
COMPARACIÓN DE EFICACIA DE ALGORITMOS PARA BÚSQUEDA DE DOCUMENTOS, CONSIDERANDO LA SECUENCIA DE LOS RESULTADOS ARROJADOS

CARLOS ROJAS KRAMER / JORGE MARTÍNEZ CARBALLIDO

RESUMEN:

Desde hace muchos años se conoce la importancia de guardar información y recuperarla. Hoy, gracias a los sistemas de cómputo, pueden almacenarse grandes cantidades de información, por lo que se hace necesario contar con medios para hallar la que es de utilidad cuando es requerida. En el campo de la recuperación de información se han logrado avances importantes mediante la aplicación de diversas estrategias para realizar las búsquedas. La medición de la eficacia de estas técnicas es interesante, en virtud que permite identificar avances reales en este campo, y para ello se han utilizado varias medidas convencionales, como las llamadas precisión, tasa de recuperación, precisión promedio y media armónica. Desafortunadamente estas medidas no toman en cuenta de manera directa algo que para el usuario de un sistema buscador es muy importante: el orden en que los documentos son arrojados por el mecanismo de búsqueda debería corresponder con el orden de relevancia real de los documentos del conjunto sobre el cual se efectúa la búsqueda. En este trabajo de investigación se propone un modelo de evaluación cuantitativa que considera la similitud de secuencia de los resultados arrojados por el algoritmo de búsqueda con respecto a una secuencia ideal, definida por expertos humanos para una mejor valoración. Esta medida, dada su naturaleza cuantitativa, permite la comparación de eficacia entre algoritmos de búsqueda diferentes, y ha sido utilizada exitosamente en un trabajo de investigación doctoral que evalúa comparativamente diversas heurísticas de búsqueda.

PALABRAS CLAVE: recuperación, información, evaluación, eficacia, secuencia.

INTRODUCCIÓN

Desde hace muchos años se conoce la importancia de guardar información y recuperarla. Hoy, gracias a los sistemas de cómputo, pueden almacenarse

grandes cantidades de información, por lo que se hace necesario contar con medios eficaces para hallar la que es de utilidad cuando ésta es requerida.

El campo de la recuperación de información o IR, como se le conoce por sus siglas en inglés (Information Retrieval), nació en los años siguientes a 1950, justamente a partir de dicha necesidad [1]. Desde entonces este campo ha madurado de manera considerable, tanto que múltiples sistemas de IR son usados diariamente por una amplia variedad de usuarios (por ejemplo, para la búsqueda de documentos en web).

La idea fundamental de las investigaciones en el campo de la IR, es hallar la manera de construir un sistema capaz de admitir diferentes formas de expresión de la necesidad de información y, al mismo tiempo, obtener solamente los documentos relevantes. El sistema ideal debe ser capaz de identificar correctamente tanto la intención de búsqueda como el contenido de las fuentes de información, a modo de obtener los documentos que satisfacen dicha necesidad de información [2].

Se han ensayado muchas estrategias para tratar de alcanzar la mayor eficacia posible en los sistemas buscadores, y también se han diseñado diversas formas de medir la eficacia, como se discute en la siguiente sección. Sin embargo, a pesar de que muchos algoritmos de búsqueda son capaces de producir resultados clasificados de acuerdo con su probabilidad estimada de relevancia, los mecanismos que han venido empleándose para medir su eficacia –y a partir de ello compararlos– solamente toman en cuenta de manera indirecta la secuencia en que los documentos son presentados al usuario, y se basan más en juicios de relevancia binarios o booleanos (relevante *vs* no relevante).

Esto permite identificar un área de oportunidad para la investigación en IR, ya que el usuario típico de un sistema de búsqueda espera que el algoritmo le presente en primer lugar los documentos que son efectivamente más relevantes y, por lo mismo, parece congruente que las evaluaciones comparativas entre algoritmos de búsqueda tomen en cuenta este factor de manera directa e importante.

Este trabajo de investigación presenta un modelo cuantitativo para la evaluación de eficacia de algoritmos de búsqueda, considerando la secuencia de los resultados arrojados, de manera que permite obtener una medida de eficacia comparada contra una secuencia ideal, definida por expertos humanos, que hace posible una comparación cuantitativa de la eficacia entre diferentes algoritmos de búsqueda.

MEDIDAS TRADICIONALES DE EFICACIA

Una vez que se tiene representada en alguna forma la colección de documentos y la necesidad de información, el siguiente paso en IR es determinar la relación de “relevancia” entre ambas.

La forma convencional es calcular, para cada documento, una medida cuantitativa que denota la similitud, la probabilidad de relevancia o el grado de cercanía con respecto a la necesidad de información. Tales medidas cuantitativas, conocidas también como valor de estado o puntaje (*score* como suele referirse en idioma inglés) de documentos, son fuertemente influidas por la elección del esquema para ponderación de términos. La determinación de cómo son identificados los documentos “relevantes”, y cómo son ponderados los términos, dependen mayormente de los supuestos de los modelos de IR, que no sólo formalizan la IR con apoyos teóricos sino que también proporcionan los razonamientos para evaluar y analizar estrategias de recuperación.

A lo largo de las últimas cinco décadas se ha desarrollado una variedad de métodos para recuperación de información. Algunos de ellos han sido derivaciones conceptuales de alguno de los métodos que le precedieron, en el intento constante de alcanzar mejores resultados. La figura 1 ilustra los métodos para recuperación de información más reconocidos y la forma en que se han derivado unos de otros [3].

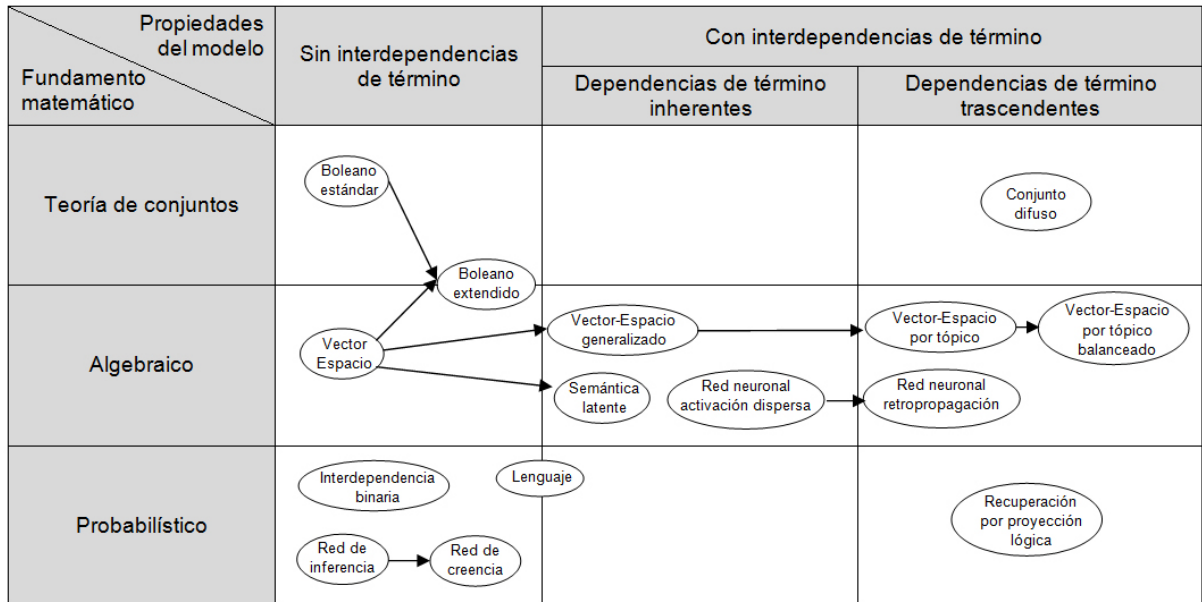


Figura 1. Modelos para recuperación de información [3]

Los sistemas para IR, independientemente del método en que se basen, pueden ser evaluados aplicando diferentes criterios, siendo los principales la eficiencia de ejecución, la eficiencia de almacenamiento, la eficacia de la recuperación de información en sí, y las facilidades de uso que se ofrecen al usuario [4].

Las medidas más frecuentemente utilizadas para evaluar la eficacia de los algoritmos de IR son las conocidas como precisión (en inglés *precision*) y tasa de recuperación (en inglés *recall*). También suelen emplearse, a fin de comparar algoritmos de búsqueda, las medidas conocidas como precisión promedio (en inglés *average precision*) y media armónica [5].

Precisión es la razón del número de documentos relevantes entre el número total de documentos arrojados por la búsqueda. Recuperación es la razón de documentos relevantes obtenidos para una consulta dada entre el total de documentos relevantes en la base de datos; aquí, con excepción de colecciones de prueba relativamente pequeñas, el denominador es generalmente desconocido y debe ser estimado mediante muestreo o algún otro método.

Tanto precisión como recuperación son medidas que toman valores entre 0 y 1, inclusive ambos.

La medida conocida como precisión promedio, y que es usada frecuentemente por la comunidad de investigadores en IR como medida de valor único para evaluar recuperación clasificada o categorizada, es calculada midiendo la precisión a intervalos predefinidos de recuperación (por ejemplo, al 0, 10%, 20%, etc.) y promediando estos valores [6].

Para otros casos donde se requiere una sola medida que combine precisión y recuperación, se tiene la media armónica F de de ambas [7], que se calcula como:

$$F(k) = \frac{2}{\frac{1}{r(k)} + \frac{1}{P(k)}}$$

donde $r(k)$ es la tasa de recuperación para el k -ésimo documento en la clasificación, $P(k)$ es la precisión para el k -ésimo documento en la clasificación, y $F(k)$ es la media armónica de $r(k)$ y $P(k)$, y por ende relativa al k -ésimo documento en la clasificación. La función F asume valores en el rango $[0,1]$. Es 0 cuando no se ha recuperado ni un documento relevante, y es 1 cuando todos los documentos en la clasificación son relevantes. La media armónica toma un valor grande solamente cuando ambas, precisión y recuperación, tienen valores altos. Por lo mismo, la determinación del máximo valor de F podría interpretarse como un intento por hallar el mejor compromiso posible entre precisión y tasa de recuperación.

En el caso particular donde k es igual al número total de documentos relevantes en la colección, sucede que $r(k)$ y $P(k)$ toman el mismo valor. La medida $F(k)$ resultante recibe en inglés el nombre de *R-precision* [8].

Debido a que precisión y recuperación son dos medidas estrechamente relacionadas, suelen discutirse conjuntamente. En cierto sentido, precisión y

recuperación están amalgamadas. Si una persona que busca información quiere únicamente los documentos que concuerdan con sus necesidades exactas, la consulta debe formularse en términos muy precisos. Sin embargo, existe el peligro de la contraparte en que, si la búsqueda es extremadamente precisa, muchos documentos relevantes no serían incluidos en el resultado. Esto explica el papel integral de la tasa de recuperación como medida de desempeño. La búsqueda debe ampliarse de modo que un número significativo de documentos queden incluidos en los resultados. Pero hay nuevamente una contraparte: si se incrementa la tasa de recuperación, es muy probable que el usuario tenga que lidiar con muchos documentos no deseados que aparecerán combinados en el resultado con los documentos verdaderamente relevantes. Lo que el usuario del sistema necesita generalmente es un subconjunto completo de documentos relevantes que no requiera un esfuerzo sustancial para librarse del material irrelevante [5].

LIMITACIONES DE LAS MEDIDAS TRADICIONALES DE EFICACIA

La experiencia de años ha mostrado que precisión y tasa de recuperación son un tanto contradictorias en la práctica. Las técnicas orientadas a mejorar la precisión suelen reducir la tasa de recuperación y viceversa. Ambas medidas son orientadas al conjunto respuesta, pero no incorporan la noción de recuperación clasificada o categorizada, por lo que los investigadores han debido usar variantes de las dos medidas para evaluaciones en estos casos. Por ejemplo, si los diseñadores consideran que la precisión es de mayor importancia para los usuarios, se podría utilizar como métrica de evaluación la precisión de los primeros diez o veinte documentos del conjunto respuesta. Por el otro lado, si la tasa de recuperación fuera más importante, podría medirse la precisión que se da en algún nivel específico de recuperación (por ejemplo al 50%), lo cual indicaría cuántos documentos no relevantes tendría que revisar el usuario para hallar dicho porcentaje de los documentos relevantes [9].

Además, desafortunadamente, estas medidas tienen limitaciones para evaluar el ordenamiento débil de los resultados [10, 11]. La única consideración sobre el orden de los resultados, que estas medidas de eficacia incluyen, está relacionada con la ausencia o presencia de resultados relevantes dentro del conjunto de documentos identificados por el algoritmo en respuesta a una expresión de búsqueda específica. Sin embargo, el nivel de relevancia relativo de un documento en comparación al de los otros no es medido directamente; y esto es cierto aún cuando se comparan algoritmos del tipo probabilístico, donde el ordenamiento de los resultados es dado en función de la probabilidad de relevancia estimada por el algoritmo.

MEDIDA DE LA SIMILITUD DE SECUENCIA

La problemática ya expuesta ha motivado a la definición de una mecánica para efectuar la evaluación de similitud de secuencia de la lista de documentos seleccionados por el algoritmo, contra otra secuencia, que para el caso se considere la ideal. Para disponer de mediciones lo más certeras posibles, la secuencia ideal debería ser definida por expertos humanos en el tema de la búsqueda.

La premisa principal para poder aplicar evaluación de eficacia considerando la similitud de secuencia, es que se dispone de una colección de documentos de prueba cuya relevancia ha sido calificada por expertos humanos, quienes además han establecido el ordenamiento ideal para los documentos relevantes. A continuación se define el mecanismo cuantitativo que se propone.

La idea subyacente para el mecanismo, comienza con extraer los documentos relevantes del conjunto respuesta, sin perder el ordenamiento arrojado por el algoritmo de búsqueda sujeto a evaluación, y después, para cada uno de los documentos en el conjunto resultante, estimar la proporción de los documentos que aparecen mejor ubicados en la clasificación dada por el algoritmo que están también mejor ubicados en la clasificación de los expertos. El resultado de este cálculo es un valor entre 0 y 1, que expresa la razón de documentos

adecuadamente posicionados en la clasificación entre el total de posibles variantes, correspondiendo los valores mayores (más cercanos a 1) a los ordenamientos más semejantes al ideal (clasificación de humanos expertos).

Este concepto permite calcular la medida de correspondencia de secuenciación para una fracción del conjunto respuesta de una consulta, siempre y cuando se tomen en cuenta los primeros k elementos contados a partir del primero (aquél que encabeza la lista de documentos arrojados).

La primera parte del cálculo consiste en obtener, del conjunto de documentos respuesta, el subconjunto de documentos correctamente identificados como relevantes:

$$MR = DT_k \cap RT$$

donde DT_k es el subconjunto compuesto por los primeros k documentos del conjunto respuesta original ordenado por ranking, y RT es el conjunto de todos los documentos relevantes identificados por los humanos expertos. El conjunto resultante MR debe obtenerse sin perder la secuencia original de los documentos en el conjunto respuesta.

Para formalizar la segunda parte del cálculo, considérese la función $RT_{pos}(D)$ que toma el valor del índice dentro la colección ordenada RT al que corresponde el documento D . Nótese que los documentos mejor posicionados en la clasificación son los que tienen los índices menores. Considérese también la función $U(D1, D2)$ que toma el valor 1 si $RT_{pos}(D1) < RT_{pos}(D2)$, o el valor 0 en caso contrario.

Considerando los primeros k documentos del conjunto respuesta original, la medida de similitud de secuencia $s(k)$ queda definida por:

$$S(k) = 1, \text{ si } |MR| \leq k$$

$$S(k) = \frac{\sum_{i=1}^{|MR|-1} \sum_{j=2}^{|MR|} U(MR_i, MR_j)}{\binom{|MR| \cdot (|MR|-1)}{2}}, \text{ si } |MR| \geq 2$$

MEDIDA DE LA EFICACIA UTILIZANDO LA SIMILITUD DE SECUENCIA

La medida de similitud de secuencia es $S(k)$, en cierta forma, un calificador adicional de la tradicional medida de precisión $P(k)$. Para incorporar estas dos medidas en una sola, se define aquí una forma modificada de esta última, a la que se nombrará “precisión secuenciada” y se representará simbólicamente como $PS(k)$, donde k es el número de elementos del conjunto respuesta considerados en el cálculo (contados en secuencia, comenzando con el primero de la clasificación):

$$PS(k) = \sqrt{P(k) \cdot S(k)}$$

Ahora, de manera semejante a la media armónica F , que mide la eficacia relativa hasta el elemento k del conjunto respuesta, usando la tasa de recuperación y la medida de precisión tradicional, puede calcularse también una medida modificada de eficacia utilizando la precisión secuenciada en vez de la tradicional. A esta medida, en lo sucesivo, se le dará aquí el nombre de media armónica modificada:

$$G(k) = \frac{2}{\frac{1}{r(k)} + \frac{1}{PS(k)}}$$

Siguiendo el mismo concepto de *R-precision* como medida de valor único que puede usarse para efectos comparativos entre algoritmos, puede definirse una medida modificada usando la función G aquí definida aplicada al caso particular en que $k=NR$ (tal como se hace para el cálculo habitual de *R-precision*), siendo NR el número de documentos relevantes en la colección de prueba, según juicio humano experto. A esta medida, en lo sucesivo, se le dará aquí el nombre de *R-precision* modificada:

$$G(NR) = \frac{2}{\frac{1}{r(NR)} + \frac{1}{\sqrt{P(NR) \cdot S(NR)}}}$$

Nótese que para el caso específico $k=NR$, que es el utilizado para el cálculo de la medida única, los valores de $r(NR)$ y $P(NR)$ resultan siempre idénticos, por lo que $G(NR)$ puede expresarse también como:

$$G(NR) = \frac{2 \cdot P(NR) \cdot \sqrt{S(NR)}}{\sqrt{S(NR)} + \sqrt{P(NR)}}$$

Nótese también, que el valor de $G(NR)$ queda indefinido si cualquiera de $r(NR)$, $P(NR)$ y/o $S(NR)$ es igual a cero. Por simplicidad, y en congruencia con el concepto de la medida de eficacia, se asume que $G(NR)$ vale cero en estos casos.

Para ilustrar el efecto de la evaluación de *R-precision* modificada $G(NR)$ en comparación a *R-precision* tradicional $F(NR)$, en las tablas 1 y 2 (al final de este documento) se presentan los valores calculados para un caso de prueba, en el que dos buscadores convencionales, Google y ht://Dig, fueron utilizados para responder a la misma expresión de búsqueda, sobre un conjunto de 100 documentos de prueba. A juicio de humanos expertos, sólo 73 de los 100 documentos eran relevantes, y se clasificaron en orden decreciente de relevancia. En cada tabla, la columna k enumera las 73 posiciones del conjunto respuesta que fueron consideradas para las evaluaciones; la segunda columna se titula con el nombre del buscador y muestra, para cada posición del conjunto respuesta arrojado por éste, la posición correspondiente del documento en la clasificación de los expertos humanos, utilizándose un guión para señalar los no relevantes. La columna N_Rel muestra la cantidad de documentos relevantes hasta la posición k . Las columnas restantes muestran los valores para las medidas calculadas: tasa de recuperación $r(k)$, precisión $P(k)$, media armónica $F(k)$, similitud de secuencia $S(k)$, precisión secuenciada $PS(k)$ y media armónica modificada $G(k)$.

Los valores de $F(k)$ y $G(k)$ en la última fila, donde $k=73$ (el número total de documentos relevantes en la colección, a juicio experto), representan respectivamente las medidas de *R-precision* y *R-precision* modificada obtenidas por cada algoritmo para el caso de prueba. Nótese que la diferencia entre estos valores resulta mayor para ht://Dig que para Google, y que el valor de $G(k)$ es mayor para Google que para ht://Dig a pesar de que el último obtuvo mayor cantidad de documentos relevantes. Ambas situaciones resultan de la consideración de la similitud de secuencia implícita en $G(k)$, que favorece la puntuación para Google sobre ht://Dig en este caso, debido a que arrojó resultados con secuencia más similar a la establecida como óptima por los expertos humanos.

DISCUSIÓN

La medida aquí propuesta para evaluación de la eficacia de algoritmos para búsqueda de documentos, considerando la secuencia de los resultados arrojados, tiene un sustento teórico que parte de dos de las medidas de eficacia más frecuentemente utilizadas, que son precisión y tasa de recuperación, combinadas en otra medida preferida, que es la media armónica de ambas, y que calculada para el punto que iguala el número de documentos relevantes del conjunto se conoce como *R-precision*.

De hecho la medida propuesta es en realidad una modificación relativamente simple de *R-precision*, donde la diferencia radica en el uso de una forma especial de la medida usual de precisión, a la cual se le ha llamado precisión secuenciada, siendo ésta la raíz cuadrada del producto de la medida tradicional de precisión y una medida de similitud de secuencia cuya mecánica de cálculo es una propuesta original de este trabajo.

La medida de evaluación de eficacia ha sido puesta a prueba en la evaluación de un conjunto de 1,210 heurísticas de búsqueda, como parte de un proyecto de investigación doctoral, y los resultados observados se han mostrado

consistentes, resultando favorable su uso para la evaluación comparativa requerida.

REFERENCIAS

1. Luhn, H.P. (1957). "A statistical approach to mechanized encoding and searching of literary information," *IBM Journal of Research and Development*, vol. 1, pp. 309-317.
2. Witten, et al. (1994). *Managing Gigabytes: Compressing and Indexing Documents and Images*, Van Nostrand Reinhold.
3. Kuroepka, D. (2004). *Modelle zur Repräsentation natürlichsprachlicher Dokumente. Ontologie-basiertes Information-Filtering und -Retrieval mit relationalen Datenbanken*, Logos Verlag Berlin.
4. Frakes, W. B. y R. Baeza-Yates, eds., (1992). *Information Retrieval: Data Structures & Algorithms*, Prentice Hall PTR.
5. Berry M.W. y M. Browne (2005). *Understanding Search Engines - Mathematical Modeling and Text Retrieval*, Society for Industrial and Applied Mathematics.
6. Salton, G. y M.J. McGill (1983). *Introduction to modern information retrieval*, McGraw-Hill.
7. Shaw, W.M. Jr., et al. (1997). "Performance standards and evaluations in IR test collections: Cluster-based retrieval models," *Information Processing & Management*, vol. 33, núm. 1, pp. 1-14.
8. Baeza-Yates y B. Ribeiro-Neto (1999). *Modern Information Retrieval*, Addison-Wesley / ACM Press.
9. Singhal, A. (2001). "Modern Information Retrieval: A Brief Overview," *IEEE Data Engineering Bulletin*, vol. 24, núm. 4, pp. 35-43.
10. Raghavan, VV (1989). al., "Retrieval System Evaluation Using Recall and Precision: Problems and Answers," *Proc. 12th ACM SIGIR Conference*, 1989.
11. R.M. Losee, (2000) "When Information Retrieval Measures Agree about the Relative Quality of Document Rankings," *Journal of the American Society for Information Science*, vol. 51, no. 9, pp. 834-840.

TABLAS

Tabla 1. Resultados para el caso ejemplo usando Google

k	Google	N	Rel	r(k)	P(k)	F(k)	S(k)	PS(k)	G(k)
1	63	1		0.014	1.000	0.027	1.000	1.000	0.027
2	72	2		0.027	1.000	0.053	1.000	1.000	0.053
3	54	3		0.041	1.000	0.079	0.333	0.577	0.077
4	30	4		0.055	1.000	0.104	0.167	0.408	0.097
5	-	4		0.055	0.800	0.103	0.167	0.365	0.095
6	-	4		0.055	0.667	0.101	0.167	0.333	0.094
7	-	4		0.055	0.571	0.100	0.167	0.309	0.093
8	4	5		0.068	0.625	0.123	0.100	0.250	0.108
9	53	6		0.082	0.667	0.146	0.200	0.365	0.134
10	-	6		0.082	0.600	0.145	0.200	0.346	0.133
11	5	7		0.096	0.636	0.167	0.190	0.348	0.150
12	55	8		0.110	0.667	0.188	0.321	0.463	0.177
13	56	9		0.123	0.692	0.209	0.417	0.537	0.201
14	6	10		0.137	0.714	0.230	0.378	0.519	0.217
15	1	11		0.151	0.733	0.250	0.309	0.476	0.229
16	57	12		0.164	0.750	0.270	0.394	0.544	0.252
17	36	13		0.178	0.765	0.289	0.397	0.551	0.269
18	-	13		0.178	0.722	0.286	0.397	0.536	0.267
19	-	13		0.178	0.684	0.283	0.397	0.521	0.265
20	58	14		0.192	0.700	0.301	0.462	0.568	0.287
21	-	14		0.192	0.667	0.298	0.462	0.555	0.285
22	37	15		0.205	0.682	0.316	0.457	0.558	0.300
23	27	16		0.219	0.696	0.333	0.433	0.549	0.313
24	2	17		0.233	0.708	0.351	0.390	0.525	0.323
25	-	17		0.233	0.680	0.347	0.390	0.515	0.321
26	10	18		0.247	0.692	0.364	0.379	0.512	0.333
27	43	19		0.260	0.704	0.380	0.398	0.529	0.349
28	25	20		0.274	0.714	0.396	0.389	0.527	0.361
29	13	21		0.288	0.724	0.412	0.381	0.525	0.372
30	-	21		0.288	0.700	0.408	0.381	0.516	0.370
31	-	21		0.288	0.677	0.404	0.381	0.508	0.367
32	-	21		0.288	0.656	0.400	0.381	0.500	0.365
33	42	22		0.301	0.667	0.415	0.398	0.515	0.380
34	7	23		0.315	0.676	0.430	0.383	0.509	0.389
35	8	24		0.329	0.686	0.444	0.373	0.506	0.399
36	59	25		0.342	0.694	0.459	0.417	0.538	0.418
37	28	26		0.356	0.703	0.473	0.418	0.542	0.430

k	Google	N	Rel	r(k)	P(k)	F(k)	S(k)	PS(k)	G(k)
38	38	27		0.370	0.711	0.486	0.430	0.553	0.443
39	39	28		0.384	0.718	0.500	0.442	0.563	0.456
40	3	29		0.397	0.725	0.513	0.416	0.549	0.461
41	-	29		0.397	0.707	0.509	0.416	0.543	0.459
42	-	29		0.397	0.690	0.504	0.416	0.536	0.456
43	60	30		0.411	0.698	0.517	0.451	0.561	0.474
44	-	30		0.411	0.682	0.513	0.451	0.554	0.472
45	-	30		0.411	0.667	0.508	0.451	0.548	0.470
46	29	31		0.425	0.674	0.521	0.449	0.550	0.479
47	-	31		0.425	0.660	0.517	0.449	0.544	0.477
48	61	32		0.438	0.667	0.529	0.480	0.566	0.494
49	31	33		0.452	0.673	0.541	0.479	0.568	0.503
50	62	34		0.466	0.680	0.553	0.506	0.587	0.519
51	40	35		0.479	0.686	0.565	0.511	0.592	0.530
52	-	35		0.479	0.673	0.560	0.511	0.586	0.528
53	-	35		0.479	0.660	0.556	0.511	0.581	0.525
54	44	36		0.493	0.667	0.567	0.519	0.588	0.537
55	64	37		0.507	0.673	0.578	0.544	0.605	0.551
56	41	38		0.521	0.679	0.589	0.545	0.608	0.561
57	65	39		0.534	0.684	0.600	0.567	0.623	0.575
58	66	40		0.548	0.690	0.611	0.587	0.636	0.589
59	32	41		0.562	0.695	0.621	0.578	0.634	0.596
60	67	42		0.575	0.700	0.632	0.597	0.646	0.609
61	-	42		0.575	0.689	0.627	0.597	0.641	0.606
62	-	42		0.575	0.677	0.622	0.597	0.636	0.604
63	-	42		0.575	0.667	0.618	0.597	0.631	0.602
64	68	43		0.589	0.672	0.628	0.615	0.643	0.615
65	-	43		0.589	0.662	0.623	0.615	0.638	0.612
66	69	44		0.603	0.667	0.633	0.631	0.649	0.625
67	46	45		0.616	0.672	0.643	0.629	0.650	0.633
68	47	46		0.630	0.676	0.652	0.628	0.652	0.641
69	33	47		0.644	0.681	0.662	0.617	0.648	0.646
70	48	48		0.658	0.686	0.671	0.617	0.650	0.654
71	-	48		0.658	0.676	0.667	0.617	0.646	0.652
72	49	49		0.671	0.681	0.676	0.617	0.648	0.660
73	70	50		0.685	0.685	0.685	0.632	0.658	0.671

Tabla 2. Resultados para el caso ejemplo usando ht://Dig

k	ht://Dig	N	Rel	r(k)	P(k)	F(k)	S(k)	PS(k)	G(k)
1	-	0		0.000	0.000	0.000	1.000	0.000	0.000
2	33	1		0.014	0.500	0.027	1.000	0.707	0.027
3	73	2		0.027	0.667	0.053	1.000	0.816	0.053
4	48	3		0.041	0.750	0.078	0.667	0.707	0.078
5	67	4		0.055	0.800	0.103	0.667	0.730	0.102
6	41	5		0.068	0.833	0.127	0.500	0.645	0.124
7	53	6		0.082	0.857	0.150	0.533	0.676	0.147
8	-	6		0.082	0.750	0.148	0.533	0.632	0.145
9	34	7		0.096	0.778	0.171	0.429	0.577	0.164
10	-	7		0.096	0.700	0.169	0.429	0.548	0.163
11	35	8		0.110	0.727	0.190	0.393	0.535	0.182
12	21	9		0.123	0.750	0.212	0.306	0.479	0.196
13	-	9		0.123	0.692	0.209	0.306	0.460	0.194
14	20	10		0.137	0.714	0.230	0.244	0.418	0.206
15	38	11		0.151	0.733	0.250	0.291	0.462	0.227
16	32	12		0.164	0.750	0.270	0.273	0.452	0.241
17	58	13		0.178	0.765	0.289	0.359	0.524	0.266
18	-	13		0.178	0.722	0.286	0.359	0.509	0.264
19	19	14		0.192	0.737	0.304	0.308	0.476	0.273
20	71	15		0.205	0.750	0.323	0.390	0.541	0.298
21	59	16		0.219	0.762	0.340	0.442	0.580	0.318
22	9	17		0.233	0.773	0.358	0.390	0.549	0.327
23	4	18		0.247	0.783	0.375	0.346	0.521	0.335
24	-	18		0.247	0.750	0.371	0.346	0.510	0.332
25	44	19		0.260	0.760	0.388	0.374	0.533	0.350
26	-	19		0.260	0.731	0.384	0.374	0.523	0.348
27	-	19		0.260	0.704	0.380	0.374	0.513	0.345
28	37	20		0.274	0.714	0.396	0.384	0.524	0.360
29	-	20		0.274	0.690	0.392	0.384	0.515	0.358
30	47	21		0.288	0.700	0.408	0.410	0.535	0.374
31	27	22		0.301	0.710	0.423	0.394	0.529	0.384
32	-	22		0.301	0.688	0.419	0.394	0.520	0.382
33	72	23		0.315	0.697	0.434	0.443	0.555	0.402
34	51	24		0.329	0.706	0.449	0.464	0.572	0.418
35	40	25		0.342	0.714	0.463	0.467	0.577	0.430
36	1	26		0.356	0.722	0.477	0.431	0.558	0.435
37	63	27		0.370	0.730	0.491	0.462	0.580	0.452

k	ht://Dig	N	Rel	r(k)	P(k)	F(k)	S(k)	PS(k)	G(k)
38	36	28		0.384	0.737	0.505	0.458	0.581	0.462
39	3	29		0.397	0.744	0.518	0.429	0.565	0.466
40	50	30		0.411	0.750	0.531	0.446	0.578	0.480
41	7	31		0.425	0.756	0.544	0.424	0.566	0.485
42	8	32		0.438	0.762	0.557	0.405	0.556	0.490
43	30	33		0.452	0.767	0.569	0.400	0.554	0.498
44	14	34		0.466	0.773	0.581	0.387	0.547	0.503
45	31	35		0.479	0.778	0.593	0.385	0.547	0.511
46	64	36		0.493	0.783	0.605	0.413	0.568	0.528
47	29	37		0.507	0.787	0.617	0.407	0.566	0.535
48	17	38		0.521	0.792	0.628	0.395	0.560	0.539
49	-	38		0.521	0.776	0.623	0.395	0.554	0.537
50	2	39		0.534	0.780	0.634	0.377	0.542	0.538
51	-	39		0.534	0.765	0.629	0.377	0.537	0.535
52	-	39		0.534	0.750	0.624	0.377	0.531	0.533
53	-	39		0.534	0.736	0.619	0.377	0.526	0.530
54	15	40		0.548	0.741	0.630	0.368	0.522	0.535
55	42	41		0.562	0.745	0.641	0.382	0.533	0.547
56	-	41		0.562	0.732	0.636	0.382	0.529	0.545
57	-	41		0.562	0.719	0.631	0.382	0.524	0.542
58	54	42		0.575	0.724	0.641	0.402	0.539	0.557
59	49	43		0.589	0.729	0.652	0.416	0.551	0.569
60	45	44		0.603	0.733	0.662	0.427	0.560	0.580
61	28	45		0.616	0.738	0.672	0.422	0.558	0.586
62	16	46		0.630	0.742	0.681	0.413	0.553	0.589
63	69	47		0.644	0.746	0.691	0.435	0.570	0.604
64	61	48		0.658	0.750	0.701	0.452	0.582	0.618
65	11	49		0.671	0.754	0.710	0.440	0.576	0.620
66	65	50		0.685	0.758	0.719	0.458	0.589	0.633
67	-	50		0.685	0.746	0.714	0.458	0.585	0.631
68	60	51		0.699	0.750	0.723	0.472	0.595	0.643
69	5	52		0.712	0.754	0.732	0.457	0.587	0.644
70	10	53		0.726	0.757	0.741	0.446	0.581	0.645
71	56	54		0.740	0.761	0.750	0.458	0.590	0.656
72	-	54		0.740	0.750	0.745	0.458	0.586	0.654
73	23	55		0.753	0.753	0.753	0.453	0.584	0.658