Abstract

As the world-wide web is increasing quickly, now a day's searching information on internet, not only the information also find the truth & related data about topic, so it is complicated to find truth details and relevancy. Unluckily, there is no assurance for the exactness of information on the web. Likewise, different websites often provide inconsistent information on a subject, such as different terms for the same product. We design a general structure for the veracity (trueness) problem, and originate an algorithm called Truth Extractor, this operates the associations among web sites and their information, i.e., a website is truthness if it runs many bits of truth material, and a bit of material is possible to be true if it is
provided by many truthness web sites. In this paper we use Truth Extractor to calculate true
details among variance information, and identify truthness web sites better than the popular
search engines.

References

  2012 IEEE].
- An Analysis of Web Document Clustering Algorithms [ISSN-Volume 1, No. 6, December
  2011].
- Web Crawling Algorithms [International Journal of Computer Science and Artificial
  Intelligence Sept. 2014, Vol. 4 Issue. 3].
- Truth Discovery with Multiple Conflicting Information Providers on the Web [IEEE
  Transactions On Knowledge And Data Engineering, Vol. 20, No. 6, June 2008].
- The MetaCrawler Architecture For Resource Aggregation on the Web [November 8,
  1996].
  e-ISSN: 2278-0661, Volume 17, Issue 2, Ver. V (Mar – Apr. 2015)].
- Information retrieval on Internet using meta search engines: A Review [(Journal of
- String Matching Algorithms and their Applicability in various Applications [(IJSCE) ISSN:
  2231-2307, Volume-I, Issue-6, January 2012].

Index Terms

Computer Science  Information Sciences

Keywords

Search Engine  Data Quality  Truth Extraction Algorithm  Ranking  Clustering