

Abstract

Spam pages use various techniques to achieve higher-than-deserved rankings in a search engine's results. Human experts can identify spam, it is too expensive to manually evaluate a large number of pages. Similar to the Trust Rank algorithm [2], Anti-TrustRank algorithm a web ranking method that seeks to overcome the link-based spam and brought the idea that spam website normally form a link farm spam that contains other spam, that is mean spam website has reference the other web spam. How the proposed method is to form the seed set that contains several web that has been known for certain status as spam website and certainly has a high PageRank value. Then from the seed set antitrust values distributed across the web that out of seed set. The number of seed set affects the outcome of the Anti TrustRank Algorithm. More of seed set will get better precision in getting a webspam that are beyond the seed set. In addition, the number of iterations also plays an important role in determining the outcome of the Anti TrustRank algorithm, because the number of iterations is a parameter used in determining the distance between the web in a graph. Experiments on the WebGraph dataset ^[12] show that our approach is very effective at detecting spam pages from a small seed set when compared with TrustRank algorithm.

Keywords : *Anti-TrustRank, TrustRank, seed set, web spam.*