4월 1주차 연구노트

RDF datasets 에서의 SPARQL 쿼리 최적화 관련 논문 조사.

관련논문 : [An analytical approach for query optimization based on hypergraph](http://ieeexplore.ieee.org/document/7207087/)

([Sangeeta Sen](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Sangeeta%20Sen.QT.&newsearch=true); [Anisha Agrawal](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Anisha%20Agrawal.QT.&newsearch=true); [Ankit Rathi](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Ankit%20Rathi.QT.&newsearch=true); [Animesh Dutta](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Animesh%20Dutta.QT.&newsearch=true); [Biswanath Dutta](http://ieeexplore.ieee.org/search/searchresult.jsp?searchWithin=%22Authors%22:.QT.Biswanath%20Dutta.QT.&newsearch=true) [2015 12th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)](http://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=7180538) Year: 2015)

**쿼리 파싱과 서술어 크기에 맞게 SPARQL 쿼리 재배열**

SPARQL 쿼리가 시스템에서 실행될 때, 쿼리를 파싱하고 hyper-edge 크기 오름차순으로 쿼리를 배열하고, 쿼리 경로를 설정하기 위해 정렬된 순서를 사용한다. 재배열 알고리즘은 다음과 같다.

|  |
| --- |
| Input: data, q: list storing the match patterns of input query.  Output: sizeEdge: dictionary storing size of each hyperedge, predicatelist: list storing rearranged match patterns and predicate according to hyperedge size. |
| sizeEdge <- 0, predicate <- 0, predicatelist <- 0, line <- 0, pred <- 0;  for all predicate in data.keys do  sizeEdge[predicate] <- data[predicate]:length=2  end for  for each line in q do  pred <- line {predicate}  predicatelist <- predicatelist  {line, pred, sizeEdge[pred]}  8: end for  sort predicatelist based on sizeEdge[pred]  10: return sizeEdge, predicatelist |