



Universidade do Minho
Escola de Ciências

Ciência



A INTERNET, O GOOGLE E A MATEMÁTICA

CIÊNCIA | RUI RALHA *

Sir Timothy John Berners-Lee, físico e cientista da computação, é o criador da “World Wide Web” que costumamos designar por Internet (embora, em rigor, a Web seja apenas um dos serviços que funcionam dentro da Internet). Em 1989, Berners-Lee propôs um projeto de hipertexto (texto com ligações, as “hyperlinks” que todos usamos quando navegamos na Internet) que permitiria aos investigadores trabalhar em conjunto, partilhando o seu conhecimento numa rede de documentos ligados. Nos finais de 1990, com a colaboração de Robert Cailliau, líder do grupo de sistemas de computação do CERN, Berners-Lee implementou a primeira comunicação bem sucedida entre dois computadores, o servidor e o cliente, usando o sistema HTTP (HyperText Transfer Protocol) de transferência de hipertexto. O primeiro “site” da Internet foi construído no CERN, tinha uma única página e ficou disponível em 1991. Três anos depois, Tim Berners-Lee fundou o consórcio W3C (World Wide Web Consortium) composto de várias empresas com o intuito de criar normas que ajudassem a desenvolver a Internet numa base tecnológica livre de quaisquer direitos de autor. Por outras palavras, o criador da Internet não patenteou a sua criação. Tal espírito altruísta nunca será demasiadamente realçado.

Assim, a Web cresceu rapidamente nos últimos anos do século XX e com o avolumar da informação disponível os utilizadores passaram a ter dificuldade em encontrar, neste enorme “livro” que é a Web, as páginas relevantes para o assunto que lhes interessava em cada momento. Para ajudar nesta tarefa, criaram-se os chamados motores de busca.

Nos dias de hoje, para muitos, fazer pesquisas na Web significa “ir ao Google” (embora existam muitos outros motores de busca) e os mais jovens internautas não conheceram a Internet antes do Google. Com efeito, foi por volta do ano 2000 que o motor de busca da Google se impôs à concorrência. A inovação tecnológica do Google consistiu essencialmente na utilização de um algoritmo matemático para



O Rank (importância) de uma página depende do rank das páginas que para ela apontam

fazer a ordenação (ranking) das páginas Web, conhecido por PageRank.

É fácil perceber a tremenda importância que tem a ordenação dos resultados (páginas Web) obtidos numa pesquisa. É que, em geral, o número de páginas associadas a uma ou mais palavras de pesquisa é muito grande (como exemplo, o termo de pesquisa “Google” produziu no próprio Google três mil e vinte milhões de resultados – e amanhã talvez já produza mais) e um utilizador terá tempo e energia para consultar apenas algumas das páginas sendo que estas serão escolhidas certamente de entre as que aparecem nos lugares de topo da lista ordenada.

Larry Page, um dos dois fundadores e atual presidente da Google, licenciou-se em Engenharia de Computadores pela Universidade do Estado do Michigan e foi para a Universidade de Stanford, em 1997, para fazer um doutoramento em Ciências da Computação. Page já imaginava nesta altura a Web como um grafo, em constante mutação e crescimento, onde as páginas da Web são os vértices (ou nós) e as hyperlinks entre essas páginas são as arestas (ou arcos) do grafo. Page estava interessado em estudar as proprie-

dades desta estrutura matemática e este foi o seu tema de investigação de doutoramento. Em Stanford conheceu Sergei Brin, o outro fundador da Google, e também estudante de doutoramento em Ciências da Computação em Stanford, e juntos desenvolveram em 1998 um protótipo do Google.

Mas afinal como funciona o PageRank? Bem, a importância (rank) de uma página da Web é determinada essencialmente pela importância das páginas cujo texto tem hyperlinks que apontam para a primeira. Complicado? Nada que um simples sistema de equações lineares (um tema da unidade curricular de Álgebra Linear que é parte integrante do plano de estudos, do 1º ano geralmente, de qualquer Licenciatura em Ciências ou em Engenharias) não consiga resolver. O problema é que o número de equações deste sistema é igual ao número de páginas da Web que se estima atualmente em alguns biliões (isto é, milhões de milhões). Por isto, a atualização do Google requer quantidades gigantescas de dados e de cálculos.

* Departamento de Matemática e Aplicações da Escola de Ciências da Universidade do Minho

B.I.

Nome: Rui Ralha

Formação Académica: Licenciado em Matemática (FCTUC, 1981); Doutoramento na Universidade de Southampton (1990) e na Universidade de Coimbra, por equivalência, em 1991, na área de Análise Numérica e Computação.

Livro Favorito: o adjetivo favorito causa-me dificuldades; entre outros géneros, gosto de romances clássicos, estou a ler agora “O Senhor Embaixador” de Érico Veríssimo.

Filme Favorito: atualmente, todo o que me conseguir manter desperto depois das 23 horas.

Cidade Favorita: passei por belas cidades na Europa e na Américas mas na maior parte delas nunca estive tempo suficiente para criar uma relação de cumplicidade. Vou eleger Coimbra e Braga como as minhas favoritas.

Músico Favorito: gosto de clássicos (outra vez). A pergunta não é sobre as estações do ano mas quero dizer que gosto das quatro, de preferência pela ordem natural, como na composição de Vivaldi.

Especialidade Culinária: sardinhas assadas, ao ar livre.

Hobbies: um deles já foi jogar xadrez. Gosto de caminhadas, de preferência em contato com ambientes naturais. De há dois anos a esta parte, disfruto sempre que posso da companhia dos netos (gémeos). Também gosto de atividades de jardinagem.

Viagem de Sonho: A qualquer lugar onde nunca tenha ido e aos outros de onde guardo boas memórias (aqui destaco o “countryside” do sul de Inglaterra).

Se não fosse cientista seria... o professor que já sou.

Quer fazer perguntas a um cientista?

Esta rubrica sobre a Escola de Ciências da Universidade do Minho tem também como objectivo criar uma relação entre leitores e investigadores. Alguma vez pensou em fazer uma pergunta a um cientista? Caso queira participar pode enviar todas as suas questões para sec@ecum.uminho.pt e verá as suas dúvidas esclarecidas.