

Initial Plan – Assessing Trust in the World Wide Web

CM3203 One Semester Individual Project (40 Credits)

Author: Owain Carpanini

Supervisor: Federico Cerutti

Moderator: David Walker

Project Description

This project will investigate the problem of aggregating sources of information with variable trust level on the World Wide Web. We will create ontologies of trust and of queries/topics, as well as analyse and implement a machine learning solution that can be used alongside our ontologies. In this project, queries are questions relating to a specific topic/source (such as a website), which will present an answer.

The Web provides individuals and organizations with a rich and diverse source of information. Through searches, it is possible to find information on any subject, in all manner of forms – from complex research papers to a ‘layman’ overview of a new scientific discovery on a news website.

However, an open and enduring problem resides in how to estimate ‘trust’ in individual websites, and how to aggregate this ‘trust’ across many websites. For example, we will trust a news article from some websites (such as the BBC or the Financial Times) far more than on other websites (such as the Daily Mail). This could be because of a variety of reasons: The reputation of a website; the language with which articles are written – is it factual or emotive, for example; or more technically, a website’s ‘PageRank’ in Google searches [1].

This project will provide an investigation towards finding a method to estimate trust by providing the first ontology of trust across websites. In addition, we will investigate different approaches for aggregating trust across websites dealing with the same topic, through additional methods such as machine learning. This will come after creating a new ontology.

The project will provide the foundation for investigations into specific case studies, such as aggregating results on controversial topics, such as the efficacy of vaccinations, or the impact that humans have on global warming. Whilst these are case-studies that could take place in future projects, we will look at one or two smaller case-studies in this project. These will show the feasibility of such future studies by the end of this project.

Ethics

This project does not require any ethical approval.

Project Aims and Objectives

As stated in the Project Description, this project will investigate the problem of aggregating sources of information with variable trust level on the World Wide Web. Below are listed the aims and objectives required to complete the project:

- Carry out literature review and case-study identification
 - Gain an overview into the socio-information available on the web, and how this information can be used to assess trust.
 - Due to how broad the project is, we must first identify potential case-studies with which we can investigate the problem on a smaller scale.
- Develop an ontology of the degree of trust in information and sources of information.
 - Use the research from our literature review and existing research in the field of trust in multi-agent systems, to identify the components necessary to assess how trustworthy a piece of information is. This will allow for the creation of our ontology of trust.
 - This will be developed using Protégé.
- Define a small set of queries to be used alongside our case-studies.
 - Ontology of queries? Again, this will be developed in Protégé.
 - These will be used when testing, to ensure our method works.
- Define a set of sources upon which queries can be run.
 - For example, a query relating to politics should be run on a politics/news source as opposed to something unrelated (such as a sports or gossip source).
- Investigate the use of machine learning to work alongside the ontology.
 - Develop a solution that makes use of machine learning.

Work Plan

The final report for this project is due on Friday 5th May. Therefore, I have 15 weeks in which to complete this project.

Below, I have created a work plan for the work I aim to complete each week, based on the aims/objectives defined in the above section. I have also included the main deliverables I will need to complete, as well as a set of review meetings at important times in the project. It is possible that more regular meetings will be required in addition to these reviews.

Week 1 – Week beginning 23rd January

- Write initial plan for project

Week 2 – 30th January

- Deliver initial plan on 30th January, by 23:00
 - Begin literature review and case-study identification.

Week 3 – 6th February

- By Friday 10th Feb, deliver results of literature review and case-study identification:
 - Appropriate case-studies
 - Queries for use in analysis
 - Sources upon which the queries will eventually be executed
 - These are the basis for our ontology
 - The literature review will be the first part of the final report, which will be written/updated continuously over the course of the project.

Week 4 – 13th February

- Begin Development of ontologies:
 - Ontology of Trust
 - Ontology of Queries/topics
 - Define competency questions with which we can assess whether the structure of the ontology is correct.

Week 5 – 20th February

- Review Meeting
- Continue Development of ontologies

Week 6 – 27th February

- Deliver final version of ontologies by Friday 3rd March

Week 7 – 6th March

- Manually evaluate our ontology against competency questions.

This will allow us to check the ontology is correct, and to see that the case-study, and therefore the project is viable on a small scale.

Week 8 – 13th March

- Investigate use of machine learning to assist in assessing trust through our ontology.
- Design a solution that uses the machine learning approach selected.

Week 9 – 20th March

- Implement solution and machine learning approach.
- Review Meeting.

Week 10 – 27th March

- Continue development of solution

Week 11 – 3rd April

- Continue development of solution
- Begin to test solution
 - Initially unit tests

Week 12 – 10th April

- Continue with tests – include user tests as well as unit tests
- Deliver code by Friday 10th April

Week 13 – 17th April

- Finish any incomplete work

Week 14 – 24th April

- Complete draft of report
 - For review in review meeting
- Final Review meeting

Week 15 – 1st May

- Make any changes agreed in review meeting
- Deliver Final Report by Friday 5th May, 23:00

References

[1] J. Gimenez-Garcia, H. Thakkar, and A. Zimmermann, "Assessing Trust with PageRank in the Web of Data," [Online]. Available: http://ceur-ws.org/Vol-1597/PROFILES2016_paper5.pdf. Accessed: Jan. 24, 2017.