Chair of Network Architectures and Services Department of Informatics Technical University of Munich

## Thesis B.Sc.

## Similarity Detection for Databases

## Motivation

Topic

Your Task

**Requirements** 

Practical courses conducted at the chair rely on the iLab-Labsystem [1] for the submission of lab answers of course participants. A major learning goal of the courses is to teach participants the ability to explain concepts

rooterory incornectab_	
	[log in] Non. 07 Jun 2021 11:13:42 +0250 [ en
guest in ister	Welcome to the iLab 2021ss!
m schedule	The ILab: Build your own Internet is a lab course offered by the Network Architectures and Services group at Technische Universität München.
b: Build your own In	In the iLab you learn to build networks equipped with all the important network protocols making up the internet of today. When you finish the course, you will be able to build up a working internet with all required technologies on your own!
urse Content	Course outline
past courses	The theoretical contents of the networking lectures are explored in practical exercises. You will set up all these services (and more) by yourself and learn how they work in detail:
20ss 19ws 19ss 10ws 18ss 17ws	We will see how machines can be provisioned with network configuration (e.g. 5LAAC). With dynamic routing protocols, such as OSPF, networks are made realinet against failures. DNS and DNSSEC are set up to provide name resolution from the not zore downwards. We will look at security protocols and see home TLS can be used security in the popular controllision with ITTL build prefix, ere will learn to protect all if traffic. To hole communication meta data, a protection not usually differed by security protocols, we create an onion notifying network wange to with the experiment of the traffic around the security protocol, we create an onion notifying network wange to with the experiment and use up more secure protocols insets. JFm in his pradhenticator, we

and results in their own words. Currently, teaching assistants check this through manual work, i.e. check given sources and compare answers within a class. However, comparing answers to answers from previous course iterations is not feasible. The problem is further amplified as the pool of given answers increases considerably with each term.

The goal of this thesis is to develop and implement a mechanism to compare answers from different classes concerning their similarity. Therefore, methods for plagiarism-checks and methods for similarity checks for free-text answers need to be analyzed. Based on the insights, a scalable mechanism for comparing answers of iLab participants to the answer pool of previous terms should be created. Functionality for the pipeline should:

- automatically identify similar questions between lab instances
- efficiently and scalably fetch and merge data from different databases
- compare large amounts of strings with respect to their similarity
- Analysis of plagiarism detection tools
- Analysis of the iLab lab system
- Implementation of an own scalable answer-comparison pipeline running in conjunction to the iLab-lab system
- Ability to write efficient and maintainable code
- Participated in any of the iLab courses (preferably iLab1)

## Sources

[1] iLab labsystem, https://github.com/m-o-p/labsystem

Contact

Lars Wüstrich wuestrich@net.in.tum.de Florian Wiedner wiedner@net.in.tum.de http://go.tum.de/080755



