

[0066] *Link-Based Criteria*

[0066] Criteri Basati sui link

[0067] According to an implementation consistent with the principles of the invention, one or more link-based factors may be used to generate (or alter) a score associated with a document. In one implementation, the link-based factors may relate to the dates that new links appear to a document and that existing links disappear. The appearance date of a link may be the first date that search engine 125 finds the link or the date of the document that contains the link (e.g., the date that the document was found with the link or the date that it was last updated). The disappearance date of a link may be the first date that the document containing the link either dropped the link or disappeared itself.

[0067] In un'applicazione coerente con i principi dell'invenzione uno o più fattori riguardanti i link possono essere usati per generare (o alterare) un punteggio associato con un documento. In un'implementazione i fattori riguardanti i link possono essere relativi alle date in cui appaiono nuovi link che puntano ai documenti e a quelle in cui link esistenti spariscono. La **data di apparizione di un link** può essere la prima data in cui il motore di ricerca trova il link o la data associata al documento che contiene il link (per esempio la data in cui viene trovato il documento con il link o la data del suo ultimo aggiornamento). La data di sparizione di un link può essere la data nella quale il documento contenente il link viene cancellato oppure elimina il link.

[0068] These dates may be determined by search engine 125 during a crawl or index update operation. Using this date as a reference, search engine 125 may then monitor the time-varying behavior of links to the document, such as when links appear or disappear, the rate at which links appear or disappear over time, how many links appear or disappear during a given time period, whether there is trend toward appearance of new links versus disappearance of existing links to the document, etc.

[0068] Queste date possono essere determinate dal motore di ricerca durante l'esplorazione del web oppure durante un'operazione di aggiornamento del suo indice. Usando queste date come riferimento il motore può monitorare le **variazioni nel tempo dei link che puntano ad un documento**, come quando i link appaiono o scompaiono, la velocità alla quale i link appaiono o scompaiono nel tempo, quanti link appaiono o scompaiono in un dato periodo di tempo, se esiste un trend verso l'apparizione di nuovi link al documento o verso la sparizione di quelli esistenti, ecc.

[0069] Using the time-varying behavior of links to (and/or from) a document, search engine 125 may score the document accordingly. For example, a downward trend in the number or rate of new links (e.g., based on a comparison of the number or rate of new links in a recent time period versus an older time period) over time could signal to search engine 125 that a document is stale, in which case search engine 125 may decrease the document's score. Conversely, an upward trend may signal a "fresh" document (e.g., a document whose content is fresh--recently created or updated) that might be considered more relevant, depending on the particular situation and implementation.

A cura di:
Stefano Becheroni di [Posizionamento Web](#)
Francesco Iamurri di [Motori e Ricerca](#)

[0069] Il motore può assegnare un punteggio al documento basandosi sulla variazione del comportamento nel tempo dei link che puntano verso il documento stesso (e/o dei suoi link in uscita). Per esempio, una tendenza calante nel numero o nel ritmo di apparizione di nuovi link nel tempo (per esempio basata sulla comparazione del numero o del ritmo di apparizione di nuovi link fra un periodo recente ed uno precedente) può segnalare al motore che il documento è "**stale**" (stantio), nel qual caso il motore abbasserà il punteggio del documento. Contrariamente una tendenza crescente può segnalare un documento "**fresco**" (per esempio un documento il cui contenuto sia stato creato o aggiornato recentemente) che potrebbe essere considerato più rilevante, a seconda della particolare situazione ed implementazione.

[0070] By analyzing the change in the number or rate of increase/decrease of back links to a document (or page) over time, search engine 125 may derive a valuable signal of how fresh the document is. For example, if such analysis is reflected by a curve that is dropping off, this may signal that the document may be stale (e.g., no longer updated, diminished in importance, superceded by another document, etc.).

[0070] Analizzando la variazione nel numero di link che puntano ad un documento o la loro tendenza a crescere o diminuire nel tempo, il motore può trarre un valido segnale di quanto sia "fresco" il documento. Per esempio, se da questa analisi si ottiene una curva calante può significare che il documento è "**stantio**" (per esempio non più aggiornato, diminuito in importanza, superato da un altro documento, ecc.).

[0071] According to one implementation, the analysis may depend on the number of new links to a document. For example, search engine 125 may monitor the number of new links to a document in the last n days compared to the number of new links since the document was first found. Alternatively, search engine 125 may determine the oldest age of the most recent y % of links compared to the age of the first link found.

[0071] L'analisi può dipendere dal **numero di nuovi backlink** (link che puntano ad un documento). Per esempio il motore può monitorare il numero di nuovi backlink negli ultimi " n " giorni e compararlo col numero dei nuovi backlink da quando è nato il documento. In alternativa, il motore di ricerca può determinare l'età del link più vecchio nella più recente percentuale " y " (y %) di backlink comparandola con l'età del primo backlink trovato .

[0072] For the purpose of illustration, consider $y=10$ and two documents (web sites in this example) that were both first found 100 days ago. For the first site, 10% of the links were found less than 10 days ago, while for the second site 0% of the links were found less than 10 days ago (in other words, they were all found earlier). In this case, the metric results in 0.1 for site A and 0 for site B. The metric may be scaled appropriately. In another exemplary implementation, the metric may be modified by performing a relatively more detailed analysis of the distribution of link dates. For example, models may be built that predict if a particular distribution signifies a particular type of site (e.g., a site that is no longer updated, increasing or decreasing in popularity, superceded, etc.).

[0072] A scopo illustrativo, consideriamo " y "=10 (vedi sopra) e 2 documenti (siti web in questo esempio) entrambi individuati per la prima volta 100 giorni fa. Per il primo sito il 10% dei link è stato trovato meno di 10 giorni fa, mentre per il secondo sito lo 0% dei link è stato trovato meno di 10 giorni fa (in altre parole i link erano stati trovati tutti prima). In questo caso il motore attribuisce un punteggio 0.1 al primo sito e un punteggio 0 al secondo sito. I punteggi possono, in seguito, essere scalati appropriatamente. In un'altra implementazione tipica la scala dei punteggi può essere variata per eseguire una analisi più dettagliata della distribuzione delle date dei link. Per esempio possono essere individuati dei modelli che permettano di attribuire una particolare distribuzione delle date dei link ad un

A cura di:
Stefano Becheroni di [Posizionamento Web](#)
Francesco Iamurri di [Motori e Ricerca](#)

particolare tipo di sito (per esempio un sito che non è più aggiornato, con popolarità in aumento o in diminuzione, superato, ecc.).

.....

[0073] According to another implementation, the analysis may depend on weights assigned to the links. In this case, each link may be weighted by a function that increases with the freshness of the link. The freshness of a link may be determined by the date of appearance/change of the link, the date of appearance/change of anchor text associated with the link, date of appearance/change of the document containing the link. The date of appearance/change of the document containing a link may be a better indicator of the freshness of the link based on the theory that a good link may go unchanged when a document gets updated if it is still relevant and good. In order to not update every link's freshness from a minor edit of a tiny unrelated part of a document, each updated document may be tested for significant changes (e.g., changes to a large portion of the document or changes to many different portions of the document) and a link's freshness may be updated (or not updated) accordingly.

[0073] L'analisi può dipendere dai **pesi assegnati ai link**. Ogni link può essere pesato attraverso una funzione che cresce all'aumentare della freschezza del link. La freschezza del link può essere determinata da:

- data di apparizione/cambiamento del link
- data di apparizione/cambiamento del testo associato a un link
- data di apparizione/cambiamento del documento contenente il link

La data di apparizione/cambiamento del documento contenente il link può essere il miglior indicatore della freschezza del link basandosi sulla teoria che quando un documento viene aggiornato un buon link può rimanere invariato se è ancora buono e rilevante. Per non dover aggiornare la freschezza di ogni link ad ogni cambiamento di una parte indipendente o molto piccola di un documento, ogni documento può essere monitorato allo scopo di rilevarne solo i cambiamenti significativi (per esempio, cambiamenti ad una grande parte del documento o cambiamenti a molte parti differenti del documento) e di conseguenza decidere se aggiornare o meno la freschezza del link.

.....

[0074] Links may be weighted in other ways. For example, links may be weighted based on how much the documents containing the links are trusted (e.g., government documents can be given high trust). Links may also, or alternatively, be weighted based on how authoritative the documents containing the links are (e.g., authoritative documents may be determined in a manner similar to that described in U.S. Pat. No. 6,285,999). Links may also, or alternatively, be weighted based on the freshness of the documents containing the links using some other features to establish freshness (e.g., a document that is updated frequently (e.g., the Yahoo home page) suddenly drops a link to a document).

[0074] I link possono essere pesati in altri modi. Per esempio, i link possono essere pesati considerando quanto i documenti che li contengono siano affidabili (per esempio ai documenti governativi può essere assegnato un alto grado di affidabilità). I link possono anche, o alternativamente, essere pesati basandosi su quanto i documenti che li contengono siano "autoritativi" (per esempio i documenti autoritativi possono essere calcolati in modo simile a quanto descritto nel brevetto U.S. Pat. No. 6,285,999). I link possono anche essere pesati basandosi sulla "freschezza" dei documenti che li contengono, usando altre caratteristiche per stabilire tale "freschezza" (per esempio un documento che è aggiornato frequentemente (per esempio la home page di Yahoo) toglie improvvisamente link ai documenti).

A cura di:
Stefano Becheroni di [Posizionamento Web](#)
Francesco Iamurri di [Motori e Ricerca](#)

[0075] Search engine 125 may raise or lower the score of a document to which there are links as a function of the sum of the weights of the links pointing to it. This technique may be employed recursively. For example, assume that a document S is 2 years olds. Document S may be considered fresh if n % of the links to S are fresh or if the documents containing forward links to S are considered fresh. The latter can be checked by using the creation date of the document and applying this technique recursively.

[0075] Il motore di ricerca può aumentare o abbassare il punteggio di un documento verso il quale puntano dei link come funzione della somma dei "pesi" di quei link. Questa tecnica può essere impiegata ricorsivamente. Per esempio, si assuma che il documento "S" abbia 2 anni. Il documento "S" può essere considerato fresco se n% dei link che puntano a "S" sono freschi o se i documenti che contengono link che puntano a "S" sono considerati freschi. Il più recente può essere individuato usando la data di creazione del documento e applicando questa tecnica ricorsivamente.

[0076] According to yet another technique, the analysis may depend on an age distribution associated with the links pointing to a document. In other words, the dates that the links to a document were created may be determined and input to a function that determines the age distribution. It may be assumed that the age distribution of a stale document will be very different from the age distribution of a fresh document. Search engine 125 may then score documents based, at least in part, on the age distributions associated with the documents.

[0076] In un'altra tecnica, l'analisi può dipendere dalla distribuzione delle età associate con i link che puntano ad un documento. In altre parole queste date possono essere determinate e immesse come input in una funzione che determina la distribuzione delle età. Si può assumere che la distribuzione dell'età di un documento "stale" (stantio) sarà molto differente da una distribuzione dell'età di un documento fresco. Il motore di ricerca può allora dare un punteggio al documento basato, almeno in parte, sulla distribuzione delle età associate con i documenti.

[0077] The dates that links appear can also be used to detect "spam," where owners of documents or their colleagues create links to their own document for the purpose of boosting the score assigned by a search engine. A typical, "legitimate" document attracts back links slowly. A large spike in the quantity of back links may signal a topical phenomenon (e.g., the CDC web site may develop many links quickly after an outbreak, such as SARS), or signal attempts to spam a search engine (to obtain a higher ranking and, thus, better placement in search results) by exchanging links, purchasing links, or gaining links from documents without editorial discretion on making links. Examples of documents that give links without editorial discretion include guest books, referrer logs, and "free for all" pages that let anyone add a link to a document.

[0077] Le date di apparizione dei link possono essere usate per rilevare spam, dove i proprietari dei documenti o dei loro colleghi creano link che puntano ai loro stessi documenti allo scopo di incrementare il punteggio assegnato da un motore di ricerca. Un tipico documento "legittimo", acquista backlink più lentamente. Un picco molto grande nella quantità di backlink può segnalare un evento di attualità (per esempio il sito CDC (<http://www.cdc.gov/>) può acquistare backlink molto velocemente dopo un evento dirompente come la SARS) oppure un tentativo di spam verso il motore (per ottenere punteggi maggiori e di conseguenza migliori piazzamenti nei risultati delle ricerche) con scambio link, acquisto link o acquisizione di link da documenti senza revisione editoriale sulla pubblicazione dei link. Esempi di documenti che concedono link senza revisione editoriale includono guest book, referrer logs e le cosiddette pagine FFA (free for all) che lasciano la possibilità a chiunque di aggiungere link a un documento.

A cura di:
Stefano Becheroni di [Posizionamento Web](#)
Francesco Iamurri di [Motori e Ricerca](#)

[0078] According to a further implementation, the analysis may depend on the date that links disappear. The disappearance of many links can mean that the document to which these links point is stale (e.g., no longer being updated or has been superseded by another document). For example, search engine 125 may monitor the date at which one or more links to a document disappear, the number of links that disappear in a given window of time, or some other time-varying decrease in the number of links (or links/updates to the documents containing such links) to a document to identify documents that may be considered stale. Once a document has been determined to be stale, the links contained in that document may be discounted or ignored by search engine 125 when determining scores for documents pointed to by the links.

[0078] In un'ulteriore implementazione, l'analisi può dipendere dalla data in cui un link scompare. La scomparsa di molti link può significare che il documento a cui questi link puntano è "stale" (stantio, non aggiornato da molto tempo, oppure rimpiazzato da un altro documento). Per esempio, il motore di ricerca può monitorare la data in cui uno o più link verso un documento scompaiono, il numero dei link che scompaiono in un dato periodo di tempo, o altri tipi di decrementi del numero dei link verso un documento nel tempo (o degli aggiornamenti/backlink dei documenti che contengono i link) per identificare documenti che possono essere considerati "stale" (stantii). Una volta che un documento è stato considerato "stale" (stantio), i link contenuti nel documento possono essere "scartati" o ignorati dal motore di ricerca quando determina punteggi per i documenti puntati dai link.

[0079] According to another implementation, the analysis may depend, not only on the age of the links to a document, but also on the dynamic-ness of the links. As such, search engine 125 may weight documents that have a different featured link each day, despite having a very fresh link, differently (e.g., lower) than documents that are consistently updated and consistently link to a given target document. In one exemplary implementation, search engine 125 may generate a score for a document based on the scores of the documents with links to the document for all versions of the documents within a window of time. Another version of this may factor a discount/decay into the integration based on the major update times of the document.

[0079] In un'altra implementazione, l'analisi può dipendere non solamente dalle età dei link che puntano a un documento, ma anche dalla loro "dinamicità". Così il motore può pesare in modo diverso documenti che ogni giorno contengono un diverso link in evidenza, nonostante questo link sia molto fresco, rispetto a documenti che vengano costantemente aggiornati e contengano link "stabili" ad altri documenti. In una tipica implementazione, il motore di ricerca può generare un punteggio per un documento basandosi sui punteggi di tutte le versioni dei documenti contenenti i link che lo puntano in una finestra di tempo. In un'altra versione di questa implementazione l'integrazione di questo fattore può essere diminuita basandosi solo sugli aggiornamenti considerati rilevanti del documento.

A cura di:
Stefano Becheroni di [Posizionamento Web](#)
Francesco Iamurri di [Motori e Ricerca](#)