# CS221 Project 3 Milestone 2

JUN GUO, ID: 83322637
ZHEN CHEN, ID: 85282960

March 11, 2017

## 1    Search Engine

A relevance scoring function is designed to calculate the relevance measure for the given query and documents. As a beginning stage, we only take $cosin$ and $isHead$ into consideration, i.e., $score = a * cosin(q, d) + b * isHead(q, d)$. The parameter $isHead$ is number of terms in query that occur in the given document. For example, when the given query is $q = "machine learning"$ and only $"learning"$ occurs in the document $d_1$, the parameter $isHead(q, d_1) = 1$. The parameter $cosin$ is calculated in terms of $tf - idf$ saved in the index. The parameters $a$ and $b$ are two constants that we can control. In the future, more relevance parameters will be added into the score measure. Moreover, we develop an interface illustrated in Fig. (1) to search our index that retrieves documents according to the relevance score. The logo of our search engine is $googal$. Similar to Google's search engine, we return not only top-$k$ relevant urls but also corresponding heads and
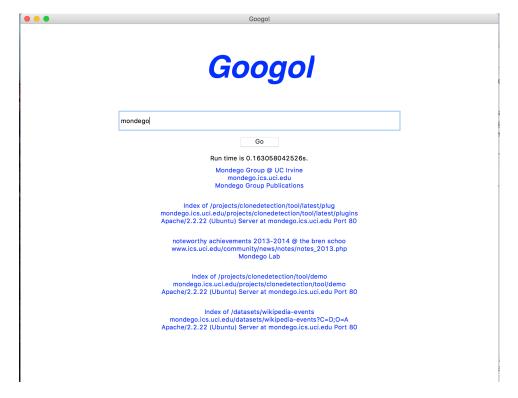


Figure 1: Search Engine Interface

snippets. To evaluate the search engine's performance, we also show the running time, which can be found in Fig. 1.

## 2    Results

The top 5 URLs for each of the 10 test queries are shown in table 1.

Table 1: Simulation Parameters

| Queries | Urls |
|---|---|
| **mondego** | mondego.ics.uci.edu |
| | mondego.ics.uci.edu/projects/clonedetection/tool/latest/plugins |
| | www.ics.uci.edu/community/news/notes/notes_2013.php |
| | mondego.ics.uci.edu/projects/clonedetection/tool/demo |
| | mondego.ics.uci.edu/datasets/wikipedia-events?C=D;O=A |
| **machine learning** | cdb.ics.uci.edu/cgibin/LearningDatasetsWeb.py |
| | archive.ics.uci.edu/ml/machine-learning-databases/uji-penchars/version2?C=S;O=A |
| | archive.ics.uci.edu/ml/machine-learning-databases/wine |
| | archive.ics.uci.edu/ml/machine-learning-databases/pima-indians-diabetes |
| | archive.ics.uci.edu/ml/machine-learning-databases/00211 |
| **software engineering** | www.ics.uci.edu/ ziv/ooad/intro_to_se/sld008.htm |
| | seal.ics.uci.edu/i-htm |
| | seal.ics.uci.edu?p=22 |
| | seal.ics.uci.edu/practical-formal-synthesis-and -automatic-enforcement-of-security-policies-for-android-dsn-2016 |
| | www.ics.uci.edu/ irus//bart/flyers/99/presentations/july99-neno/tsld025.htm |
| **security** | www.ics.uci.edu/ ejw/authoring/washington/jim/sld016.htm |
| | www.ics.uci.edu/ goodrich/teach/ics8/hw |
| | www.ics.uci.edu/ theory/workshop/michael.html |
| | www.ics.uci.edu/ theory/workshop |
| | sprout.ics.uci.edu/projects/privacy-dna/code.html |
| **student affairs** | www.ics.uci.edu/ugrad/sao/SAO_Events.php |
| | students.uci.edu |
| | www.ics.uci.edu/community/alumni/mentor//../../../../about/search/search_sao.php |
| | www.ics.uci.edu/prospective/en/contact/student-affairs/contact/ contact/contact/contact/contact/contact/contact/student-affairs/ contact/student-affairs/contact/contact/student-affairs/contact/student-affairs |
| | www.ics.uci.edu/about/annualreport/2005-06/sao.php |
| **graduate courses** | www.ics.uci.edu/grad/courses/index/faculty/highlights |
| | www.ics.uci.edu/ vitus/ics1c/doc/schedule.html |
| | www.ics.uci.edu/ kay/choosing.courses.html |
| | cml.ics.uci.edu/2014/07/anima-anandkumar-gives-short-course-at-mlss |
| | cml.ics.uci.edu/courses |
| **Crista Lopes** | www.ics.uci.edu/ lopes |
| | www.ics.uci.edu/ lopes/patents.html |
| | sdcl.ics.uci.edu?p=31 |
| | sdcl.ics.uci.edu/category/conference |
| | www.ics.uci.edu/ corps/people.html |
| **REST** | www.ics.uci.edu/ eppstein/cgt/dodgem.html |
| | archive.ics.uci.edu/ml/support/Molecular+Biology+(Promoter+Gene+Sequences) |
| | www.ics.uci.edu/ kay/pubs/bandwagons.html |
| | www.ics.uci.edu/ thornton/ics22/CourseReference.html |
| | www.ics.uci.edu/pub/chimera/software/chimera.el |
| **computer games** | fano.ics.uci.edu/cites/Organization/Int-Computer-Games-Assoc.html |
| | cgvw.ics.uci.edu/tag/ethnography |
| | cgvw.ics.uci.edu/tag/3d-collaboration-environments |
| | cgvw.ics.uci.edu/2012/05 |
| | vcp.ics.uci.edu/content/exploring-potential-computer-games-decentralized-command-and-control |
| **information retrieval** | www.ics.uci.edu/ kay/courses/i141/hw/index.html |
| | www.ics.uci.edu/ vitus/ics1c |
| | www.ics.uci.edu/ kay/courses/i141/refs.html |
| | www.ics.uci.edu/computing/policy |
| | www.ics.uci.edu/computing/services/restore.php |