Efficient Graph Kernels for RDF data using Spark

Bernhard Japes[1] and Shinho Kang[2]

[1] Informatik III, Universität Bonn, Germany
bernhard.japes@uni-bonn.de
[2] Informatik III, Universität Bonn, Germany
TODO

Abstract

In this paper we study the application of graph kernels for RDF data using the popular Apache Spark[1] engine in combination with the SANSA-Stack[2] data flow utilities. We focus on an implementation of the Intersection Tree Path (ITP) Kernel, published by Gerben Klaas Dirk de Vries and Steven de Rooij in [1], that is based on the concept of constructing a tree for each instance and counting the number of paths in that tree.

TODO: Add further information about implementation and/or results

1 Introduction

2 Approach

3 Implementation

4 Evaluation

5 Conclusion

5.1 Project timeline

5.2 Further ideas

References

[1] Gerben Klaas Dirk De Vries and Steven De Rooij. A fast and simple graph kernel for rdf. In *Proceedings of the 2013 International Conference on Data Mining on Linked Data - Volume 1082*, DMoLD'13, pages 23–34, Aachen, Germany, Germany, 2013. CEUR-WS.org.

[1] http://spark.apache.org
[2] http://www.sansa-stack.net