Study of conflicts in Wikipedia during massive edition

Contract type: Internship

Level of qualifications required: Bachelor's degree or equivalent

Function: Internship Research

Context

This internship subject is in the context of PEPR eNSEMBLE (https://pepr-ensemble.fr/). The internship will be co-supervised by Nicolas Jullien, LUSSI, Marsouin-LEGO, IMT Atlantique, Nicolas.Jullien@imt-atlantique.fr, Web: https://www.imt-atlantique.fr/fr/personne/nicolas-jullien

The selected candidate will have the possibility to continue the internship with a PhD thesis on a related topic.

Assignment

Motivation and context:

Wikipedia is the result of the largest scale collaboration and has become the dominant information source for an entire generation of Internet users. It offers large-scale data produced by a large number of contributors done with a high velocity on a large number of articles. We are interested in studying the dynamicity of conflicts, i.e., concurrent changes on the same article, and how users resolve these conflicts during collaborative editing tasks in Wikipedia.

MediaWiki does not automatically merge parallel modifications on the same page. This becomes critical when changes have to be performed very quickly such as during breaking news when several users contribute at the same time to the same Wikipedia page. The conflict resolution mechanism in Wikipedia requires that a user manually merges two concurrent versions of the same page. If two users are concurrently editing the same article, when they try to save their changes, the changes of only one user are published. The other user will be presented with two versions of the wiki page: the one that the user tried to publish and the recently published version. Conflicts have to be manually resolved by users by retyping or copying and pasting the changes they performed in the last version that was published. Moreover, users are unaware of other concurrent updates on the same article until they try to publish their own changes. Resolving conflicts might become very tedious, and this can be critical when a change has to be published almost instantly. Our hypothesis is that this is an inefficient conflict resolution mechanism, frustrating for users and which decreases the collaboration quality.

Goal:

We aim to analyse the frequency of user editions on the same article in Wikipedia and show that at certain moments of massive edition of an article the collaborative editing model offered by MediaWiki is not well suited. Based on the results obtained we will propose implications for design for the collaborative editing in Wikipedia.

Main activities

- Study the existing literature on conflicts in Wikipedia (e.g. [1,6]), in real-time collaborative editing (e.g. [2,4,5]) and in version control systems (e.g. [3]).
- Compute the number of user editions on the same articles during a time window (15s, 30s, 1 min) and analyse the frequency of concurrent editions and their distribution on different articles.
- Analyse whether concurrent editions that are done during a time window on a same article are close in space such as in the same paragraph, sentence or word
- Analyse discussion pages to detect whether users complain about the management of conflicts particularly related to periods with a lot of contributions
- Discuss implications for design, i.e. whether adopting real-time collaborative editing for the periods with massive changes would be a more appropriate approach for encouraging user contributions.

Bibliography:


**Skills**

- Ideally for an internship during the Master 2 degree in Computer science / Applied mathematics with an experience in computer networks
- Engineering or Master 1 degree in Computer science / Applied mathematics with an experience in computer networks.
- Theoretical expertise: distributed systems, data analysis
- Good collaborative and networking skills, excellent written and oral communication in English
- Good programming skills
- Strong analytical skills

**Benefits package**

- Subsidized meals
- Partial reimbursement of public transport costs
- Leave: 7 weeks of annual leave + 10 extra days off due to RTT (statutory reduction in working hours) + possibility of exceptional leave (sick children, moving home, etc.)
- Possibility of teleworking (after 6 months of employment) and flexible organization of working hours
- Professional equipment available (videoconferencing, loan of computer equipment, etc.)
- Social, cultural and sports events and activities
- Access to vocational training
- Social security coverage

**Remuneration**

4.05 per hour of training

**General Information**

- **Theme/Domain**: Data and Knowledge Representation and Processing
- **Data production, processing, analysis (BAP D)**
- **Town/city**: Villers lès Nancy
- **Inria Center**: Centre Inria de l'Université de Lorraine
- **Starting date**: 2024-03-01
- **Duration of contract**: 6 months
- **Deadline to apply**: 2024-01-21

**Contacts**

- **Inria Team**: COAST
- **Recruiter**: Ignat Claudia-lavinia / claudia.ignat@inria.fr

**About Inria**

Inria is the French national research institute dedicated to digital science and technology. It employs 2,600 people. Its 200 agile project teams, generally run jointly with academic partners, include more than 3,500 scientists and engineers working to meet the challenges of digital technology, often at the interface with other disciplines. The Institute also employs numerous talents in over forty different professions. 900 research support staff contribute to the preparation and development of scientific and entrepreneurial projects that have a worldwide impact.

**Warning**: you must enter your e-mail address in order to save your application to Inria. Applications must be submitted online on the Inria website. Processing of applications sent from other channels is not guaranteed.
Instruction to apply

Defence Security:
This position is likely to be situated in a restricted area (ZRR), as defined in Decree No. 2011-1425 relating to the protection of national scientific and technical potential (PPST). Authorisation to enter an area is granted by the director of the unit, following a favourable Ministerial decision, as defined in the decree of 3 July 2012 relating to the PPST. An unfavourable Ministerial decision in respect of a position situated in a ZRR would result in the cancellation of the appointment.

Recruitment Policy:
As part of its diversity policy, all Inria positions are accessible to people with disabilities.