Accepted Manuscript

Cognitive spammer: A Framework for PageRank analysis with Split by Over-sampling and Train by Under-fitting

Aaisha Makkar, Neeraj Kumar

PII: S0167-739X(18)30570-3
DOI: https://doi.org/10.1016/j.future.2018.07.046
Reference: FUTURE 4364

To appear in: *Future Generation Computer Systems*

Received date : 17 March 2018
Revised date : 17 June 2018
Accepted date : 22 July 2018

Please cite this article as: A. Makkar, N. Kumar, Cognitive spammer: A Framework for PageRank analysis with Split by Over-sampling and Train by Under-fitting, *Future Generation Computer Systems* (2018), https://doi.org/10.1016/j.future.2018.07.046

Cognitive Spammer: A Framework for PageRank analysis with Split by Over-sampling and Train by Under-fitting

Aaisha Makkar[a], Neeraj Kumar[b]

[a] *Computer Science Engineering Department, Thapar Institute of Engineering and Technology, Patiala*
[b] *Computer Science Engineering Department, Thapar Institute of Engineering and Technology, Patiala*

Abstract

From the past few years, there is an exponential increase in one of the most popular technologies of the modern era called as Internet of Things (IoT). In IoT, various objects perform the tasks of sensing and computation for providing uninterrupted services (e.g., e-health, e-transportation, security access, etc.) to the end users. In this era, Cognitive Internet of Things (CIoT) is an another paradigm of IoT developed to enhance the capabilities of intelligence in IoT objects where these objects can take independent decisions in any environment. **IoT follows the service oriented architecture (SOA), in which the application layer is the topmost layer. The application layer enables the IoT objects to interact across the globe.** With the power of learning, thinking, and understanding by these objects, can make the information access more accurate and reliable. However, Web spam is one of the challenges while accessing information from the web. It has been observed from the literature review that, search engines are preferred mostly by the people for accessing information. **The efficient ranking by the search engines can reduce the computational cost of information exchange by IoT objects.** Search engines should be able to prevent the spam from being injected into the web. But, the existing techniques for this problem target in finding the spam after its occurrence in search engine result pages. So, in this proposal, we present an intelligent cognitive spammer framework which eliminates the spam pages during the web page rank score calculation by search engines. The framework update the Google's ranking algorithm, *PageRank* in such a way that it automatically prevents link spam by consid-