



República Federativa do Brasil
Ministério do Desenvolvimento, Indústria
e do Comércio Exterior
Instituto Nacional da Propriedade Industrial.

(21) PI0708397-1 A2



* B R P I 0 7 0 8 3 9 7 A 2 *

(22) Data de Depósito: 08/02/2007
(43) Data da Publicação: 31/05/2011
(RPI 2108)

(51) Int.CI.:
G06Q 10/00 2006.01
G06F 17/30 2006.01
G06F 17/40 2006.01

(54) Título: EXTRAÇÃO DE COMPORTAMENTO DE USUÁRIO DE BUSCA NA REDE PARA APERFEIÇOAR RELEVÂNCIA DE BUSCA NA REDE

(30) Prioridade Unionista: 14/07/2006 US 11/457,733, 02/03/2006 US 60/778,650, 02/03/2006 US 60/778,650, 14/07/2006 US 11/457,733

(73) Titular(es): Microsoft Corporation

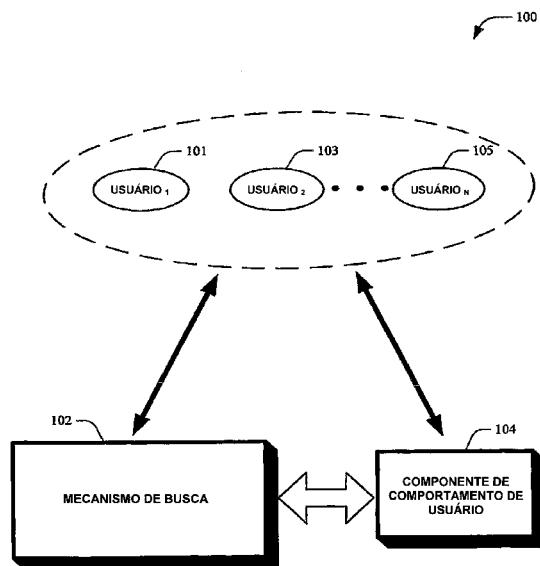
(72) Inventor(es): Eric D. Brill, Robert J. Ragno, Susan T. Dumais, Yevgeny E. Agichtein

(74) Procurador(es): Nellie Anne Daniel-Shores

(86) Pedido Internacional: PCT US2007003530 de 08/02/2007

(87) Publicação Internacional: WO 2007/106269de 20/09/2007

(57) Resumo: EXTRAÇÃO DE COMPORTAMENTO DE USUÁRIO DE BUSCA NA REDE PARA APERFEIÇOAR RELEVÂNCIA DE BUSCA NA REDE Sistemas e métodos que estimam preferência de usuário via interpretação automática de comportamento de usuário. Um componente de comportamento de usuário associado com um mecanismo de busca pode automaticamente interpretar comportamento coletivo de usuários (por exemplo, usuários de busca na rede). Tal componente de retorno pode incluir características de comportamento do usuário e modelos preditivos (por exemplo, a partir de um componente de comportamento de usuário) que são robustos a ruído, que podem estar presentes em interações de usuário observadas com os resultados da busca (por exemplo, atividade de usuário maliciosa e/ou irracional).



PI0708397-1

"EXTRAÇÃO DE COMPORTAMENTO DE USUÁRIO DE BUSCA NA REDE PARA APERFEIÇOAR RELEVÂNCIA DE BUSCA NA REDE"

Fundamentos da Invenção

- Dada a popularidade da Rede Mundial e da Internet, os usuários podem adquirir informação relacionada a quase qualquer tópico de uma grande quantidade de fontes de informação. De modo a encontrar informação, os usuários geralmente aplicam vários mecanismos de busca à tarefa de retorno de informação. Os mecanismos de busca permitem que usuários encontrem páginas da Rede contendo informação ou outro material na Internet que contém palavras ou frases específicas.
- Em geral, uma busca de palavras chave pode encontrar, para o melhor da capacidade de um computador, todos os sítios da Rede que têm qualquer informação neles relacionada a quaisquer palavras chave e frases que são especificadas. Um sítio de mecanismo de busca terá uma caixa para usuários inserirem palavras chave e um botão para pressionar para iniciar a busca. Muitos mecanismos de busca têm dicas sobre como usar palavras chave para buscar efetivamente. Tipicamente, tais dicas auxiliam os usuários a definir estreitamente termos de busca, tal que informação estranha e não relacionada não é retornada e o processo de retorno de informação não é desordenado. Tal estreitamento manual de termos pode suavizar o recebimento de vários milhares de sítios para classificar quando procurando por informação específica.
- Em alguns casos, tópicos de busca são pré-arranjados em áreas de tópico e sub-tópico. Por exemplo, o "Yahoo" fornece uma lista pré-determinada hierarquicamente arranjada de tópicos possíveis (por exemplo, negócios, governo, ciência, etc.) onde o usuário selecionará um tópico e então adicionalmente escolherá um sub-tópico na lista. Um outro exemplo de listas pré-determinadas de tópicos é comum em utilitários de ajuda de computadores pessoais de mesa, onde uma lista de tópicos de ajuda e sub-tópicos relacionados é fornecida ao usuário. Enquanto essas hierarquias pré-determinadas podem ser úteis em alguns contextos, os usuários frequentemente necessitam buscar/exigir informação fora e/ou não incluída nessas listas pré-determinadas. Assim, os mecanismos de busca ou outros sistemas de busca são frequentemente empregados para habilitar que os usuários direcionem consultas, para encontrar informação desejada. Contudo, durante as buscas de usuário, muitos resultados não relacionados são retornados, desde que os usuários podem estar incertos de como autorar ou construir uma consulta particular. Além disso, tais sistemas usualmente exigem que os usuários modifiquem continuamente consultas, e refinem resultados de busca retornados para obter um número razoável de resultados para examinar.
- Não é incomum digitar uma palavra ou frase em um campo de consulta de entrada em sistema de busca, e então retornar vários milhões de resultados como candidatos potenciais. Para compreender o grande número de candidatos retornados, o usuário frequente-

mente experimentará outras combinações de palavras, para adicionalmente estreitar a lista.

Em geral, o sistema de busca classificará os resultados de acordo com a relevância predita de resultados para a consulta. A classificação é tipicamente baseada em uma função que combina muitos parâmetros incluindo a similaridade de uma página da Rede com uma consulta bem como a qualidade intrínseca do documento, frequentemente inferida a partir de uma informação de topologia de Rede. A qualidade da experiência de busca do usuário está diretamente relacionada à qualidade da função de classificação, à medida que os usuários tipicamente não visualizam os resultados classificados mais abaixo.

Em geral, o sistema de busca tentará alcançar ou encontrar todos os tópicos relacionados à entrada de consulta do usuário sem considerar se os tópicos “buscados” têm qualquer relação contextual com a área ou categoria de tópico do que o usuário está realmente interessado. Como um exemplo, se um usuário que estava interessado em astronomia inserisse a consulta “Saturno” em um sistema de busca convencional, todos os tipos de resultados não relacionados são provavelmente retornados incluindo aqueles relacionados a carros, comerciantes de carros, jogos de computador, e outros sítios tendo a palavra “Saturno”. Um outro problema com as implementações de busca convencionais é que os mecanismos de busca operam da mesma forma para todos os usuários sem considerar as diferentes necessidades e circunstâncias dos usuários. Assim, se dois usuário inserem a mesma consulta de busca, eles tipicamente obtêm os mesmos resultados, sem considerar os seus interesses ou características, histórico de busca anterior, contexto de computação atual (por exemplo, arquivos abertos), ou contexto ambiental (por exemplo, localização, máquina sendo usada, hora do dia, dia da semana).

Sintonizando as funções de classificação de busca para retornar resultados relevantes no topo geralmente exige esforço significativo. Uma aproximação geral para modernos mecanismos de busca é treinar funções de classificação e ajustar parâmetros de função e pesos automaticamente baseado em exemplos de resultados de busca manualmente avaliados. Anotadores humanos podem explicitamente avaliar um conjunto de páginas para uma consulta de acordo com relevância percebida, e criar o “padrão de ouro” contra o qual diferentes algoritmos de classificação podem ser retornados e avaliados. Entretanto, classificações humanas explícitas são caras e difíceis de obter, frequentemente resultando em funções de classificação sub-ótimas e incompletamente treinadas.

Sumário da Invenção

O seguinte apresenta um sumário simplificado de modo a fornecer um entendimento básico de alguns aspectos do assunto reivindicado. Este sumário não é uma visão extensiva. Ele não pretende identificar elementos chave/críticos ou delinear o escopo do assunto reivindicado. Seu único propósito é apresentar alguns conceitos de uma forma simplificada como um prelúdio para a descrição mais detalhada que é apresentada posteriormente.

A inovação em questão aperfeiçoa as classificações de busca em um sistema de retorno de informação, via empregar um componente de comportamento de usuário que facilita uma interpretação automática para o comportamento coletivo de usuários, estimar as preferências de usuário para um item sobre um outro item. Tais preferências podem então ser empregadas para vários propósitos, tal como para aperfeiçoar a classificação dos resultados. O componente de comportamento de usuário pode interagir com um mecanismo(s) de busca e incluir características de retorno que suavizam ruído que tipicamente acompanham o comportamento do usuário (por exemplo, atividade de usuário maliciosa e/ou irracional). Explorando o comportamento agregado de usuários (por exemplo, não tratando cada usuário como um indivíduo profissional), a inovação em questão pode suavizar ruído e gerar julgamentos de relevância a partir do retorno de usuários. O componente de comportamento de usuário pode empregar retorno implícito ou explícito de usuários e suas interações com os resultados de consultas anteriores. As características comportamentais chave incluem características de apresentação que podem ajudar um usuário a determinar se um resultado é relevante olhando o título do resultado e descrição; navegando por características como tempo de permanência em uma página, maneira de alcançar resultados de busca (por exemplo, através de outros vínculos), desvio do tempo médio no domínio, e seus similares; características de clique, tal como o número de cliques em um resultado particular para a consulta. Para um dado par consulta-resultado, a inovação em questão fornece múltiplos valores de característica observados e derivados para cada tipo de característica.

O componente de comportamento de usuário pode empregar um modelo acionado por dados do comportamento do usuário. Por exemplo, o componente de comportamento de usuário pode modelar o comportamento de busca na rede do usuário como se fosse gerado por dois componentes: um componente de “fundo” (tal como usuários clicando indiscriminadamente), e um componente de “relevância” (tal como comportamento específico de consulta que é influenciado pela relevância do resultado para a consulta).

De acordo com um aspecto adicional da inovação em questão, o componente de comportamento de usuário pode gerar e/ou modelar os desvios a partir do comportamento de usuário esperado. Portanto, as características derivadas podem ser computadas, onde tais características derivadas explicitamente abordam o desvio do valor de característica observado para um dado resultado de busca a partir dos valores esperados para um resultado, sem informação dependente de consulta.

Além disso, o componente de comportamento de usuário da inovação em questão pode empregar modelos tendo dois tipos de característica para descrever o comportamento do usuário, ou seja: direto e desviado, onde o primeiro é os valores diretamente medidos, e o ultimo é o desvio dos valores esperados estimados a partir das distribuições gerais (independentes de consulta) para as características diretamente observadas correspondentes.

Conseqüentemente, o valor observado o de uma característica f para uma consulta q e resultado r , pode ser expresso como uma mistura de dois componentes:

$$o(q, r, f) = C(r, f) + rel(q, r, f)$$

onde $C(r, f)$ é a distribuição de “fundo” anterior para valores de f agregados através

- 5 de todas as consultas correspondentes a r , e $rel(q, r, f)$ é o componente “relevância” do comportamento influenciado pela relevância do resultado para a consulta. Por exemplo, uma estimativa de relevância do comportamento de usuário pode ser obtida com característica de clique, via uma subtração de distribuição de fundo a partir da freqüência de clique em uma dada posição. Para suavizar o efeito de variações de usuário individuais em comportamento, a inovação em questão pode tirar a média de valores de característica através de todos os usuários e sessões de busca para cada par consulta-resultado. Tal agregação pode fornecer robustez adicional, onde interações de usuário “ruidosas” adicionais não são consideradas.
- 10

- 15 Conseqüentemente, o comportamento de usuário para um par consulta-resultado pode ser representado por um vetor característica que inclui ambas as características diretamente observadas e os valores de características “corrigidos” derivados. Várias técnicas de aprendizagem de máquina podem também ser empregadas em conjunto com algoritmos de classificação de treinamento para sistemas de retorno de informação. Por exemplo, julgamentos de relevância humana explícitos podem ser inicialmente fornecidos para várias
- 20 consultas de busca e empregados para subseqüentes algoritmos de classificação de treinamento.

- 25 Em um aspecto relacionado, o comportamento coletivo de usuários interagindo com um mecanismo de busca na rede pode ser automaticamente interpretado de modo a prever futuras preferências de usuário; portanto, o sistema pode se adaptar para mudar padrões de comportamento de usuário e diferentes configurações de busca re-treinando automaticamente o sistema com os dados de comportamento de usuário mais recentes.

- 30 Para alcançar os fins anteriores e relacionados, certos aspectos ilustrativos do assunto reivindicado são descritos aqui em conjunto com a descrição seguinte e os desenhos em anexo. Esses aspectos são indicativos de várias formas nas quais o assunto pode ser praticado, todas as quais pretendem estar no escopo do assunto reivindicado. Outras vantagens e novas características podem se tornar aparentes a partir da seguinte descrição detalhada quando consideradas em conjunto com os desenhos.

Breve Descrição dos Desenhos

- A Fig. 1 ilustra um diagrama de bloco de um componente de comportamento de usuário de acordo com um aspecto exemplificado da inovação em questão.

A Fig. 2 ilustra um diagrama de bloco de um sistema que incorpora um componente de comportamento de usuário e interage com um modelo de treinamento de um mecanismo

de busca de acordo com um aspecto da inovação em questão.

A Fig. 3 ilustra um diagrama de bloco de um sistema que incorpora um componente classificador operativamente conectado a um componente de comportamento de usuário, e um mecanismo de busca de acordo com um aspecto exemplificado da inovação em questão.

A Fig. 4 ilustra uma tabela de características que representam atividades de navegação de usuário de acordo com um aspecto da inovação em questão.

A Fig. 5 ilustra um sistema de retorno de informação automatizado que pode empregar um componente de aprendizado de máquina com um aspecto da inovação em questão.

A Fig. 6 ilustra um componente de comportamento de usuário que interage com uma pluralidade de características de sistema, que representa ação do usuário de acordo com um aspecto particular da inovação em questão.

A Fig. 7 ilustra uma metodologia exemplificada de interpretar o comportamento do usuário para estimar preferências do usuário de acordo com um aspecto da inovação em questão.

A Fig. 8 ilustra uma metodologia de implementar o comportamento de usuário como parte de classificação de valor de acordo com um aspecto da inovação em questão.

A Fig. 9 ilustra um ambiente exemplificado para implementar vários aspectos da inovação em questão.

A Fig. 10 é um diagrama de bloco esquemático de um ambiente de computação adicional que pode ser empregado para implementar vários aspectos da inovação em questão.

Descrição Detalhada da Invenção

Os vários aspectos da inovação em questão são agora descritos com relação aos desenhos em anexo, onde números similares se referem a elementos similares ou correspondentes. Dever-se-ia entender, entretanto, que os desenhos e a descrição detalhada relacionada a esses não pretendem limitar o assunto reivindicado à forma particular descrita. De preferência, a intenção é abranger todas as modificações, equivalentes, e alternativas que caem no espírito e escopo do assunto reivindicado.

Como usado aqui, os termos “componente”, “sistema”, “característica” e seus similares pretendem também se referir a uma entidade relacionada a computador, ou hardware, uma combinação de hardware e software, software, ou software em execução. Por exemplo, um componente pode ser, mas não está limitado a ser, um processo executando em um processador, um processador, um objeto, um executável, um processo de execução, um programa, e/ou um computador. A título de ilustração, ambos uma aplicação executando em computador e o computador podem ser um componente. Um ou mais componentes podem

residir em um processo e/ou tarefa de execução e um componente pode estar localizado em um computador e/ou distribuído entre dois ou mais computadores.

A palavra “exemplificado” é usada aqui para significar servir como um exemplo, caso, ou ilustração. Qualquer aspecto ou projeto descrito aqui como “exemplificado” não é necessariamente construído como preferencial ou vantajoso sobre outros aspectos ou projetos.

Além disso, o assunto descrito pode ser implementado como um sistema, método, aparelho, ou artigo de fabricação usando programação padrão e/ou técnicas de engenharia para produzir software, suporte lógico inalterável, hardware, ou qualquer combinação desses para controlar um dispositivo baseado em computador ou processador para implementar aspectos detalhados aqui. O termo programa de computador como usado aqui pretende abranger um programa de computador acessível a partir de qualquer dispositivo legível por computador, portadora ou meios. Por exemplo, os meios legíveis por computador podem incluir, mas não estão limitados a dispositivos de armazenamento magnético (por exemplo, disco rígido, disco flexível, fitas magnéticas, ...), discos ópticos (por exemplo, disco compacto (CD), disco versátil digital (DVD), ...), cartões inteligentes, e dispositivos de memória rápida (por exemplo, cartão, módulo). Adicionalmente, dever-se-ia apreciar que uma onda portadora pode ser empregada para carregar dados eletrônicos legíveis por computador tal como aqueles usados na transmissão e recebimento de correio eletrônico ou no acesso a uma rede tal como a Internet ou uma rede de área local (LAN). É claro, aqueles versados na técnica reconhecerão que muitas modificações podem ser feitas a essa configuração sem abandonar o escopo ou espírito do assunto reivindicado.

Voltando inicialmente para a Fig. 1, um diagrama de bloco de um sistema 100 é ilustrado, o qual incorpora um componente de comportamento de usuário que interage com um mecanismo de busca de acordo com um aspecto exemplificado da inovação em questão. O componente de comportamento de usuário 104 associado com o mecanismo de busca 102 pode automaticamente interpretar o comportamento coletivo de usuários 101, 103, 105 (1 a N, onde N é um inteiro). Tal componente de comportamento de usuário 104 pode incluir características de retorno que suavizam ruído, que tipicamente acompanham o comportamento de usuário (por exemplo, atividade de usuário maliciosa e/ou irracional). Explorando o comportamento agregado dos usuários 101, 103, 105 (por exemplo, não tratando cada usuário como um indivíduo especialista), o sistema 100 pode suavizar ruído, e gerar julgamentos de relevância a partir de retorno de usuários.

O componente de comportamento de usuário 104 pode interagir com o componente de classificação. Para uma dada consulta, o componente de comportamento de usuário 104 retorna as previsões derivadas de um modelo de comportamento anteriormente treinado para essa consulta, e re-ordena os resultados para a consulta, tal que os resultados que

parecem relevantes para usuários anteriores são classificados mais altos. Por exemplo, para uma dada consulta q , a pontuação implícita IS_r pode ser computada para cada resultado r a partir das características de interação de usuário disponível, resultando na classificação implícita I_r , para cada resultado. Uma pontuação combinada $SM(r)$ pode ser computada para r combinando as classificações obtidas a partir do retorno implícito, I_r , com a classificação original de r , O_r :

$$SM(r, I_r, O_r, w_I) = \begin{cases} w_I \frac{1}{I_r + 1} + \frac{1}{O_r + 1}, & \text{se retorno implícito existe para } r \\ \frac{1}{O_r + 1}, & \text{de outra forma} \end{cases}$$

O peso w_I é um fator de escala heuristicamente sintonizado que representa a “importância” relativa do retorno implícito. Os resultados de consulta podem ser ordenados em valores decrescentes de $SM(r)$ para produzir a classificação final. Um caso particular de tal modelo aparece quando configurando w_I para um valor muito grande, efetivamente forçando resultados clicados como sendo classificados mais altos do que resultados não clicados – uma heurística efetiva e intuitiva que pode ser empregada como uma linha de base. Em geral, a aproximação acima assume que não há interações entre as características subjacentes produzindo a classificação de busca na rede original e as características de retorno implícitas. Outros aspectos da inovação em questão relaxam tal hipótese interagindo as características de retorno implícitas diretamente no processo de classificação, como descrito em detalhes a seguir. Além disso, aprecia-se que comportamento de usuário mais sofisticado e algoritmos de combinação de classificador podem ser empregados, e estão bem no âmbito da inovação em questão.

A Fig. 2 ilustra um aspecto adicional da inovação em questão, onde o mecanismo de busca 202 adiciona um modelo de treinamento 204 de acordo com um aspecto da inovação em questão. O modelo de treinamento 204 pode ademais compreender tipos de modelo adicional para descrever comportamento de usuário, ou seja: uma característica de comportamento observado 201 e uma característica de comportamento derivado 203. As características de comportamento observadas 201 são os valores diretamente medidos, e a característica de comportamento derivada 203 é o desvio dos valores esperados estimados a partir dos valores esperados estimados a partir das distribuições gerais (independentes de consulta) para as características diretamente observadas correspondentes. Conseqüentemente, o valor observado o de uma característica f para uma consulta q e resultado r , pode ser expresso como uma mistura de dois componentes:

$$o(q, r, f) = C(r, f) + rel(q, r, f)$$

onde $C(r, f)$ é a distribuição de “fundo” anterior para valores de f agregados através de todas as consultas correspondentes a r , e $rel(q, r, f)$ é o componente do comportamento

influenciado pela relevância dos resultados. Por exemplo, uma estimativa de relevância do comportamento de usuário pode ser obtida com característica de clique, via uma subtração de distribuição de fundo (por exemplo, ruído) da freqüência de clique observada em uma dada posição. Para suavizar o efeito de variações de usuário individuais em comportamento,

5 a inovação em questão pode tirar a média de valores de característica direta através de todos os usuários e sessões de busca para cada par consulta-URL. Tal agregação pode fornecer robustez adicional, onde interações de usuário “ruidosas” individuais não são consideradas. Conseqüentemente, o comportamento de usuário para um par consulta-URL pode ser representado por um vetor de característica que inclui ambas as características diretamente observadas e, valores de características “corrigidas” derivadas.

10

A Fig. 3 ilustra um diagrama de bloco de um sistema 300 que incorpora um componente classificador 310 operacionalmente conectado a um componente de comportamento de usuário 315 e um mecanismo de busca 340 de acordo com um aspecto exemplificado da inovação em questão. Tipicamente, o mecanismo de busca 340 pode classificar resultados de busca 350 baseado em um grande número de características, incluindo características baseadas em conteúdo (por exemplo, o quanto uma consulta se aproxima do texto ou título ou texto âncora do documento), e características de qualidade de página independente de consulta (por exemplo, Classificação da Página do documento ou do domínio), como descrito em detalhes a seguir. Além disso, o mecanismo de busca 340 pode empregar métodos automáticos (ou semi-automáticos) para sintonizar a função de classificação específica que combina tais valores de característica. Por exemplo, pode-se assumir que um usuário que faz uma consulta 360 executará ações particulares. Tais ações podem incluir clicar, navegar, enviar refinamentos de consulta até encontrar um documento relevante, e seus similares. Mediante encontrar o documento relevante, o usuário pode se tornar satisfeito e mudar o comportamento (por exemplo, ler o documento). A inovação em questão habilita descobrir um conjunto suficientemente rico de características que permitiriam a detecção de quando o usuário está satisfeito com um resultado retornado. Tais características são dependentes das consultas feitas, e, portanto, são específicos de consulta. Por exemplo, características/atividades de usuário podem ser categorizadas em características de apresentação, características de navegação, e características de clique, como descrito com relação à Fig. 4.

15

20

25

30

A Fig. 4 ilustra uma tabela de características 400 que representam atividades de navegação de usuário. As características de apresentação 410 são tipicamente projetadas para representar a experiência do usuário à medida que elas afetam alguns ou todos os aspectos do comportamento (por exemplo, um usuário pode decidir clicar em um resultado baseado nas características de apresentação). Para modelar tal aspecto de experiência de usuário, a inovação em questão pode empregar características tais como sobreposição em

35

palavras no título e palavras na consulta (Sobreposição de Título) e a fração de palavras compartilhadas pela consulta e o sumário de resultado, à medida que esses são freqüentemente considerados por usuários quando tomando uma decisão se clicar em um sumário de resultado para visualizar o documento completo.

5 Igualmente, a característica de navegação 420 pode capturar e quantificar aspectos das interações de página da rede de usuário. Por exemplo, a inovação em questão pode computar o desvio de tempo de permanência a partir do tempo de permanência da página esperada para uma consulta, o que permite modelar a diversidade intra-consulta de comportamento de navegação de página. Esse pode ademais incluir ambas características diretas e
 10 10 as derivadas, como descrito em detalhes a seguir. Igualmente, as características de clique 430 são um exemplo de interação de usuário com os resultados de mecanismo de busca. Por exemplo, as características de clique podem incluir o número de cliques para um par consulta-resultado, ou o desvio da probabilidade de clique esperada.

15 Como ilustrado na Fig. 4, o clique ilustra um aspecto de interações de usuário com um mecanismo de busca na rede. A inovação em questão pode empregar modelos de comportamento de usuário preditivos automaticamente derivados. Conseqüentemente, para uma dada consulta, cada resultado pode ser representado com as características na Tabela da Fig. 4. As preferências de usuário relativas podem então ser estimadas usando o modelo de comportamento de usuário aprendido, como descrito em detalhes acima. O uso de tais
 20 20 modelos de comportamento de usuário habilita o mecanismo de busca a se beneficiar da compreensão de multidões interagindo com os resultados da busca bem como características mais ricas caracterizando comportamento de navegação além da página de resultados de busca.

25 A Fig. 5 ilustra um sistema de retorno de informação automatizado 500 que pode empregar um componente de aprendizado de máquina 535 de acordo com um aspecto da inovação em questão. Uma estratégia de interpretação de retorno implícito geral pode ser empregada para automaticamente aprender um modelo de preferências de usuário (por exemplo, ao invés de contar com heurísticas ou intuições). O sistema 500 inclui um componente de classificação 510 que pode ser treinado a partir de um registro de dados 520 ou interações com o componente de comportamento de usuário 515, por exemplo. Os dados no registro 520 podem ser reunidos a partir de fontes de dados locais ou remotas e incluem informação relacionada aos dados de busca anteriores ou atividades 530 a partir de uma pluralidade de usuários. Depois de treinar, o componente classificador 510 pode interagir com o mecanismo de busca 540 para facilitar ou aperfeiçoar resultados de busca futuros
 30 30 que são indicados como resultados relevantes 550. Por exemplo, uma ou mais novas consultas de busca 560 podem ser processadas pelo mecanismo de busca 540, baseadas em parte no treinamento a partir de dados de busca anteriores 530, e/ou informação a partir do

componente de comportamento de usuário 515. Em geral, o sistema 500 pode empregar várias técnicas de mineração de dados para aperfeiçoar a relevância do mecanismo de busca. Tais técnicas podem incluir empregar classificadores de relevância no componente classificador 510, para gerar dados de treinamento de alta qualidade para classificadores de 5 tempo de execução, que são empregados com o mecanismo de busca 540 para gerar os resultados de busca 550. A Fig. 6 ilustra um componente de comportamento de usuário 610 que interage com uma pluralidade de características de sistema, que representam a ação do usuário. Em um aspecto, a inovação em questão considera os comportamentos de busca na rede como uma combinação de um componente de “fundo” (por exemplo, ruído independente de consulta e de relevância no comportamento do usuário, e seus similares), e um componente de “relevância” (por exemplo, comportamento específico de consulta indicativo da relevância de um resultado para uma consulta). Tal arranjo pode levar vantagem de comportamento de usuário agregado, onde o conjunto de características é compreendido de características diretamente observadas (computadas diretamente a partir de observações para 10 cada consulta), bem como características derivadas específicas de consulta, computadas como o desvio da distribuição independente de consulta total de valores para os valores de característica diretamente observados correspondentes. Como ilustrado na Fig. 6, características de sistema exemplificadas tais como: característica(s) de clique 612, características de navegação 614, e características de apresentação 616, que podem ser empregadas para 15 representar interações de usuário com resultados de busca na rede, através do componente de comportamento de usuário 610. Além disso, características tal como o desvio do número de cliques observados para um dado par consulta-URL a partir do número esperado de cliques em um resultado na dada posição, podem também ser consideradas. Além disso, o comportamento de navegação pode ser modelado, por exemplo, depois que um resultado é 20 clicado, então o tempo de permanência em página médio para um dado par consulta-URL dado, bem como seu desvio a partir do tempo de permanência (médio) esperado, é empregado para tal modelo. Adicionalmente, por exemplo, usuário de busca na rede podem frequentemente determinar se um resultado é relevante olhando no título do resultado, URL e sumário – em muitos casos, olhar no documento original é tipicamente não necessário. Para 25 modelar esse aspecto de experiência de usuário, características tais como: sobreposições em palavras no título e palavras na consulta, podem também ser empregadas.

A Fig. 7 ilustra uma metodologia exemplificada 700 de interpretar o comportamento do usuário para estimar as preferências do usuário de acordo com um aspecto da inovação em questão. Enquanto o método exemplificado é ilustrado e descrito aqui como uma série 30 de blocos representativos de vários eventos e/ou ações, a inovação em questão não está limitada pelo ordenamento ilustrado de tais blocos. Por exemplo, algumas ações ou eventos podem ocorrer em diferentes ordens e/ou ao mesmo tempo com outras ações ou eventos,

separado do ordenamento ilustrado aqui, de acordo com a inovação. Em adição, nem todos os blocos ilustrados, eventos ou ações, podem ser exigidos para implementar uma metodologia de acordo com a inovação em questão. Além disso, será apreciado que o método exemplificado e outros métodos de acordo com a inovação podem ser implementados em associação com o método ilustrado e descrito aqui, bem como em associação com outros sistemas e aparelhos não ilustrados ou descritos. Inicialmente e em 710, dados relacionados à interação de usuário com o mecanismo de busca, tal como comportamento de usuário pós busca podem ser adquiridos. Subseqüentemente e em 720, o comportamento do usuário pode ser agregado, por exemplo, empregando técnicas de análise estatística. Em 730, o aprendizado em máquina pode então ser empregado para treinar modelo de preferência de usuário. Subseqüentemente e em 740, as predições de preferência podem ser fornecidas para resultado de consultas futuras.

A Fig. 8 ilustra uma metodologia 800 de implementar o comportamento de usuário como parte da classificação de acordo com um aspecto da inovação em questão. Inicialmente e em 810, os dados relacionados ao comportamento do usuário podem ser coletados. Tal comportamento do usuário pode então ser empregado para treinar e/ou automaticamente gerar um modelo de comportamento em 820. Tal modelo (por exemplo, modelo de comportamento preditivo) pode então ser incorporado como parte de um mecanismo de busca para classificar resultados e/ou gerar julgamentos de relevância implícitos a partir do retorno de usuários, em 830. Subseqüentemente, e 840, baseado em parte na informação de modelo comportamental gerado e/ou treinado retornada pelo mecanismo de busca pode então ser classificado.

De modo a fornecer um contexto para os vários aspectos do assunto descrito, as Figs. 9 e 10, bem como a seguinte discussão pretendem fornecer uma breve descrição geral de um ambiente adequado no qual os vários aspectos do assunto descrito podem ser implementados. Enquanto o assunto foi descrito acima no contexto geral de instruções executáveis por computador de um programa de computador que executa em um computador e/ou computadores, aqueles versados na técnica reconhecerão que a inovação também pode ser implementada em combinação com outros módulos de programa. Geralmente, os módulos de programa incluem rotinas, programas, componentes, estruturas de dados, etc., que executam tarefas particulares e/ou implementam tipos de dados abstratos particulares. Além disso, aqueles versados na técnica apreciarão que os métodos inovadores podem ser praticados com outras configurações de sistema de computador, incluindo sistemas de computador de único processador ou de multiprocessador, dispositivos de mini-computação, computadores de grande porte, bem como computadores pessoais, dispositivos de computação portáteis (por exemplo, assistente digital pessoal (PDA), telefone, relógio, ...), eletrônicos industriais ou de consumo programáveis ou baseados em microprocessador, e seus

similar. Os aspectos ilustrados podem também ser praticados em ambientes de computação distribuídos onde tarefas são executadas por dispositivos de processamento remoto que são ligados através de uma rede de comunicações. Entretanto, alguns, de não todos os aspectos da inovação podem ser praticados em computadores autônomos. Em um ambiente

- 5 de computação distribuído, os módulos de programa podem ser localizados em ambos dispositivos de armazenamento em memória remoto ou local.

Com relação à Fig. 9, um ambiente exemplificado 910 para implementar vários aspectos da inovação em questão é descrito, o qual inclui um computador 912. O computador 912 inclui uma unidade de processamento 914, uma memória de sistema 916, e um barramento de sistema 918. O barramento de sistema 918 acopla componentes de sistema incluindo, mas não limitados a, memória de sistema 916 à unidade de processamento 914. A unidade de processamento 914 pode ser qualquer dos vários processadores disponíveis. Os microprocessadores duplos e outras arquiteturas de multiprocessador também podem ser empregados como a unidade de processamento 914.

15 O barramento de sistema 918 pode ser qualquer um dos vários tipos de estrutura(s) de barramento incluindo barramento de memória ou controlador de memória, um barramento periférico ou barramento externo, e/ou um barramento local usando qualquer variedade de arquiteturas de barramento disponíveis incluindo, mas não limitadas a, barramentos de 11 bits, Arquitetura Padrão de Indústria (ISA), Arquitetura de Micro Canal (MCA), ISA Estendido (EISA), Eletrônica de Açãoamento Inteligente (IDE), Barramento Local VESA (VLB), Componentes Periféricos Interconectados (PCI), Barramento Serial Universal (USB), Porta Gráfica Avançada (AGP), barramento de Associação Internacional de Placa de Memória de Computador Pessoal (PCMCIA), e Interface de Sistemas de Pequeno Computador (SCSI).

A memória de sistema 916 inclui memória volátil 920 e memória não volátil 922. O sistema de entrada/saída básico (BIOS), contendo as rotinas básicas para transferir informação entre elementos no computador 912, tal como durante a inicialização, é armazenado em memória não volátil 922. A título de ilustração, e não limitação, a memória não volátil 922 pode incluir memória somente de leitura (ROM), ROM programável (PROM), ROM eletricamente programável (EPROM), ROM eletricamente apagável (EEPROM), ou memória rápida.

30 A memória volátil 920 inclui memória de acesso aleatório (RAM), que age como memória cache externa. A título de ilustração e não limitação, a RAM está disponível em muitas formas, tal como RAM síncrona (SRAM), RAM dinâmica (DRAM), DRAM síncrona (SDRAM), SDRAM de taxa dupla de dados (DDR SDRAM), SDRAM aperfeiçoada (ESDRAM), DRAM de ligação síncrona (SLDRAM), e RAM Rambus direta (DRRAM).

35 O computador 912 também inclui meios de armazenamento em computador removíveis/não removíveis, voláteis/não voláteis. A Fig. 9 ilustra, por exemplo, um armazenamento em disco 924. O armazenamento em disco 924 inclui, mas não está limitado a, dispositi-

vos como uma unidade de disco magnético, unidade de disco flexível, unidade de fita, unidade de Jaz, unidade de Zip, unidade LS-60, cartão de memória rápida, ou módulo de memória. Em adição, o armazenamento em disco 924 pode incluir meios de armazenamento separadamente ou em combinação com outros meios de armazenamento incluindo, mas

5 não limitados a, uma unidade de disco óptico tal como dispositivo ROM de disco compacto (CD-ROM), unidade gravável de CD (Unidade de CD-R), unidade regravável de CD (Unidade de CD-RW) ou uma unidade ROM de disco versátil digital (DVD-ROM). Para facilitar a conexão dos dispositivos de armazenamento em disco 924 ao barramento de sistema 918, uma interface removível ou não removível é tipicamente usada tal como a interface 926.

10 Aprecia-se que a Fig. 9 descreve software que age como um intermediário entre usuários e os recursos básicos de computador descritos no ambiente operacional adequado 910. Tal software inclui um sistema operacional 928. O sistema operacional 928, que pode ser armazenado no armazenamento em disco 924, age para controlar e alocar recursos do sistema de computador 912. Os aplicativos de sistema 930 levam vantagem do gerencialmente de recursos pelo sistema operacional 928 através de módulos de programa 932 e dados de programa 934 armazenados ou na memória de sistema 916 ou no armazenamento em disco 924. Aprecia-se que vários componentes descritos aqui podem ser implementados com vários sistemas operacionais ou combinações de sistemas operacionais.

15 Um usuário insere comandos ou informação no computador 912 através do dispositivo(s) de entrada 936. Os dispositivos de entrada 936 incluem, mas não estão limitados a, um dispositivo de apontamento tal como um mouse, mouse estacionário, caneta gráfica, mesa sensível ao toque, teclado, microfone, comando de jogos, mesa de jogos, antena de satélite, digitalizador, placa sintonizadora de TV, câmera digital, câmera de vídeo digital, câmera de rede, e seus similares. Esses e outros dispositivos de entrada se conectam à unidade de processamento 914 através do barramento de sistema 918 via a porta(s) de interface 938. As portas de interface 938 incluem, por exemplo, uma porta serial, uma porta paralela, uma porta de jogos, e um barramento serial universal (USB). O dispositivo(s) de saída 940 usa algum do mesmo tipo de portas do dispositivo(s) de entrada 936. Assim, por exemplo, uma porta USB pode ser usada para fornecer entrada para o computador 912, e 20 para emitir informação a partir do computador 912 a um dispositivo de saída 940. O adaptador de saída 942 é fornecido para ilustrar que há alguns dispositivos de saída 940 como monitores, alto-falantes, e impressoras, dentre outros dispositivos de saída 940 que exigem adaptadores especiais. Os adaptadores de saída 942 incluem, a título de ilustração e não limitação, placas de vídeo e som que fornecem um dispositivo de conexão entre o dispositivo de saída 940 e o barramento de sistema 918. Dever-se-ia notar que outros dispositivos e/ou sistemas de dispositivos fornecem ambas as capacidades de entrada e saída tal como 25 o computador(es) remoto 944. O computador remoto 944 pode ser um computador pessoal, 30 35

um servidor, um roteador, um PC em rede, uma estação de trabalho, uma ferramenta baseada em microprocessador, um dispositivo par ou outro nó de rede comum e seus similares, e tipicamente inclui muitos ou todos os elementos descritos em relação ao computador 912. Para propósitos de brevidade, somente um dispositivo de armazenamento em memória 946 5 é ilustrado com o computador remoto 944. O computador remoto 944 é logicamente conectado ao computador 912 através de uma interface de rede 948 e então fisicamente conectado via a conexão de comunicação 950. A interface de rede 948 abrange redes de comunicação tais como redes de área local (LAN) e redes de área ampla (WAN). As tecnologias LAN 10 incluem Interface de Dados Distribuídos por Fibra Ótica (FDDI), Interface de Dados Distribuídos por Cabos de Par Trançado (CDDI), Ethernet/IEEE 802.3, Anel de Sinal/IEEE 802.5 e seus similares. As tecnologias WAN incluem, mas não são limitadas a, conexões ponto a ponto, redes de comutação de circuito como Redes Digitais de Serviços Integrados (ISDN) e variações nestas, redes de comutação de pacote, e Linhas de Assinante Digital (DSL).

A conexão de comunicação 950 refere-se ao hardware/software empregado para 15 conectar a interface de rede 948 ao barramento 918. Enquanto a conexão de comunicação 950 é mostrada para esclarecimento ilustrativo dentro do computador 912, ela pode ser externa ao computador 912. O hardware/software necessário para conexão à interface de rede 948 inclui, para propósitos exemplificados somente, tecnologias internas e externas tais como, modems incluindo modems de qualidade de telefone regular, modems a cabo e modems DSL, adaptadores ISDN, e placas Ethernet.

Como usado aqui, os termos “componente”, “sistema” e seus similares pretendem se referir a uma entidade relacionada a computador, ou hardware, uma combinação de hardware e software, software, ou software em execução. Por exemplo, um componente 25 pode ser, mas não é limitado a ser, um processo executando em um processador, um processador, um objeto, um executável, uma tarefa de execução, um programa, e/ou um computador. A título de ilustração, ambas uma aplicação executando no computador e o computador podem ser um componente. Um ou mais componentes podem residir em um processo e/ou tarefa de execução e um componente pode estar localizado em um computador e/ou distribuído entre dois ou mais computadores. A palavra “exemplificado” é usada aqui para 30 significar servir como um exemplo, caso, ou ilustração. Qualquer aspecto ou projeto descrito aqui como “exemplificado” não é necessariamente construído como preferencial ou vantajoso sobre outros aspectos ou projetos.

Além disso, o assunto descrito pode ser implementado como um sistema, método, aparelho ou artigo de fabricação usando técnicas de programação padrão e/ou de engenharia para produzir software, suporte lógico inalterável, hardware ou qualquer combinação desses para controlar um dispositivo baseado em computador ou processador para implementar aspectos detalhados aqui. O termo programa de computador como usado aqui pre-

tende abranger um programa de computador acessível a partir de qualquer dispositivo legível por computador, portadora ou meios. Por exemplo, os meios legíveis por computador podem incluir, mas não estão limitados a, dispositivos de armazenamento magnético (por exemplo, disco rígido, disco flexível, fitas magnéticas, ...), discos ópticos (por exemplo, disco compacto (CD), disco versátil digital (DVD), ...), cartões inteligentes, e dispositivos de memória rápida (por exemplo, cartão, módulo). Adicionalmente, dever-se-ia apreciar que uma onda portadora pode ser empregada para carregar dados eletrônicos legíveis por computador tal como aqueles usados na transmissão e recebimento de correio eletrônico ou no acesso a uma rede, tal como a Internet ou uma rede de área local (LAN). É claro, aqueles versados na técnica reconhecerão que muitas modificações podem ser feitas a essa configuração sem abandonar o escopo ou espírito do assunto reivindicado.

A Fig. 10 é um diagrama de bloco esquemático de um ambiente de computador de amostra 1000 que pode ser empregado para estimar preferência de usuário via o componente de comportamento de usuário de acordo com um aspecto da inovação em questão. O sistema 1000 inclui um ou mais cliente(s) 1010. O cliente(s) 1010 pode ser hardware e/ou software (por exemplo, tarefas, processos, dispositivos de computação). O sistema 1000 também inclui um ou mais servidor(es) 1030. O servidor(es) 1030 pode também ser hardware e/ou software (por exemplo, tarefas, processos, dispositivos de computação). Os servidores 1030 podem alojar tarefas para executar transformações empregando os componentes descritos aqui, por exemplo. Uma comunicação possível entre um cliente 1010 e um servidor 1030 pode ser na forma de um pacote de dados adaptado para ser transmitido entre dois ou mais processos de computador. O sistema 1000 inclui uma estrutura de comunicação 1050 que pode ser empregada para facilitar comunicações entre o cliente(s) 1010 e o servidor(es) 1030. Os clientes 1010 são operacionalmente conectados a um ou mais armazenedores de dados de cliente 1060 que podem ser empregados para armazenar informação local no cliente(s) 1010. Similarmente, os servidores 1030 são operacionalmente conectados a um ou mais armazenedores de dados de servidor 1040 que podem ser empregados para armazenar informação local nos servidores 1030.

O que foi descrito acima inclui vários aspectos exemplificados. Não é possível, é claro, descrever cada combinação concebível de componentes ou metodologias para propósitos de descrever esses aspectos, mas um versado na técnica pode reconhecer que muitas combinações adicionais e permutas são possíveis. Conseqüentemente, os aspectos descritos aqui pretendem abranger todas tais alterações, modificações, e variações que caem no espírito e escopo das reivindicações em anexo.

Além disso, no que diz respeito ao termo “inclui”, este é usado ou na descrição detalhada ou nas reivindicações, tal termo pretende ser inclusivo de uma maneira similar ao termo “compreende”, como “compreende” é interpretado quando empregado como uma pa-

lavra transitória em uma reivindicação.

REIVINDICAÇÕES

1. Sistema implementado por computador, **CARACTERIZADO** pelo fato de que compreende os seguintes componentes executáveis por computador:
 - um componente de comportamento de usuário (104, 35, 515, 610) que facilita a interpretação automática de comportamento coletivo de usuários (101, 103, 105) para estimar preferências de usuário de resultados de busca (350, 550) e
 - um mecanismo de busca (102, 202, 340, 540) que incorpora o comportamento coletivo para determinação de relevância e classificação de resultados de busca retornados (350, 550).
- 5 10 2. Sistema implementado por computador, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o componente de comportamento de usuário adicionalmente compreende um componente de fundo e um componente de relevância.
- 15 3. Sistema implementado por computador, de acordo com a reivindicação 1, **CARACTERIZADO** adicionalmente pelo fato de que compreende um componente de aprendizado de máquina.
4. Sistema implementado por computador, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o componente de comportamento de usuário adicionalmente compreende um modelo acionado por dados de comportamento de usuário.
- 20 5. Sistema implementado por computador, de acordo com a reivindicação 4, **CARACTERIZADO** pelo fato de que o mecanismo de busca adicionalmente compreende um modelo de comportamento de usuário com características diretamente observadas e características de comportamento derivadas.
- 25 6. Sistema implementado por computador, de acordo com a reivindicação 4, **CARACTERIZADO** adicionalmente pelo fato de que compreende um registro de dados que inclui dados de busca anteriores.
7. Sistema implementado por computador, de acordo com a reivindicação 1, **CARACTERIZADO** pelo fato de que o mecanismo de busca adicionalmente compreende um componente classificador que classifica os resultados de busca.
- 30 8. Sistema implementado por computador, de acordo com a reivindicação 5, **CARACTERIZADO** adicionalmente pelo fato de que compreende um componente de aprendizado de máquina que treina o modelo de comportamento de usuário.
9. Sistema implementado por computador, de acordo com a reivindicação 5, **CARACTERIZADO** pelo fato de que o modelo adicionalmente compreende características de clique, características de apresentação e características de navegação.
- 35 10. Método implementado por computador, **CARACTERIZADO** pelo fato de que compreende as seguintes ações executáveis por computador:
 - obter comportamento de usuário durante a interação com o mecanismo de busca

- (102, 202, 340, 540);
agregar o comportamento de usuário para uma análise desse; e
estimar as preferências de usuário para resultados retornados (350, 550).
11. Método implementado por computador, de acordo com a reivindicação 10,
5 **CARACTERIZADO** adicionalmente pelo fato de que compreende classificar informação re-
tornada baseada em preferências de usuário.
12. Método implementado por computador, de acordo com a reivindicação 10,
CARACTERIZADO adicionalmente pelo fato de que compreende treinar um modelo para
classificar a informação.
- 10 13. Método implementado por computador, de acordo com a reivindicação 10,
CARACTERIZADO adicionalmente pelo fato de que compreende automaticamente gerar o
modelo a partir do comportamento de usuário.
14. Método implementado por computador, de acordo com a reivindicação 10,
CARACTERIZADO adicionalmente pelo fato de que compreende desenvolver um conjunto
15 de características relacionadas à interação de usuário com a informação retornada.
15. Método implementado por computador, de acordo com a reivindicação 10,
CARACTERIZADO adicionalmente pelo fato de que compreende empregar aprendizado de
máquina para incorporar o comportamento de usuário.
16. Método implementado por computador, de acordo com a reivindicação 10,
20 **CARACTERIZADO** adicionalmente pelo fato de que compreende prever o comportamento
do usuário.
17. Método implementado por computador, de acordo com a reivindicação 10,
CARACTERIZADO adicionalmente pelo fato de que compreende minar comportamento de
usuário agregado para classificar os resultados da busca.
- 25 18. Método implementado por computador, de acordo com a reivindicação 10,
CARACTERIZADO adicionalmente pelo fato de que compreende empregar características
diretamente observadas a partir de interações de usuário com os resultados da busca para
estimar as preferências de usuário.
19. Método implementado por computador, de acordo com a reivindicação 10,
30 **CARACTERIZADO** adicionalmente pelo fato de que compreende suavizar ruído associado
com comportamento de usuário agregado.
20. Sistema implementado por computador, **CARACTERIZADO** pelo fato de que
compreende os seguintes componentes executáveis por computador:
Dispositivo (102, 202, 340, 540) para coletar retorno implícito de usuários; e
35 Dispositivo (104, 315, 515, 610) para estimar preferências de usuário.

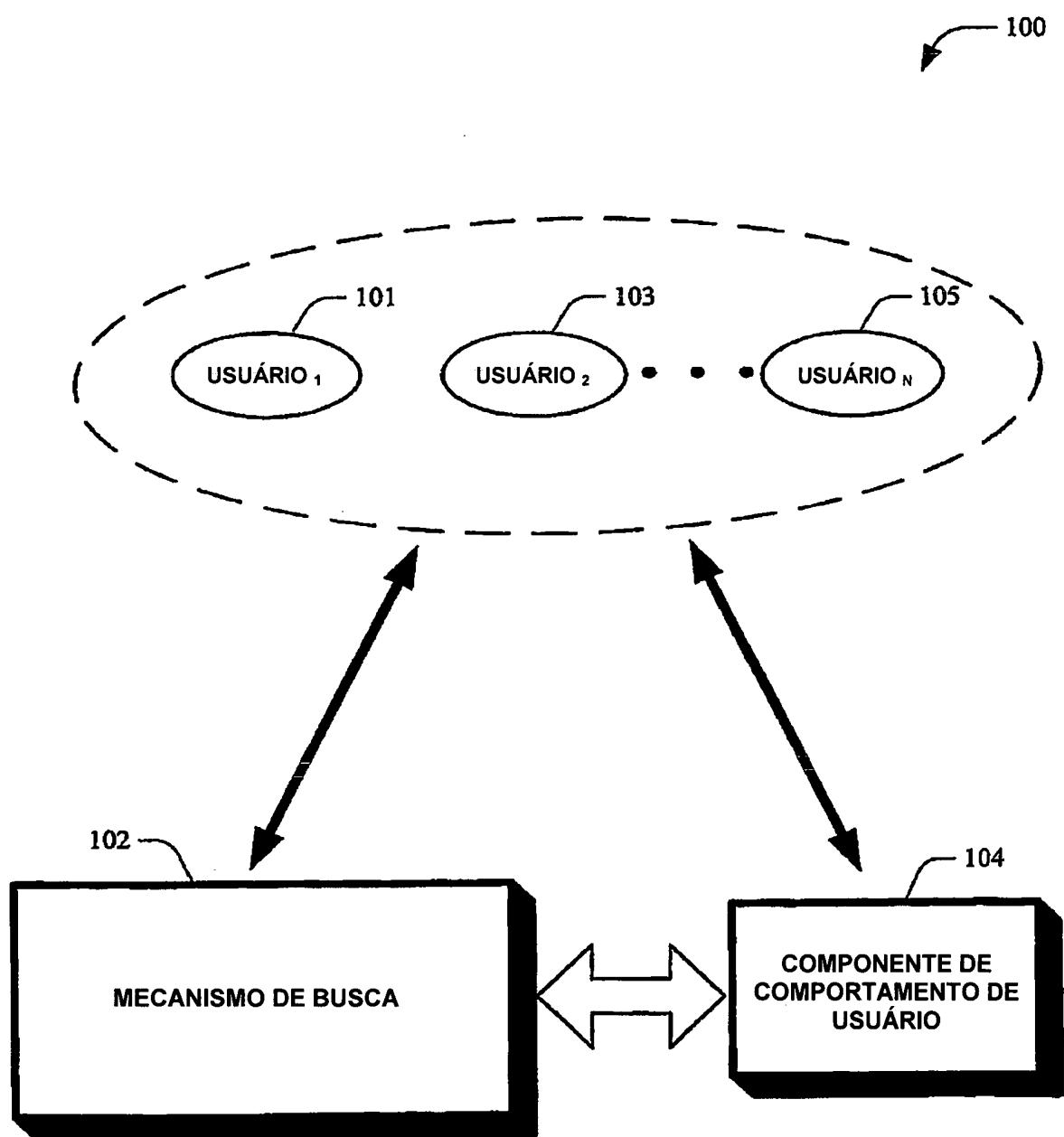


Fig. 1

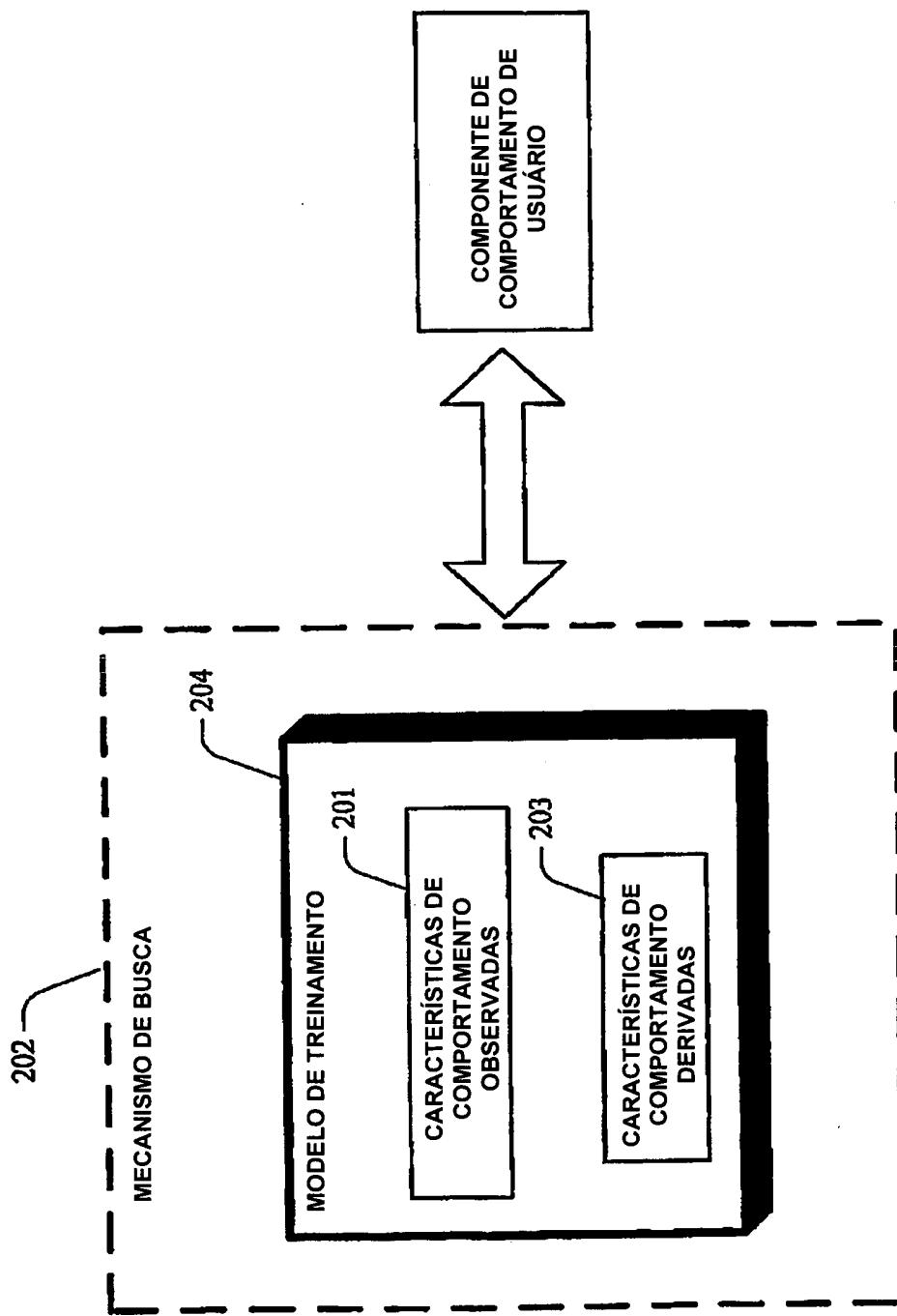


Fig. 2

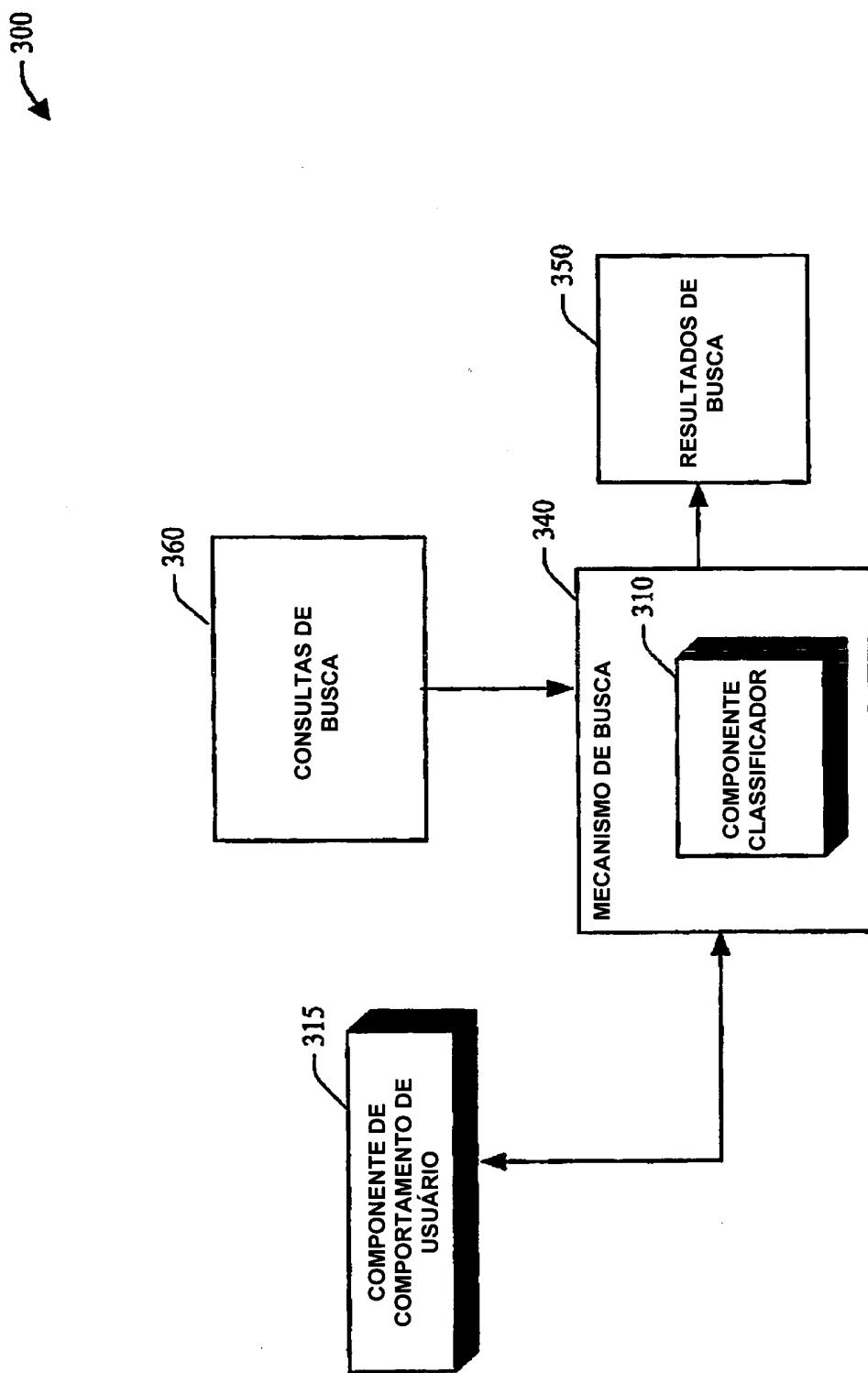


Fig. 3

410 ↗

430 →	características de clique
	position posição da url na classificação atual
	clickfrequency número de cliques para esse par consulta-url
	clickprobability probabilidade de um clique para essa consulta e url
	clickdeviation desvio da probabilidade de clique esperada
	isnextclicked 1, se há um clique na próxima posição, 0 de outra forma
	ispreviousclicked 1, se há um clique na posição anterior, 0 de outra forma
	isclickabove 1, se há um clique acima, 0 de outra forma
	isclickbelow 1, se há um clique abaixo, 0 de outra forma
420 →	características de navegação
	timeonpage tempo de permanência na página
	cumulativeimeonpage tempo cumulativo para todas as páginas subsequentes depois da busca
	timeondomain tempo de permanência cumulativo para esse domínio
	timeonshorturl tempo cumulativo em parâmetros de soltar, prefixo de url
	isfollowedlink 1, se vínculo seguido para resultado, 0 de outra forma
	isexacturlmatch 0, se normalização agressiva, 1 de outra forma
	isredirected 1, se url inicial é a mesma da url final, 0 de outra forma
	ispathfromsearch 1, se somente vínculos seguidos depois da consulta, 0 de outra forma
	clicksfromsearch número de saltos para alcançar a página da consulta
	averagedweltim tempo médio na página para essa consulta
	dwelltimedeviation desvio do tempo de permanência médio geral na página
	cumulativedeviation desvio do tempo cumulativo médio na página
	domaindeviation desvio do tempo médio no domínio
	shorturldeviation desvio no tempo médio na url curta
410 →	características de apresentação
	titleoverlap fração de palavras compartilhadas entre consulta e título
	summaryoverlap fração de palavras compartilhadas entre consulta e sumário
	queryurloverlap fração de palavras compartilhadas entre consulta e url
	querydomainoverlap fração de palavras compartilhadas entre consulta e domínio
	querylength número de sinais na consulta
	querynextoverlap fração média de palavras compartilhadas com a próxima consulta

FIG. 4

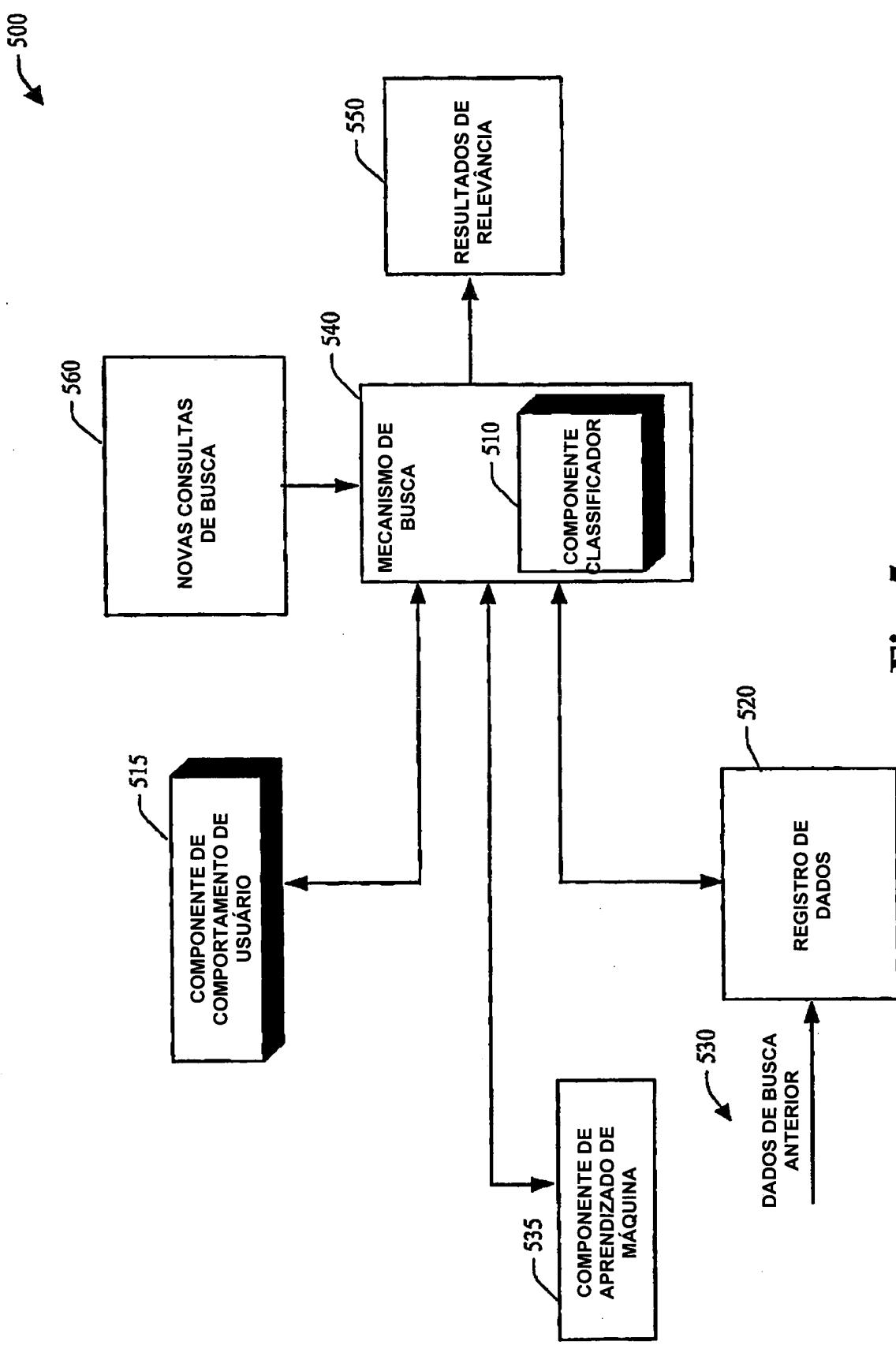


Fig. 5

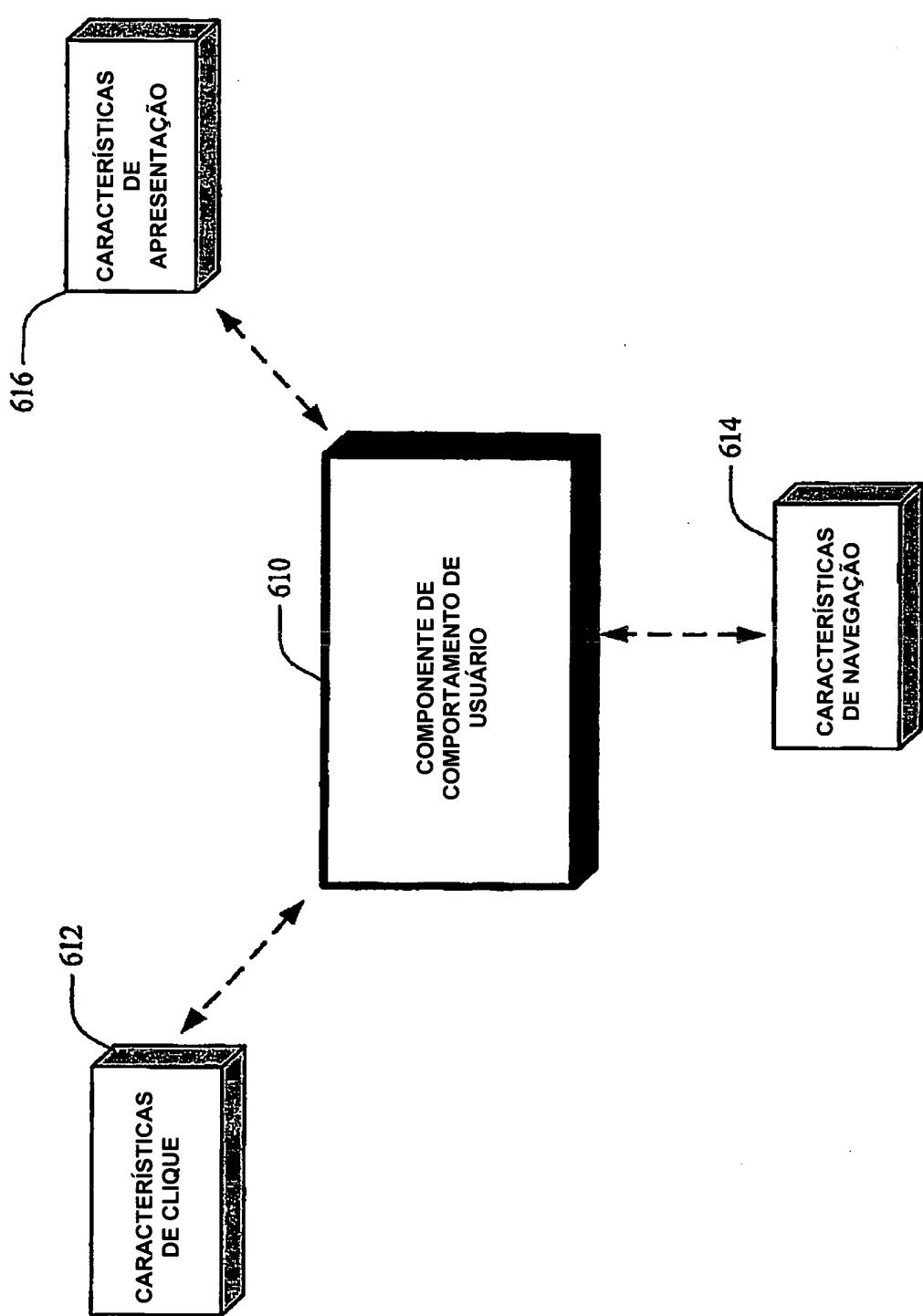


Fig. 6

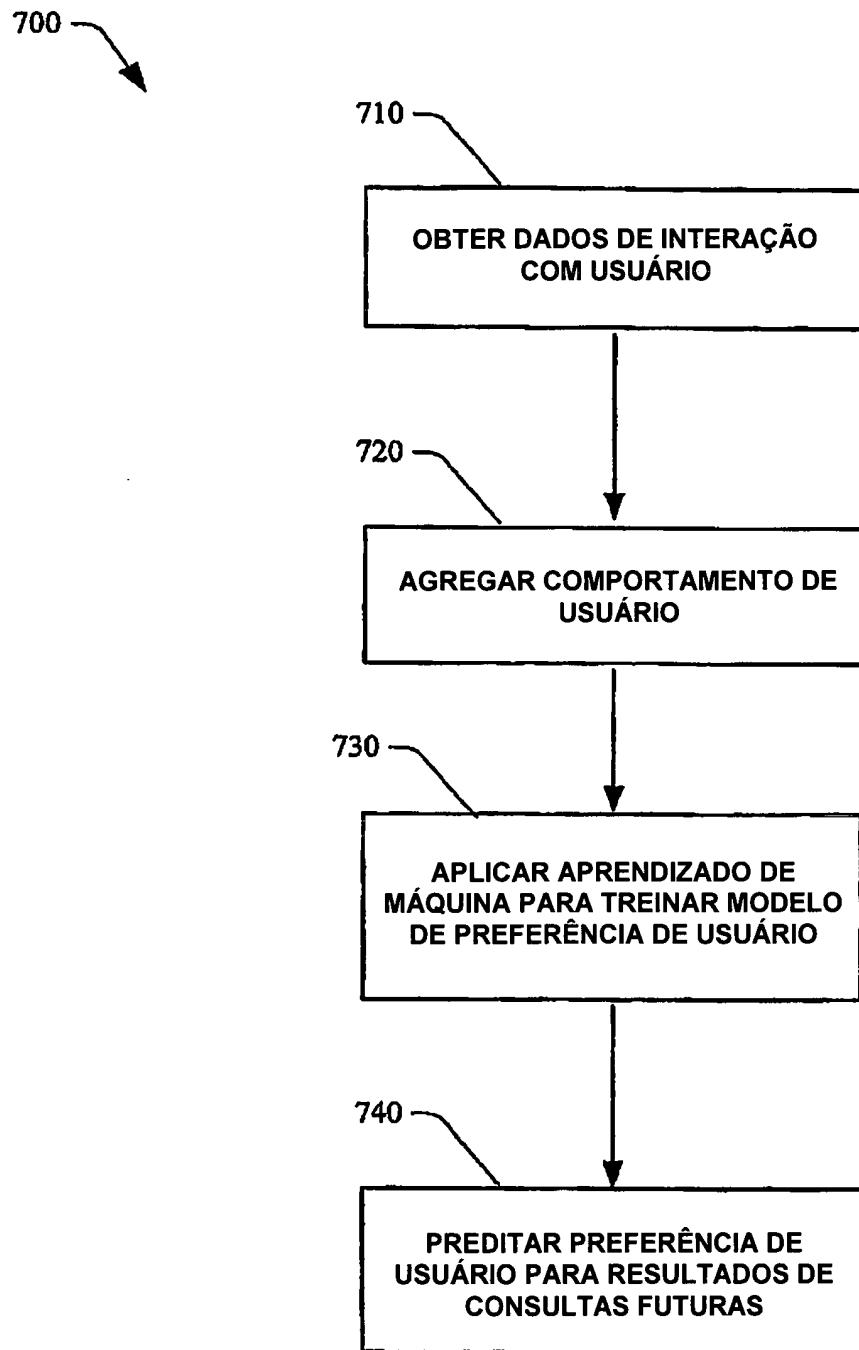


Fig. 7

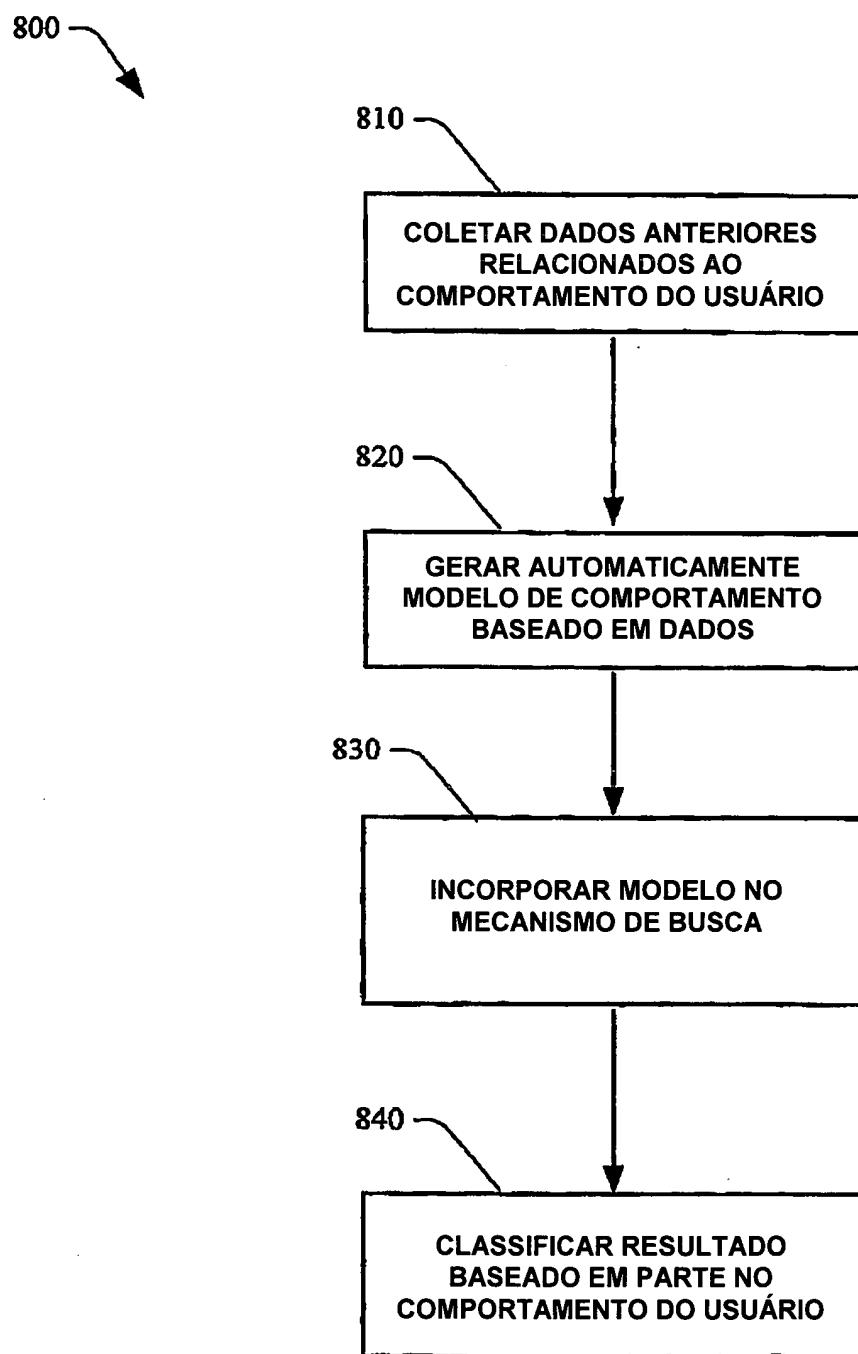


Fig. 8

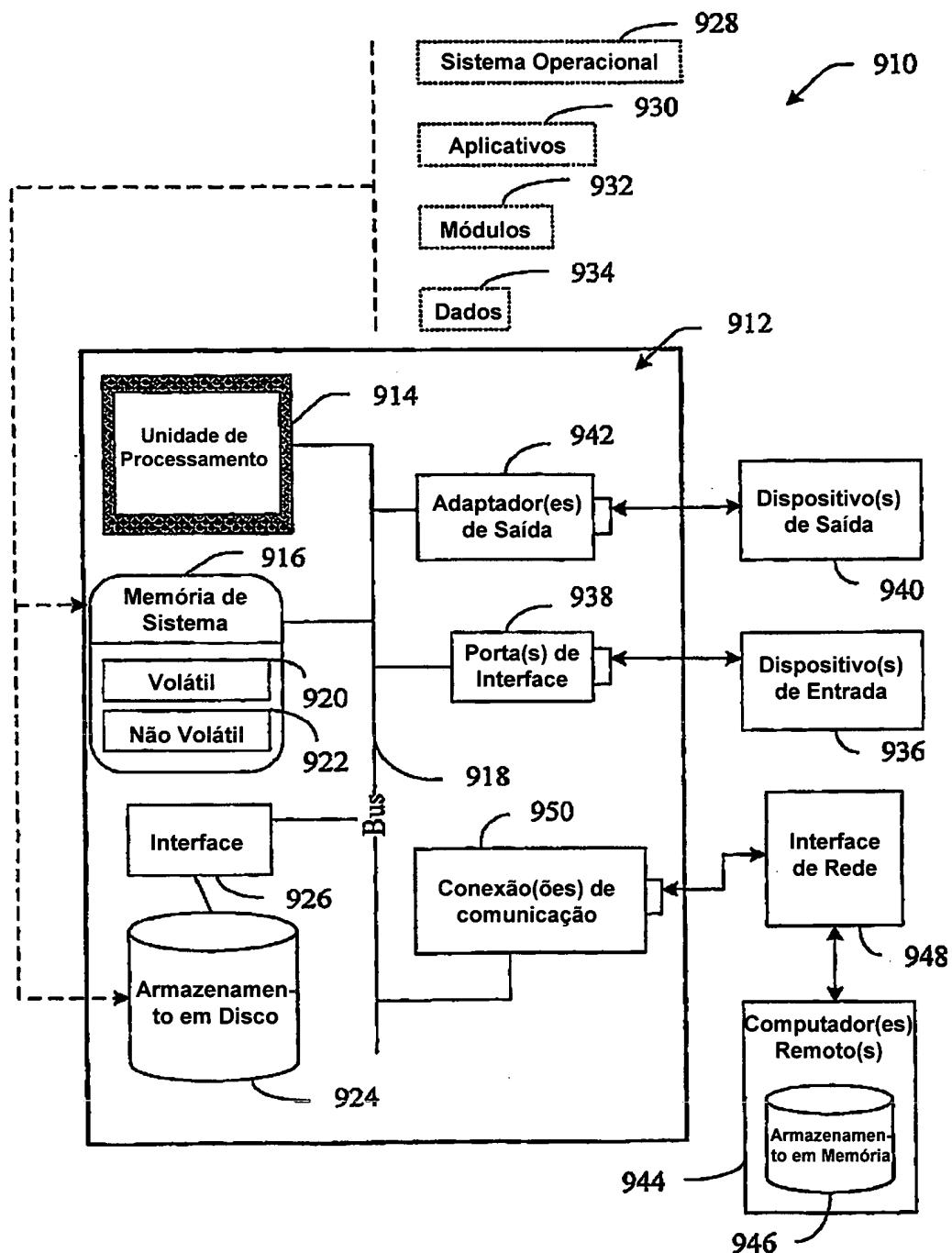


Fig. 9

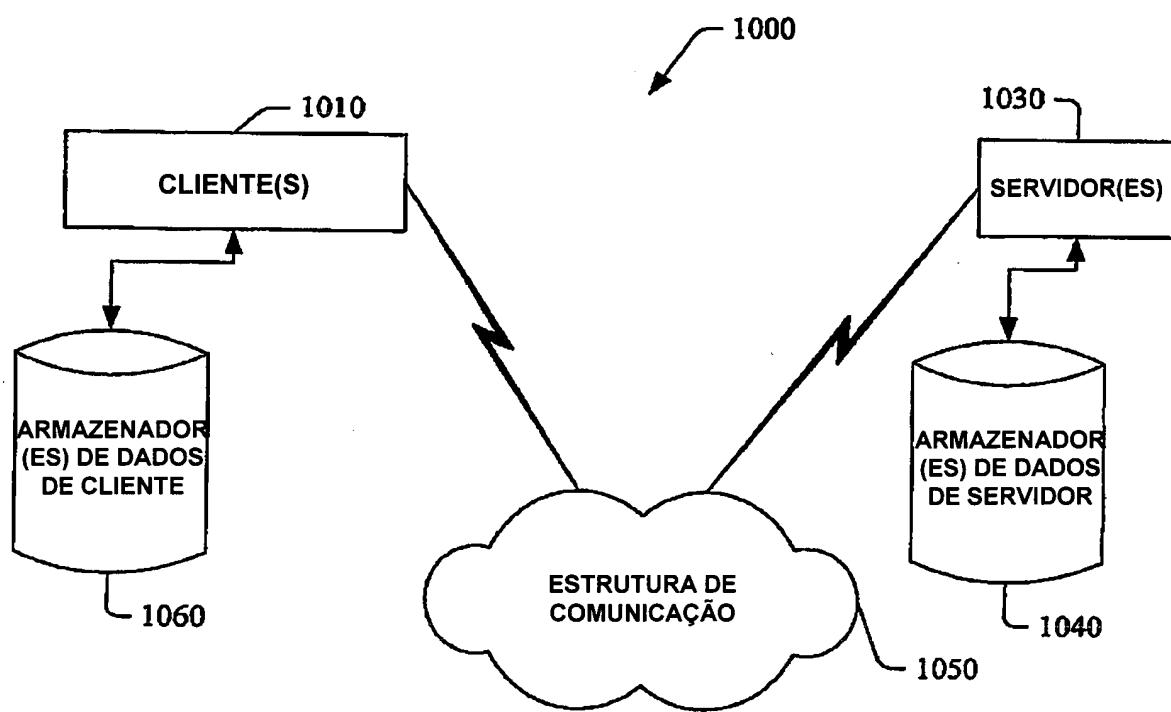


Fig. 10

RESUMO

"EXTRAÇÃO DE COMPORTAMENTO DE USUÁRIO DE BUSCA NA REDE PARA APERFEIÇOAR RELEVÂNCIA DE BUSCA NA REDE"

Sistemas e métodos que estimam preferência de usuário via interpretação automática de comportamento de usuário. Um componente de comportamento de usuário associado com um mecanismo de busca pode automaticamente interpretar comportamento coletivo de usuários (por exemplo, usuários de busca na rede). Tal componente de retorno pode incluir características de comportamento do usuário e modelos preditivos (por exemplo, a partir de um componente de comportamento de usuário) que são robustos a ruído, que podem estar presentes em iterações de usuário observadas com os resultados da busca (por exemplo, atividade de usuário maliciosa e/ou irracional).