A Distributed Architecture for Rule Engine to Deal with Big Data

Siyuan Zhu[1], Second Author[1], Third Author[1]

School of Computer Science and Technology

Beijing University and Telecommunications ,China

thinkingfar@163.com, {hhuang,zlei}@bupt.edu.cn

***Abstract*—Rule engine, which acknowledges facts and draws conclusions by repeatedly matching facts with rules, is a good way of knowledge representation and inference. However, because of its low computational efficiency and the limitation of single machine's capacity, it cannot deal well with big data. As traditional MapReduce architecture can only address this problem in certain conditions, we have made some improvements and therefore proposed a distributed implementation of the rule engine using MapReduce-based architecture. It is designed to deal with a large amount of data in a parallel and distributed way by using a computing cluster that consists of multiple machines, on which certain part of the Rete algorithm would be operated. In the phase of splitting rules and the Rete-net, Apriori algorithm is also improved and adopted so as to gain a better system performance. This paper not only describes details of the design and its implementation, but also shows its high performance through several experiments.**

***Keyword*—rule engine, big data, rete algorithm, map-reduce, apriori algorithm.**

Siyuan Zhu is a master student in computer science and technology at the Beijing University of Posts and Telecommunications of China (BUPT). He was born in China in 1990. He received his bachelor degree at BUPT in 2013. His research interests include cloud computing, rule-based computing and internet of things.

Hai Huang is a lecturer in the School of Compuer Science and Technology at the Beijing University of Posts and Telecommunications of China (BUPT). He was born in China in 1979. He received his P.H.D. degree of engineering in computer science at BUPT. His research interests include internet of things , cloud computing and service software.

Lei Zhang is a pofessor in the School of Compuer Science and Technology at the Beijing University of Posts and Telecommunications of China (BUPT). He was born in China in 1962. He received his Master degree of engineering in computer science at BUPT in 1988. His research interests include distributed systems, cloud computing and internet of things. He is the author of a great deal of research studieds published at national and international journals, conference proceedings as well as book chapters