” www.ips.edu.ar

- Instituto Politécnico de Leiria www.ipl.pt
- Instituto Politécnico da Guarda www.ipg.pt
- Instituto Politécnico de Coimbra www.ipc.pt
- Universidade Politécnica de Cartagena www.upct.es
- Escola Politécnica da Universidade de São Paulo www.poli.usp.br
- Instituto Politécnico da Universidade do Estado do Rio de Janeiro www.iprj.uerj.br

Excluídas as instituições supracitadas, estaríamos perante uma tabela com 20 instituições (anexo 5).

Após a selecção, voltámos a criar um gráfico e eis o resultado:

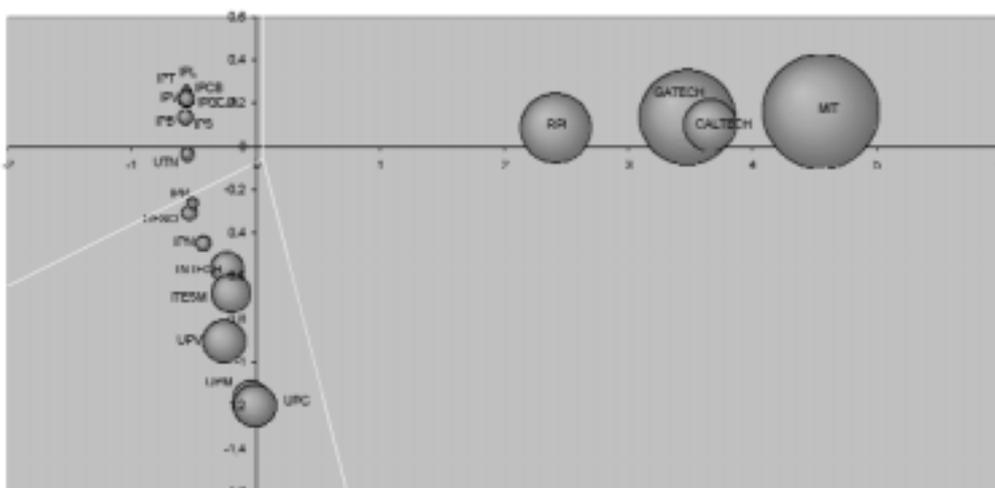


GRÁFICO MDS
"CO-CITAÇÃO ENTRE INSTITUIÇÕES DO ENSINO SUPERIOR POLITÉCNICO"

CONCLUSÕES

Conforme proposto na introdução do presente trabalho, tentou-se estabelecer a relação existente entre as instituições de ensino superior politécnico seleccionadas para o estudo. Após a construção da matriz, usando o MDS e construindo um gráfico de bolhas, foi-nos possível definir três agrupamentos: A, B e C.

O **agrupamento A** é completamente homogéneo, sendo composto por quatro instituições estado-unidenses: o RPI (Rensselaer Polytechnic Institute), o GATECH (Georgia Institute of Technology), o CALTECH (California Institute of Technology) e o MIT (Massachusetts Institute of Technology). De entre estas quatro instituições, o GATECH e o CALTECH encontram-se muito próximas, sendo que as restantes, apesar de constituírem o mesmo agrupamento, estão um pouco mais distanciadas. Este agrupamento ter-se-á definido pela proximidade geográfica e linguística, bem como pela proximidade nas áreas de ensino e de investigação.

O **agrupamento B** é, também ele homogéneo, sendo composto por oito instituições, sete das quais são portuguesas: o IPL (Instituto Politécnico de Leiria), IPT (Instituto Politécnico de Tomar), IPV (Instituto Politécnico de Viseu), IPB (Instituto Politécnico de Bragança), IPS (Instituto Politécnico de Setúbal), IPBeja (Instituto Politécnico de Beja), IPCB (Instituto Politécnico de Castelo Branco) e uma é argentina, a UTN (Universidade Tecnológica Nacional). Também aqui se pode explicar a distribuição e a proximidade das instituições que compõem este agrupamento pela sua origem e pela sua área de intervenção. Para o posicionamento do UTN neste grupo não foi encontrada qualquer explicação.

O **agrupamento C** é mais heterogéneo, sendo composto por uma instituição portuguesa, o IPP (Instituto Politécnico do Porto), três mexicanas, o ITESO (Instituto Tecnológico e de Estudos Superiores de Occidente), o IPN (Instituto Politécnico Nacional) e o ITESM (Instituto Tecnológico de Monterrey), uma estado-unidense, o TNTECH (Tennessee Tech University) e três espanholas, a UPC (Universidade Politécnica de Catalunya), a UPV (Universidade Politécnica de Valencia) e a UPM (Universidade Politécnica de Madrid). A presença do IPP neste agrupamento explica-se pelo facto de esta instituição apostar na internacionalização, distanciando-se, por isso, das suas congéneres portuguesas, e aproximando-se das suas congéneres estrangeiras, com as quais existe uma proximidade linguística e de áreas de estudo e pesquisa. Também neste grupo

encontramos uma instituição para a qual não foi possível encontrar qualquer explicação para o seu posicionamento, o TNTECH, que se encontra distante das suas congéneres estado-unidenses.

Relativamente às instituições para as quais não foi possível encontrar qualquer explicação para o seu posicionamento através do uso do MDS, seria aconselhável recorrer à metodologia da análise de redes sociais³ uma vez que este método poderia auxiliar na procura de uma explicação e interpretação para os resultados obtidos.

NOTAS

- ¹ *links* que ligam páginas Web à página que está a ser estudada.
- ² Espaço Web é entendido como domínio de país ou domínio institucional.
- ³ A análise de redes pode ser aplicada ao estudo de diferentes situações e questões sociais. Esta estabelece um novo paradigma na pesquisa direccionada para a estrutura social. É usada para estudar o modo como os comportamentos ou as opiniões dos indivíduos dependem das estruturas nas quais se inserem, a unidade de análise não são os atributos individuais (classe, sexo, idade, género), mas o conjunto de relações que os indivíduos estabelecem através das suas interações uns com os outros. A estrutura é apreendida concretamente como uma rede de relações e de limitações que pesa sobre as escolhas, as orientações, os comportamentos e as opiniões dos indivíduos. A análise de redes não constitui um fim em si mesmo. Ela é o meio usado para realizar uma análise estrutural cujo objectivo, por sua vez, é mostrar como a forma da rede é explicativa dos fenómenos analisados. O objectivo é demonstrar que a análise de uma diade (interacção entre duas pessoas) só tem sentido relativamente ao conjunto das outras diades da rede, porque a sua posição estrutural tem necessariamente um efeito sobre a sua forma, o seu conteúdo e a sua função. Portanto, a função de uma relação depende da posição estrutural dos elos, o mesmo ocorrendo com o *status* e o papel de um actor. Uma rede não se reduz a uma simples soma de relações, sendo que a sua forma exerce uma influência sobre cada relação.

BIBLIOGRAFIA

- ALMIND, T.; INGWERSEN, P. – “Informetric analyses on the World Wide Web: methodological approaches to ‘Webometrics’”. *Journal of Documentation*, 53, 404-426. 1997.
- AGUILLO, I.; GRANADINO, B.; LLAMAS, G. – “Posicionamiento en el Web del sector académico iberoamericano”. *Interciencia*, 30(12), 735-738. 2005.
- BAR-ILAN, J. – “Search engine results over the time”. *Cybermetrics* 2/3(1), 1-16. 1998.
- ____ – “Data collections methods on the Web for informetric purposes: a review and analysis”. *Scientometrics* 50(1), 7-32. 2001.
- BJÖRNERBON, L.; INGWERSEN, P. – “Perspectives of Webometrics”. *Scientometrics* 50(1), 65-82. 2001.
- CARMEN VÁSQUEZ, Maria del – *Análisis de cocitación de universidades españolas en relación con algunos casos de Francia y Reino Unido*. Universidade de Granada.
- CHEN, C. – “Bridging the gap: the use of pathfinder networks in visual navigation”. *Journal of Visual Languages and Computing*, 9, 267-286. 1998.
- ____ – “Visualising semantic spaces and author co-citation networks in digital libraries”. *Information Processing & Management*, 35, 401-420. 1999.
- ____ – “Visualizing a knowledge domain’s intellectual structure”. *IEEE Computer*, 34(3), 65-71. 2001.
- DING, Y.; CHOWDHURY, G.; FOO, S. – “Journal as markers of intellectual space: journal co-citation analysis of information retrieval area, 1987-1997”. *Scientometrics*, 47(1), 55-73. 2000.
- GUERRERO-BOTE, V.; MOYA-ANEGÓN, F.; HERRERO-SOLANA, V. – “Document organization using Kohonen’s algorithm”. *Information Processing & Management*, 38(1), 79-89. 2002.
- GUERRERO-BOTE, V.; REYES-BARRAGÁN, M.; MOYA-ANEGÓN, F.; HERRERO-SOLANA, V. – “Method for the análisis of the uses of scientific information: the case of the University of Extremadura (1996-7)”. *Libri: International Journal of Libraries and Information Services*, 52(2), 99-109. 2002.

- HERRERO SOLANA, Victor – *A representação gráfica da informação como meio de acesso*. 30 Universidade de Granada. 2001.
- INGWERSEN, P. – “The calculation of Web impact factors”. *Journal of Documentation*, 54(2), 236-243. 1998.
- INTERNETLAB. *World Universities ranking on the Web*. [disponível em <http://www.webometrics.info/index.html>] 2006.
- LARSON, R. R. – “Bibliometric analysis of the World Wide Web: an exploratory analysis of the intellectual structure of cyberspace”. 59TH ANNUAL MEETING OF ASIS: *Global complexity: information, chaos and control*. 1996.
- LYNCH, C. – “Searching the Internet: combining the skills of the librarian and the computer scientist may help organize the anarchy of the Internet”. *Scientific American*, 276(3), 44-48.
- MARTELETO, Regina Maria – “Análise de redes sociais – aplicação nos estudos de transferência da informação”. *Ciência da Informação* 30(1), 71-81. 2001.
- McCAIN, K. W. (1990): “Mapping authors in intellectual space: a technical overview”. *Journal of the American Society for Information Science* (JASIS), 41(6), 433-443.
- ___ – “Mapping economics through the journal literature: an experimental in journal cocitation analysis”. *Journal of the American Society for Information Science* (JASIS), 42(4), 290-296. 1991.
- MOYA-ANEGÓN, F.; HERRERO-SOLANA, V. – “Visibilidad internacional de la producción científica iberoamericana en bibliotecología y documentación (1991-2000)”. *Ciência da Informação*, 31(3), 54-65. 2002.
- MUÑOZ FERNÁNDEZ, Francisco José – *Aplicación del MDS a un caso concreto de cocitación de Webs universitarias*. Universidade de Granada.
- NOYONS, E. C. M. – “Mapping scientometrics, informetrics, and bibliometrics: bibliometric mapping as a policy support tool”. In *STI Conference*.
- PELAÉZ JIMÉNEZ, Monica – *Universidades españolas: análisis de cocitación Web*. Universidade de Granada. 2002.
- ROUSSEAU, R. – “Sitations: an exploratory study”. *Cybermetrics*, (1)1, 1-7. 1997.
- SCHARNRST, A.; THELWALL, Mike – “Citation and hyperlink networks”. *Current Science*, 89(9), 1518-1523. 2005.
- SMITH, A. – “A tale two Web spaces: comparing sites using Web impact factors”. *Journal of Documentation*, 55(5), 577-592. 1999.
- THELWALL, Mike – “Extracting macroscopic information from Web links”. *Journal of the American Society for Information Science and Technology*, 52(13), 1157-1168. 2001.
- ___ – “A research and institutional size based model for national university Web site interlinking”. *Journal of Documentation*, 58(6), 683-694. 2002.
- ___ – “Conceptualizing documentation on the Web: an evaluation of different heuristic-based models for counting links between university Web sites”. *Journal of the American Society for Information Science and Technology*, 53(12), 995-1005. 2002.
- ___ – “Methods for reporting on the targets of links from national systems of university Web sites”. *Information Processing & Management*, (40)1, 125-144. 2004.
- WHITE, H. – “Pathfinder networks and author cocitation analysis: a remapping of paradigmatic information scientists”. *Journal of the American Society for Information Science and Technology*, 54(5), 423-434. 2003.
- WHITE, H. D.; McCAIN, K. W. – “Visualizing a discipline: an author co-citation analysis of information science, 1972-1995”. *Journal of the American Society for Information Science* (JASIS), 49(4), 327-355. 1998.
- WHITE, H.; BUZYDŁOWSKI, J.; LIN, X. – “Co-cited author maps as interfaces to digital libraries: designing pathfinder networks in the humanities”. Fourth International Conference on Information Visualization. 2000.

ANEXO 1 INSTITUIÇÕES SELECIONADAS PARA O ESTUDO

País		INSTITUIÇÃO	URL
PORTUGAL	1	INSTITUTO POLITÉCNICO DO PORTO	WWW.IPP.PT
	2	INSTITUTO POLITÉCNICO DE LISBOA	WWW.IPL.PT
	3	INSTITUTO POLITÉCNICO DE BRAGANÇA	WWW.IPB.PT
	4	INSTITUTO POLITÉCNICO DE VISEU	WWW.IPV.PT
	5	INSTITUTO POLITÉCNICO DE SETÚBAL	WWW.IPS.PT
	6	INSTITUTO POLITÉCNICO DA GUARDA	WWW.IPG.PT
	7	INSTITUTO POLITÉCNICO DE LEIRIA	WWW.IPLEIRIA.PT
	8	INSTITUTO POLITÉCNICO DE COIMBRA	WWW.IPC.PT
	9	INSTITUTO POLITÉCNICO DE TOMAR	WWW.IPT.PT
	10	INSTITUTO POLITÉCNICO DE CASTELO BRANCO	WWW.IPCB.PT
	11	INSTITUTO POLITÉCNICO DE BEJA	WWW.IPBEJA.PT
	12	INSTITUTO POLITÉCNICO DE VIANA DO CASTELO	WWW.IPVC.PT
ESPAÑA	13	UNIVERSIDADE POLITÉCNICA DE MADRID	WWW.UPM.ES
	14	UNIVERSIDADE POLITÉCNICA DE CATALUNHA	WWW.UPC.ES
	15	UNIVERSIDADE POLITÉCNICA DE VALÈNCIA	WWW.UPV.ES
	16	UNIVERSIDADE POLITÉCNICA DE CARTAGENA	WWW.UPCT.ES
	17	CENTRO POLITÉCNICO SUPERIOR	WWW.CPS.UNIZAR.ES
BRASIL	18	ESCOLA POLITÉCNICA DA UNIVERSIDADE DE SÃO PAULO	WWW.POLL.USP.BR
	19	INSTITUTO POLITÉCNICO DA UNIVERSIDADE ESTATAL DO RIO DE JANEIRO	WWW.IPRJ.UERJ.BR
	20	INSTITUTO POLITÉCNICO DA UNIVERSIDADE CATÓLICA	WWW.IPUC.PUCMINAS.BR
MÉXICO	21	INSTITUTO TECNOLÓGICO DE MONTERREY	WWW.ITESM.MX
	22	INSTITUTO POLITÉCNICO NACIONAL	WWW.IPN.MX
	23	INSTITUTO TECNOLÓGICO DE SALTILLO	WWW.ITS.MX
	24	INSTITUTO TECNOLÓGICO E DE ESTUDOS SUPERIORES DE OCIDENTE	WWW.ITESO.MX
	25	INSTITUTO TECNOLÓGICO DE NOGALES	WWW.ITN.MX
ARGENTINA	26	UNIVERSIDADE TECNOLÓGICA NACIONAL	WWW.UTN.EDU.AR
	27	INSTITUTO POLITÉCNICO SUPERIOR "GENERAL SAN MARTÍN"	WWW.IPS.EDU.AR
EUA	28	MASSACHUSETTS INSTITUTE OF TECHNOLOGY	WWW.MIT.EDU
	29	CALIFORNIA INSTITUTE OF TECHNOLOGY	WWW.CALTECH.EDU
	30	TENNESSEE TECH UNIVERSITY	WWW.TNTECH.EDU
	31	GEORGIA INSTITUTE OF TECHNOLOGY	WWW.GATECH.EDU
	32	RENSSELEAR POLYTECHNIC INSTITUTE	WWW.RPI.EDU

ANEXO 2 "MATRIZ DE CO-CITAÇÃO"

IPP.PT	IPL.PT	IPB.PT	IPV.PT	IPS.PT	IPG.PT	IPLEIRIA	IPC.PT	IPT.PT	IPCB.PT	IPBEJA.P	IPVC.PT	UPM.ES	UPC.ES	UPV.ES	UPCT.ES	POLL.USP	IPRJ.UERJ	IPUC.PUC	ITESM.M	IPN.MX	ITS.MX	ITESO.MX	IPS.ED	UTN.ED	IPS.ED	.MIT.ED	CALTEC	TNTECH	GATECH	RPI.EDU
IPP.PT	53546	16	22	24	16	9	3	10	13	12	10	9	43	42	48	1	0	0	0	46	39	0	35	1	0	215	41	3	76	16
IPL.PT	16	53546	16	12	0	0	1	9	10	9	7	7	4	6	4	0	0	0	0	0	1	0	0	1	0	12	4	3	5	2
IPB.PT	22	16	53546	23	15	8	3	10	12	16	20	7	8	12	8	3	0	0	0	7	1	0	0	0	0	12	2	0	6	2
IPV.PT	24	12	23	53546	14	13	1	10	11	14	14	7	6	1	1	0	0	0	0	6	0	0	0	0	0	16	1	2	5	5
IPS.PT	16	0	15	14	53546	6	1	8	9	9	10	4	8	8	11	1	0	0	0	3	0	0	0	0	0	11	3	0	5	4
IPG.PT	9	0	8	13	6	53546	1	8	8	9	8	3	2	1	1	0	0	0	0	0	0	0	0	0	0	7	0	0	5	5
IPLEIRIA.PT	3	1	3	1	1	1	53546	0	1	2	1	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
IPC.PT	10	9	10	10	8	8	0	53546	6	9	6	5	0	0	7	1	0	0	0	0	0	0	0	0	0	1	0	0	1	1
IPT.PT	13	10	12	11	9	8	1	6	53546	9	9	4	1	3	2	1	0	0	0	1	1	0	0	0	0	4	0	0	2	1
IPCB.PT	12	9	16	14	9	9	2	9	9	53546	12	8	1	1	5	1	0	0	0	1	0	0	0	0	0	2	0	0	1	1
IPBEJA.PT	10	7	20	14	10	8	1	6	9	12	53546	4	6	1	4	0	0	0	0	3	0	0	1	0	0	12	2	0	8	3
IPVC.PT	9	7	7	7	4	3	1	5	4	8	4	53546	0	0	2	0	0	0	0	1	0	0	0	0	0	1	0	0	0	0
UPM.ES	43	4	8	6	8	2	0	0	1	1	6	0	53546	1601	1418	107	12	2	0	405	89	3	41	26	1	2336	553	15	682	409
UPC.ES	42	6	12	1	8	1	0	0	3	1	1	0	1601	53546	1372	97	6	2	0	326	113	2	23	20	0	2231	752	47	926	481
UPV.ES	48	4	8	1	11	1	0	7	2	5	4	2	1418	1372	53546	95	4	1	0	311	46	2	23	14	1	1248	328	23	413	221
UPCT.ES	1	0	3	0	1	0	0	1	1	1	0	0	107	97	95	53546	0	0	0	11	2	0	1	1	0	12	4	0	7	1
POLL.USP.BR	0	0	0	0	0	0	0	0	0	0	0	0	12	6	4	0	53546	0	0	5	5	0	0	0	0	28	5	3	19	8
IPRJ.UERJ.BR	0	0	0	0	0	0	0	0	0	0	0	0	2	2	1	0	0	53546	0	3	0	0	0	0	0	10	3	0	5	4
IPUC.PUCMI	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	53546	0	0	0	0	0	0	0	0	0	0	0
ITESM.MX	46	0	7	6	3	0	0	0	1	1	3	1	405	326	311	11	5	3	0	53546	525	22	346	23	0	1659	522	35	587	357
IPN.MX	39	1	1	0	0	0	0	0	1	0	0	0	89	113	46	2	5	0	0	525	53546	23	185	11	0	222	40	16	93	40
ITS.MX	0	0	0	0	0	0	0	0	0	0	0	0	3	2	2	0	0	0	0	22	22	53546	12	3	0	1	0	0	0	0
ITESO.MX	35	0	0	0	0	0	0	0	0	0	1	0	41	23	23	1	0	0	0	346	185	12	53546	2	0	126	33	7	42	19
UTN.EDU.AR	1	1	0	0	0	0	0	0	0	0	0	0	26	20	14	1	0	0	0	23	11	3	2	53546	0	42	13	2	13	11
IPS.EDU.AR	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	0	0	0	0	0	0	53546	2	0	0	0	0
MIT.EDU	215	12	12	16	11	7	0	1	4	2	12	1	2336	2231	1248	12	28	10	0	1659	222	1	126	42	2	53546	26773	1179	23871	12867
CALTECH	41	4	2	1	3	0	0	0	0	0	2	0	553	752	328	4	5	3	0	522	40	0	33	13	0	26773	53546	475	10242	5055
TNTECH.EDU	3	3	0	2	0	0	0	0	0	0	0	0	15	47	23	0	3	0	0	35	16	0	7	2	0	1179	475	53546	692	513
GATECH.ED	76	5	6	5	5	5	0	1	2	1	8	0	682	926	413	7	19	5	0	587	93	0	42	13	0	23871	10242	692	53546	5430
RPI.EDU	16	2	2	5	4	5	0	1	1	1	3	0	409	481	221	1	8	4	0	357	40	0	19	11	0	12867	5055	513	5430	53546

ANEXO 3 TABELA DE ESCALONAMENTO MUTIDIMENSIONAL

CONFIGURATION DERIVED IN 2 DIMENSIONS

STIMULUS COORDINATES

DIMENSION

STIMULUS NUMBER	STIMULUS NAME	1	2
1	VAR00001	-,5624	-,2143
2	VAR00002	-,6016	,2618
3	VAR00003	-,6046	,1688
4	VAR00004	-,6002	,2845
5	VAR00005	-,6043	,1863
6	VAR00006	-,5978	,3378
7	VAR00007	-,5962	,3884
8	VAR00008	-,5991	,3357
9	VAR00009	-,6009	,3019
10	VAR00010	-,6004	,3130
11	VAR00011	-,6015	,2682
12	VAR00012	-,5977	,3593
13	VAR00013	-,1021	-1,0510
14	VAR00014	-,0779	-1,0806
15	VAR00015	-,3241	-,7940
16	VAR00016	-,5790	-,3055
17	VAR00017	-,5999	,2362
18	VAR00018	-,5977	,3302
19	VAR00019	-,5945	,4075
20	VAR00020	-,2432	-,5889
21	VAR00021	-,4700	-,4036
22	VAR00022	-,6039	,2215
23	VAR00023	-,5860	-,2742
24	VAR00024	-,6002	-,0201
25	VAR00025	-,5947	,3934
26	VAR00026	4,3292	,1286
27	VAR00027	3,4793	,0873
28	VAR00028	-,2689	-,4577
29	VAR00029	3,2979	,1074
30	VAR00030	2,3023	,0721

ANEXO 4 TABELA DE VALORES X, Y, T

INSTITUIÇÃO	X	Y	T
WWW.UTN.EDU.AR	-0,5599	-0,0321	9584
WWW.UPV.ES	-0,2612	-0,8969	93509
WWW.UPM.ES	-0,0433	-1,1595	67036
WWW.UPCT.ES	-0,5349	-0,3348	1566
WWW.UPC.ES	-0,0166	-1,1965	98913
WWW.TNTECH.EDU	-0,2338	-0,5567	58228
WWW.RPL.EDU	2,4119	0,0893	261414
WWW.POLL.USP.BR	-0,5625	0,1996	1113
WWW.MIT.EDU	4,5501	0,1635	718870
WWW.ITS.MX	-0,5687	0,1807	37
WWW.ITESO.MX	-0,542	-0,3038	13149
WWW.ITESM.MX	-0,2073	-0,6744	76797
WWW.IPVC.PT	-0,5644	0,3273	736
WWW.IPV.PT	-0,5661	0,2452	5577
WWW.IPUC.PUCMINAS.BR	-0,5605	0,4076	219
WWW.IPT.PT	-0,5669	0,2607	3238
WWW.IPS.PT	-0,5685	0,1452	5355
WWW.IPS.EDU.AR	-0,562	0,358	20
WWW.IPRJ.UERJ.BR	-0,5645	0,2874	1668
WWW.IPP.PT	-0,5181	-0,2616	5890
WWW.IPN.MX	-0,432	-0,4449	11176
WWW.IPLEIRIA.PT	-0,5622	0,3846	533
WWW.IPL.PT	-0,5671	0,2205	10440
WWW.IPG.PT	-0,5647	0,2952	986
WWW.IPCB.PT	-0,5667	0,2699	2618
WWW.IPC.PT	-0,5652	0,3078	248
WWW.IPBEJA.PT	-0,5672	0,2261	11101
WWW.IPB.PT	-0,5685	0,1384	11738
WWW.GATECH.EDU	3,4677	0,1367	482628
WWW.CALTECH.EDU	3,6531	0,1026	149453

ANEXO 5 TABELA DE VALORES X, Y, T

INSTITUIÇÃO	X	Y	T
WWW.UTN.EDU.AR	-0,5599	-0,0321	9584
WWW.UPV.ES	-0,2612	-0,8969	93509
WWW.UPM.ES	-0,0433	-1,1595	67036
WWW.UPC.ES	-0,0166	-1,1965	98913
WWW.TNTECH.EDU	-0,2338	-0,5567	58228
WWW.RPL.EDU	2,4119	0,0893	261414
WWW.MIT.EDU	4,5501	0,1635	718870
WWW.ITESO.MX	-0,542	-0,3038	13149
WWW.ITESM.MX	-0,2073	-0,6744	76797
WWW.IPV.PT	-0,5661	0,2452	5577
WWW.IPT.PT	-0,5669	0,2607	3238
WWW.IPS.PT	-0,5685	0,1452	5355
WWW.IPP.PT	-0,5181	-0,2616	5890
WWW.IPN.MX	-0,432	-0,4449	11176
WWW.IPL.PT	-0,5671	0,2205	10440
WWW.IPCB.PT	-0,5667	0,2699	2618
WWW.IPBEJA.PT	-0,5672	0,2261	11101
WWW.IPB.PT	-0,5685	0,1384	11738
WWW.GATECH.EDU	3,4677	0,1367	482628
WWW.CALTECH.EDU	3,6531	0,1026	149453