

REVOCAÇÃO (RECALL) E PRECISÃO (PRECISION) NO SDI/CIN/CNEN*

Eratóstenes E. fí. de Araújo
Departamento de Sistemas e Computação
Universidade Federal da Paraíba
Campina Grande, PB

1 INTRODUÇÃO

Um SDI ideal é aquele capaz de dar a cada usuário, todas as referências relevantes existentes no arquivo pesquisado, sem lhe fornecer aquelas que são irrelevantes. Um sistema com essas características dificilmente pode ser encontrado, e assim sendo, as falhas de recuperação devem ser analisadas para melhorar o seu desempenho.

Esse desempenho pode ser avaliado através de dois parâmetros que medem as falhas de recuperação, definidas por Leggate³ como sendo:

- a) uma referência que foi julgada relevante para o usuário e não foi recuperada pelo sistema;
- b) uma referência que foi julgada irrelevante pelo usuário e que foi recuperada pelo sistema.

* SDI/CIN/CEN
Sistema de Disseminação Seletiva de Informações
Centro de Informações Nucleares
Comissão Nacional de Energia Nuclear

Ci. Inf., Rio de Janeiro, 8 (1):47 -50, 1979

RESUMO

Valores para a revocação e precisão no SDI/CIN/CNEN são calculados empregando-se uma nova metodologia para se determinar a relevância das referências não recuperadas do arquivo consultado. Estabelece-se a idéia de distância entre respostas para se determinar a aproximação do valor calculado para a revocação daquele que seria o valor real.

Descritores: Sistemas de informação; Parâmetros de avaliação; Desempenho; Revocação.

Os dois parâmetros são conhecidos como recall (R) e precision (P) respectivamente, Eles foram primeiro utilizados por Cleverdon² em seu trabalho no Aslib Cranfield Research Project e são definidos como:

$$R = \frac{\text{nº de referências relevantes e recuperadas}}{\text{no de referências relevantes no arquivo}} \quad (1)$$

$$P = \frac{\text{nº de referências relevantes e recuperadas}}{\text{total de referências recuperadas}} \quad (2)$$

onde:

a = referências relevantes e recuperadas

b = referências relevantes não recuperadas

c = referências irrelevantes e recuperadas

Utilizando-se esses conceitos, calculamos os valores do recall e da precision para o SDI/CIN/CNEN.

2 NÚMERO DE REFERÊNCIAS E PERFIS UTILIZADOS

Nesse estudo 18.847 referências foram pesquisadas correspondendo a um período de seis meses, equivalente a dez notificações, para um total de 35 perfis de usuários selecionados entre 213 perfis existentes no sistema.

Houve, portanto, 350 consultas ao arquivo, utilizando-se a estratégia de pesquisa por peso atribuído ao descritor e a área de interesse do usuário. A essa estratégia chamamos de critério do patamar.

3 AVALIAÇÃO DO RECALL E DA PRECISION

A avaliação desses dois parâmetros é feita medindo-se a relevância de cada referência em relação ao interesse do usuário. No SDI/CIN/CNEN, cada referência recuperada é enviada ao usuário através de notificações, e é solicitado o julgamento segundo os critérios:

- a) muito interesse (M);
- b) relativo interesse (R);
- c) nenhum interesse (N).

Dessa forma os valores de a e c na equação 2 são conhecidos. A dificuldade é se determinar o valor de b na equação 1. Num arquivo pequeno, poderia se recorrer ao usuário para que julgasse também as referências não recuperadas. Para um arquivo grande, este método não pode ser empregado, como é o caso do SDI/CIN/CNEN. Faz-se então através de medidas estimativas. Miller⁴ em seu trabalho na University of Newcastle, examinou vários métodos para estimar o valor do recall e considerou-os todos insatisfatórios. Nesse trabalho

apresentamos uma nova metodologia para se determinar a relevância das referências não recuperadas, descritas a seguir.

Essa metodologia utiliza uma avaliação exaustiva de cada referência do arquivo consultado pelo perfil do usuário, feita por especialistas (indexadores) que trabalham no sistema, responsáveis pela indexação das referências brasileiras que são colocadas no arquivo. Isto foi realizado da seguinte maneira.

Inicialmente se fez com que cada indexador se identificasse o máximo possível com o interesse do usuário atribuído para seu julgamento. Para isso, foi entregue ao indexador o perfil do usuário, onde estão indicadas as principais atividades do usuário, suas áreas de interesse (no máximo cinco) e os descritores escolhidos (no máximo vinte) e as respostas dadas as referências recuperadas. Após uma fase de análise, que chamamos de conhecimento do interesse do usuário, o indexador passou a julgar as referências recuperadas e não recuperadas. É importante salientar que o julgamento pelo indexador, das referências já recuperadas e, portanto, já avaliadas pelo usuário, se fez necessário para se medir o grau de acerto do indexador.

Os resultados encontrados foram considerados satisfatório, pois se verificou que o indexador conseguiu se aproximar bastante do interesse do usuário, conforme mostra a figura 1.

Para medir o grau de acerto do indexador se utilizou a idéia de distância entre suas respostas e do usuário, dadas as referências recuperadas. Estabelecemos assim que as respostas do tipo M—M, R—R, N—N e M—R, dadas pelo indexador e pelo usuário têm distância igual a zero, isto é, a relevância atribuída é a mesma. Para respostas do tipo R—N atribuiu-se distância igual a um, e finalmente as respostas do tipo M—N, definimos a distância como sendo igual a dois.

Distância	nº de Referências	%
0 (M-M, R-R, N-N, M-R)	3275	84,87
1 (R-N)	416	10,78
2 (M-N)	168	4,35

Figura 1 - Resultado da avaliação executada pelo indexador e pelo usuário.

Revocação (Recall) e Precisão (Precision) no SDI/CIN/CNEN
Eratóstenes E. R. de Araújo

Podemos notar que o grau de acerto do indexador é apreciável, 84,87%, com divergências para as respostas de distância 1 de 10,78% e de apenas 4,35% para as respostas de distância 2.

Pelas equações 1 e 2 podemos ver que apenas o recall é afetado pelas distorções encontradas entre a resposta do usuário e a do indexador, uma vez que só o parâmetro b da equação 1 tem seu valor calculado através da resposta do indexador. Assim sendo, observando-se os resultados da figura 1, onde a distorção apresentada na distância 1 mais aquela na distância 2 é igual a 15,00% aproximadamente, podemos afirmar que o valor do recall, se calculado com a resposta do usuário (RU) estaria entre os limites:

$$a + 1,15b \leq RU \leq a + 0,85b$$

onde b é o obtido com a resposta do indexador.

4 DESEMPENHO DO SDI/CIN/CNEN

O desempenho do SDI/CIN/CNEN é mostrado graficamente através da curva R X P, figura 2, traçada com os valores obtidos para o recall e a precision, utilizando-se as equações 1 e 2. Os pontos assinalados representam a média dos 35 perfis utilizados na pesquisa, e estão situados acima da curva proposta por Wagner⁵ e neste caso, a performance do sistema é aceitável.

Na representação individual do comportamento de cada perfil, figura 3, alguns deles precisam ser analisados, pois se encontram abaixo da curva.

Apesar da metodologia para o cálculo do recall ser diferente da utilizada por Cleverdon², é importante salientar que a sua afirmação de que há uma relação inversa entre o recall e a precision, também foi comprovada nesse trabalho.

5 CONCLUSÃO

Para centros de informação que possuam um grupo de indexadores, pensamos que a metodologia apresentada para avaliação dos documentos não recuperados pode ser inteiramente aplicada. Além da boa aproximação obtida para o cálculo do recall, essa metodologia funciona como um instrumento de capacitação dos indexadores, que passam a ter em mãos uma análise comparativa entre o que está descrito no perfil do usuário e o seu interesse real.

REFERÊNCIAS BIBLIOGRÁFICAS

- ARAÚJO, E.E.R., *Atualização de perfis em um sistema de disseminação seletiva de informação*. Rio de Janeiro, IME. 98p. 1977. Tese de Mestrado.
- CLEVERDON, C.W., *Report on testing and analysis of investigation into comparative efficiency of indexing systems*. Aslib-Cranfield Research Report, Cranfield England, 1962.
- LEGGAT, P. et alii, Evaluation of an SDI service based on the index chemicus registry system, *J. C. Docuni.*, 13: 192-203, 1973.
- MILLER, W.L., The extension of user's literatura awareness as a measure of retrieval performance, and its application to MEDLARS, *J. Doc.*, 27 125-135, Jun 1971.
- WAGNER, R.H., A selective current-awareness system using engineering index's plastic data base. Parte I - Performance, *J. C. Docum.*, 9:85-8, 1969.

ABSTRACT:

Precision and Recall values in SDI/CIN/CNEN are calculated using a new method to determine the relevance of non-recurable references from the called file. A relation between answers was established to determine the approximation of the calculated value to the one which would be the real value.

Revocação (Recall) e Precisão (Precision) no SDI/CIN/CNEN
 Eratóstenes E. R. de Araújo

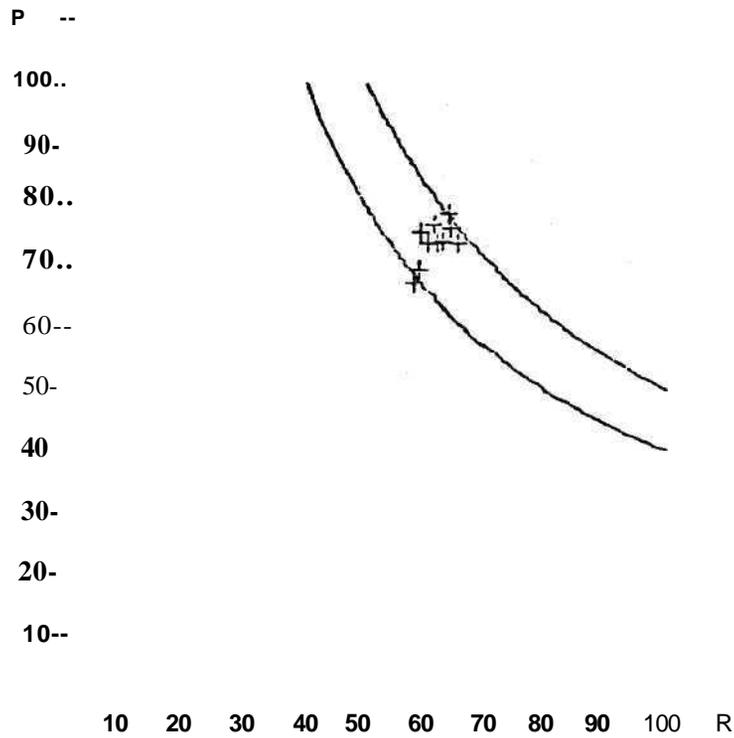


Figura 2 - R X P. média da amostra

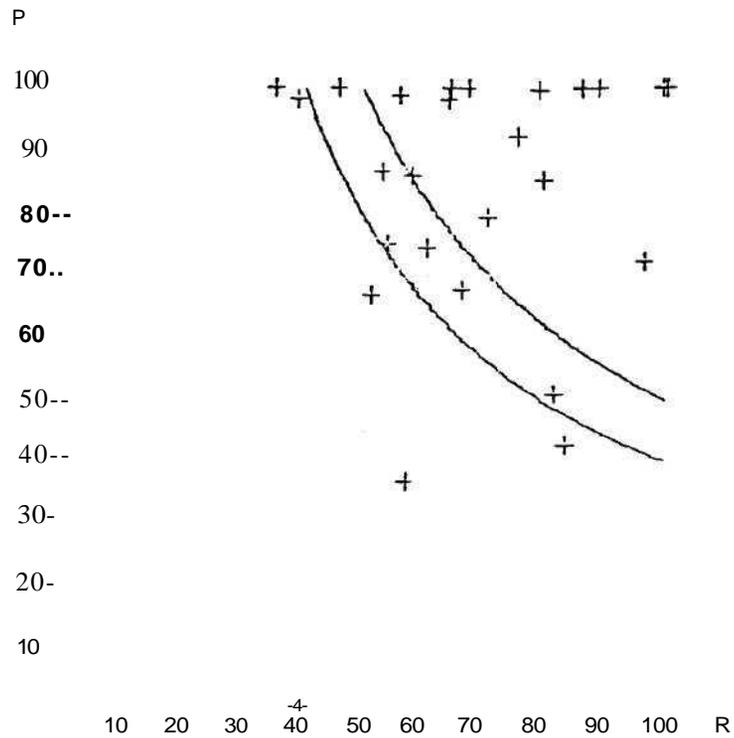


Figura 3 — Incidência de perfis no gráfico R X P