Truth Discovery with *AllegatorTrack*

http://dafna.qcri.org

D. A. Waguih, N. Goel, H. M. Hammady, L. Berti-Equille

Data Analytics, Qatar Computing Research Institute

معهد قطر لبحوث الحوسبة
Qatar Computing Research Institute
Member of Qatar Foundation عضو في مؤسسة قطر

MOTIVATION

The amount of information available on the Web, social networks, digital media, corporate or scientific applications has been growing at a dramatic pace in recent years. Conflicting information, rumors, erroneous and fake contents can be easily spread across multiple sources, making it hard to distinguish between what is true and what is not.

A major challenge is to determine the veracity of conflicting data claimed by multiple sources.

How do you figure out that a value is true or false? Is a source reliable or not? A lie has been told often enough that it is now considered to be true? How many lying sources are required to introduce confusion in what we knew before to be the truth? To answer these questions, we present ***AllegatorTrack***, a unique truth discovery system.

CONTRIBUTIONS

At QCRI Data Analytics group, our goal is to design an advanced truth discovery system able to extract factual information items claimed by multiple sources, structure them as integrated data, automate fact-checking, and provide clear explanations to end-users on how data veracity has been inferred.

AllegatorTrack is an operational truth discovery system that currently enables the users:

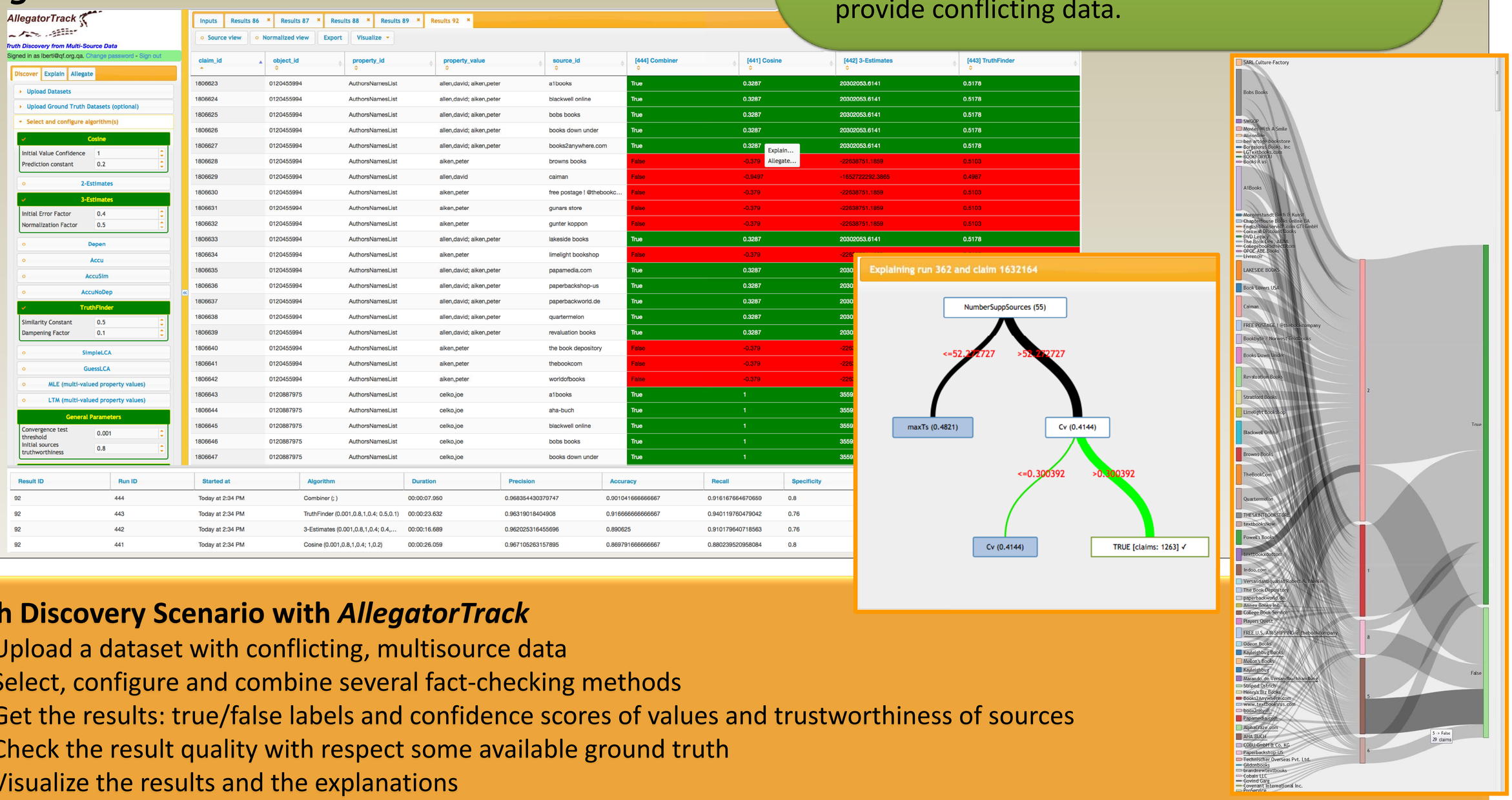
(1) To know the veracity of conflicting data
(2) To compare 12 existing methods for fact-checking
(3) To combine of their results with Bayesian modeling
(4) To generate allegations that can falsify true claims
(5) To get meaningful explanations of truth discovery results.

AllegatorTrack

APPLICATIONS

AllegatorTrack is intended to be used for all data-centered applications where truth and data quality play a critical role: e.g., journalism, cyber security, homeland security, police investigation, data forensics as long as multiple sources provide conflicting data.

AllegatorTrack in Action



Truth Discovery Scenario with *AllegatorTrack*

(1) Upload a dataset with conflicting, multisource data
(2) Select, configure and combine several fact-checking methods
(3) Get the results: true/false labels and confidence scores of values and trustworthiness of sources
(4) Check the result quality with respect some available ground truth
(5) Visualize the results and the explanations
(6) Generate allegations to falsify true claims and test the robustness of truth discovery methods.

معهد قطر لبحوث الحوسبة
Qatar Computing Research Institute
Member of Qatar Foundation عضو في مؤسسة قطر

PUBLICATIONS

- D. A. Waguih, N. Goel, H. M. Hammady, L. Berti-Equille. AllegatorTrack: Combining and Reporting Results of Truth Discovery from Multi-source Data. *The 31st International Conference on Data Engineering (ICDE), Seoul, Korea, 2015.*
- D. A. Waguih, L. Berti-Equille. Truth Discovery Algorithms - An Experimental Evaluation. *Technical Report QCRI, 2014* CoRR abs/1409.6428