

Optimisation des requêtes temporelles sur le web

Rim Faiz, Nizar Khayati et Khaled Mellouli

LARODEC, IHEC, 2016 Carthage-Présidence, Tunisie

Rim.Faiz@ihec.rnu.tn

Nizar.Khayati@isg.rnu.tn

Khaled.Mellouli@ihec.rnu.tn

La recherche d'informations pertinentes sur le web est considérée comme un nouveau besoin de la société de l'information. Les méthodes de traitement d'informations fondées sur les statistiques ne sont plus suffisantes pour répondre aux besoins des utilisateurs afin de manipuler (rechercher, traduire, résumer...) les informations sur le web. Un constat tend à s'imposer : introduire *plus de sémantique* pour la recherche d'informations pertinentes issues des textes. En effet, lorsque l'utilisateur du web lance sa requête, il s'attend généralement à trouver précisément ce qu'il cherche, c'est à dire trouver « l'information pertinente », sans qu'il soit submergé par un volume de réponses non-maîtrisables et ingérables.

Dans ce travail, nous présentons une approche permettant la recherche de documents sur le web prenant en compte des critères temporels. Notre approche est parente des projets en cours visant à améliorer les résultats des moteurs de recherche [Lawrence *et al.*, 2001] [Glover *et al.*, 2001].

Nous avons constaté que les requêtes recherchant un ou plusieurs événements ayant lieu à une date donnée ou pendant une période de temps bien déterminée, ne permettent pas d'obtenir les résultats attendus. Par exemple : *Les découvertes scientifiques depuis 1940*. Pour cette requête, l'utilisateur veut chercher les découvertes scientifiques **depuis 1940** jusqu'à aujourd'hui et non pas pour l'année 1940, il s'agit de traiter un intervalle de temps.

Notre étude consiste, d'une part, à reformuler des requêtes recherchant des documents textuels ayant un aspect *événementiel* [Faiz et Biskri, 2002] et ce en tenant compte du contexte de la requête [Desclés, 1993], et d'autre part, à obtenir des résultats pertinents répondant précisément à la requête lancée. Une requête de type « **Événement X + [date / intervalle de temps]** » lancée par l'internaute sera analysée et segmentée en deux parties et ce par la détection d'un indicateur temporel (*au mois de, pendant, depuis le, etc.*). Le système proposé, permettra d'effectuer une recherche d'information événementielle avec un niveau de granularité plus fin et ce par le biais d'une transformation de la requête lancée.

Références

- [Desclés, 1993] J.-P. Desclés. L'exploration Contextuelle : Une méthode linguistique et informatique pour l'analyse automatique du texte. *ILN'93*, 339-351, 1993.
- [Faiz et Biskri, 2002] R. Faiz et I. Biskri. Hybrid approach for the assistance in the events extraction in great textual data bases. *IEEE SMC 2002*, 6-9 octobre, Tunisie, 2002.
- [Glover et al., 2001] E. Glover, G. Flake, S. Lawrence, W. Birmingham. Improving category specific web search by learning query modifications. *SAINT-2001*, 23-31, 2001.
- [Lawrence et al., 2001] S. Lawrence, F. Coetzee, E. Glover, D. Pennock, G. Flake. Persistence of Web References in Scientific Research. *IEEE Computer*, 34:26-31, 2001.