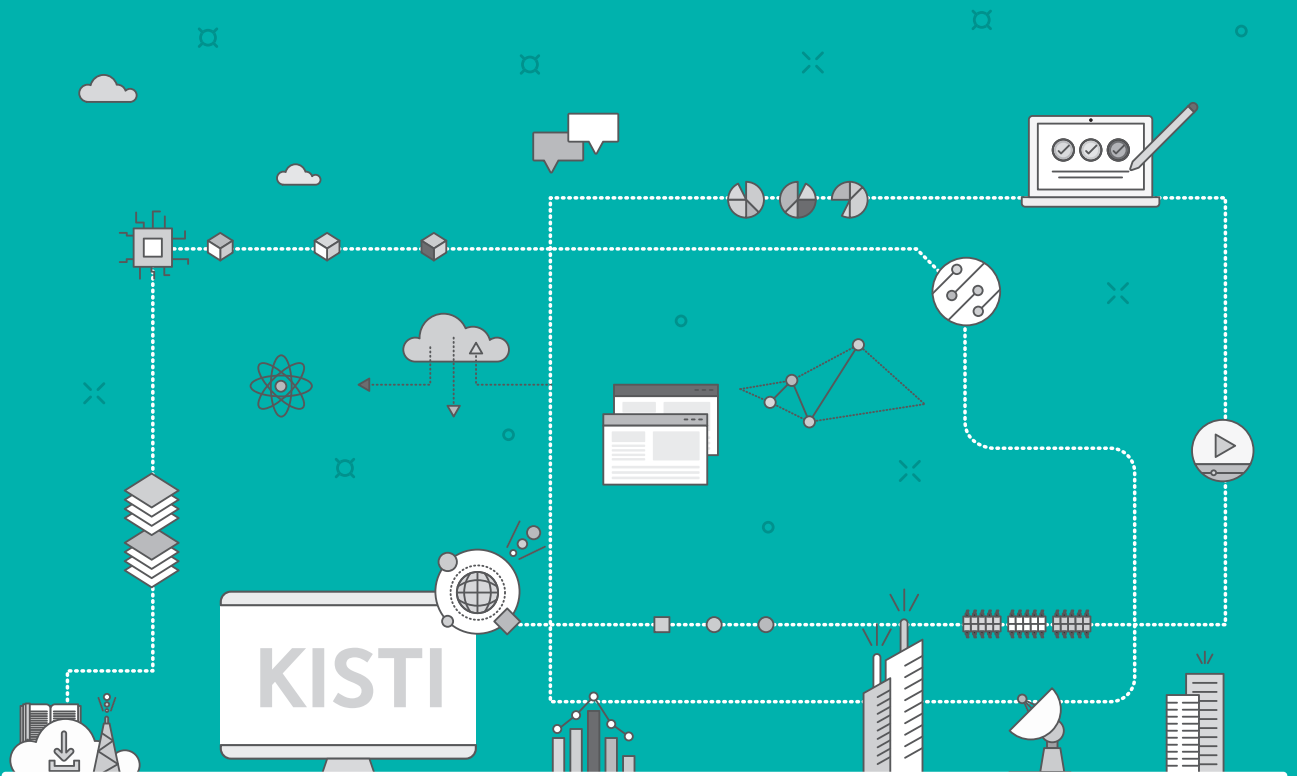


대용량 네트워크 분석용 가시화 시스템 및 패키징 기술





Keyword 빅데이터 분석, 네트워크 군집화, 네트워크 탐색
연구책임자 서동민
기술 완성단계(TRL) 5단계(시제품제작 및 성능 평가)

기술개요

최근 증가하고 있는 빅데이터 분석 수요에 대응하기 위해 대용량 데이터를 빠르고 정확하게 다차원 분석할 수 있는 네트워크 군집화 및 네트워크 탐색 알고리즘, 다차원 분석 플랫폼 패키지

기존 기술의 문제점

선진국 대비 낮은 클라우드 컴퓨팅 산업 수준 및 기술 경쟁력

- 미국 대비 66%의 낮은 산업 경쟁력, 미국 대비 4.1년의 기술격차 존재, 안정성의 이유로 시장의 외산 솔루션 비율이 높음

가상화 솔루션, 빅데이터 플랫폼 등 컴퓨팅 원천기술 미확보로 해외기술에 대한 의존성 심화

- 국내기업은 원천기술 기반 응용개발에 주력, 해외 원천기술에 대한 라이선스 비용으로 많은 금액을 지불

외산 선호현상과 국내제품에 대한 정보 부재, 레퍼런스 부족으로 낮은 국내제품 보급률

- 공공기관 클라우드 구축 공고 시 외산/특정 제품 공고로 국내기업 진입 차단 사례 발생

기술 내용 및 차별성

기존 대용량 네트워크 분석 시스템 대비 정확도 및 실행속도 향상

기술 내용

- 네트워크 군집화 기술 : 네트워크 압축 및 분산 병렬 처리 기술을 활용한 군집화 기술
 → coarsening 기법활용, 멀티코어 이용 병렬처리
- 네트워크 탐색 기술 : 노드 우선순위 및 분산 병렬처리 기술을 활용한 탐색 기술
 → 탐색시작 정점 기준 문제 분할, 스레드 할당, 중복출력 제거

차별성

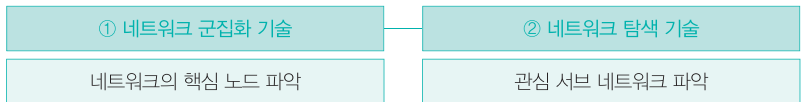
- 네트워크 군집화 기술
 - MCL 대비 정확도 110% 이상, 실행 속도 120% 이상 향상
 - 대용량 네트워크 As-Skitter에 대한 군집화 지원
- 네트워크 탐색 기술
 - VF2 대비 정확도 유지, 실행속도 110% 향상
 - 대용량 네트워크 As-skitter에 대한 탐색 지원

※ As-Skitter : 스탠포드 대학에서 공개한 네트워크 데이터, 대용량 네트워크 데이터 처리 기술의 성능 평가를 위해 표준 실험 데이터로 활용



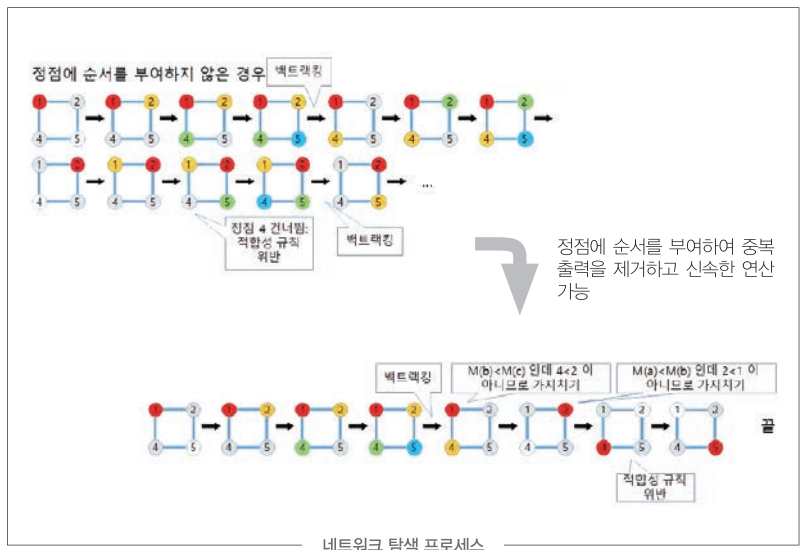
주요기술 구성 및 구현방법

| 주요기술 구성 |



| 구현방법 |

- 네트워크 군집화 기술 : 네트워크의 핵심 노드 파악
 - 그래프 분할기법인 METIS의 coarsening 기법 활용 MCL 기반 알고리즘의 성능 향상
 - ※ Coarsening : 여러 노드를 하나의 슈퍼 노드로 합치는 방식
 - ※ MCL : 오하이오 대학의 군집화 알고리즘, 무료 소프트웨어로 현재 가장 많은 사용자 보유
- MCL과 그의 변형기법들의 핵심연산인 행렬곱셈, Inflate, Prune 등의 병렬화로 확장성 향상
 - ex) 대규모 생물학적 네트워크에서 질병(치매)의 핵심 인자의 신속한 파악 가능
- 네트워크 탐색 기술 : 네트워크에서 관심있는 서버네트워크 파악
 - 데이터 그래프와 패턴 그래프의 정점에 각각 순서를 부여하여 VF2 알고리즘의 매치 결과 중복 출력 문제 해결
 - ex) 특정 질병에 효과가 있는 약물 빠르게 파악 가능
 - ※ VF2 : 로마 사피엔자 대학의 네트워크 탐색 기술



기술/시장 동향

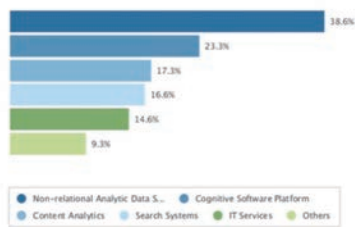
| 기술 동향 |

- Cytoscape
 - 많은 커뮤니티를 통해 네트워크 군집화, 탐색, 가시화 기능을 제공하며 분석 툴도 모듈 별로 제공
 - 노드와 간선수의 총합이 70,000~150,000 일 경우 1G 이상의 메모리를 요구하여 대용량 네트워크 분석 기능은 취약
- MCL(Markov Clustering)
 - 오하이오 주립대학의 네트워크 군집화 알고리즘으로 플로우 기반 그래프 클러스터링 방식 사용
 - 간단하고 직관적인 알고리즘이나 확장성이 떨어지며 세부 클러스터들이 많이 생성되는 문제점 보유
- VF2
 - 로마 사피엔자 대학의 네트워크 탐색 기술로 대표적인 그래프 매칭 알고리즘
 - 패턴 그래프 순서로 정점 매칭 시 중복 출력되어 후처리 작업 및 추가 연산이 필요

| 시장 동향 |

- 빅데이터 분석 분야의 사회·경제적인 기술 수요 급증
 - 최근 4차 산업혁명에 대한 관심이 고조되고 클라우드 컴퓨팅 기술이 큰 주목을 받으면서 빅데이터 분석에 대한 수요 급증
 - 빅데이터 분석을 통해 고객만족도를 향상시키고 내부 업무의 효율성을 극대화 하려는 기업들의 수요 급증
- 세계 빅데이터 시장규모는 2020년까지 매년 11.9% 성장 예상, 국내 시장규모는 10% 증가 전망
 - 2016년 세계 빅데이터 분석 시장은 전년대비 12.4% 성장하여 1,508억 달러 전망
 - 2020년까지 연평균 11.9% 성장하여 2020년에는 2,100 달러 규모 전망
- 전체 데이터베이스 시장 내 그래프 데이터베이스 분야의 가장 높은 성장률 전망(40.8%)
 - 전체 데이터베이스 시장은 2016년 전년도 대비 7.3% 증가, 2019년까지 매년 6~9% 증가하여 2020년에는 약 52조원 예상
 - NoSQL 데이터베이스 시장은 2020년까지 연평균 35.1% 수준의 성장 예상, 그래프 데이터베이스가 차지하는 비중은 NoSQL 시장의 3% 수준으로 연평균 40.8% 성장 예상

Top Technology Category Based on 5 Year CAGR (2015 - 2020)



Source: IDC Worldwide Semianual Big Data and Analytics Spending Guide, 2016Q2

활용분야 및 권리현황

| 기술활용분야 |

기술 수요처	적용처
생명과학 분야 민간 기업 및 연구소	바이오 빅데이터 네트워크 분석
병원, 의료 기관, 제약기업	질병 효과물질 및 약물 탐색, 신약 재창출
정보서비스 업체	빅데이터 분석, 소셜 네트워크 분석

| 권리현황 |

- 국내 등록특허 2건

발명의 명칭	특허번호	비고
네트워크 클러스터링 방법 및 장치	10-1700108	등록
그래프 매칭 방법 및 장치	10-1747854	등록

추가기술정보

기술분류	빅데이터 분석
시장전망	그래프데이터베이스 분야는 데이터 분석 시장 중 가장 높은 40.8%의 시장 성장 예상
기술문의	서동민 책임연구원 (과학기술데이터본부) 042-869-1796 dmseo@kist.re.kr 윤신혜 행정원 (성과확산실 기술이전 담당) 042-869-1832 shyoon@kist.re.kr