

네트워크 탐색 기술을 기반으로 한 무역 거래 위험 요소 적발 시스템 개발

Development of a System to Detect the Risk Factors of Trade based on Network Search Technology

서 동 민*,**, 김 재 수*,**, 송 정 아***, 박 문 일***
한국과학기술정보연구원*,
과학기술연합대학원대학교**, 데이터월드***

Seo Dongmin*,**, Kim Jaesoo*,**, Song Jeong a***,
Park Moon il***

Korea Institute of Science and Technology
Information*, Univ. of Science & Technology**,
DataWord Co.***

요약

빅데이터 분석에 활용되는 원천 데이터는 네트워크 형태이며, 최근 소셜 네트워크 분석을 통한 효과적인 상품 광고, 핵심 유전자 발굴, 신약 재창출 등 다양한 영역에서 네트워크 분석 기술이 사회와 인류에게 가치 있는 정보를 제공할 수 있는 가능성을 제시하면서 네트워크 분석 기술의 중요성이 부각되고 있다. 또한, 세계화와 정보통신기술의 급격한 발전으로 빠르게 변화하는 무역 환경 속에서 신속하고 정확한 무역 거래에 대한 안전 관리의 요구가 점차 증가하고 있다. 그래서 본 논문에서는 네트워크 탐색 기술을 기반으로 한 무역 거래 위험 요소 적발 기술을 제시했다.

I. 서론

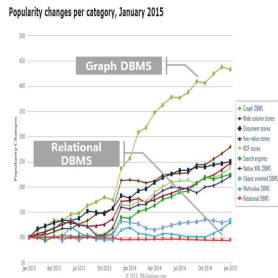
네트워크는 하나 이상의 데이터를 데이터 간 관계를 기반으로 연결시킨 자료구조를 의미하고, 네트워크 구조를 분석하면 밝혀지지 않은 데이터 간 특성 및 패턴을 발견할 수 있기 때문에 대부분의 빅데이터 분석은 네트워크 분석을 기반으로 한다[1]. 최근 빅데이터가 다양한 분야에서 활용되면서 방대한 양의 네트워크 데이터가 생성되고 있고, 소셜 네트워크 분석을 통한 효과적인 상품 광고, 핵심 유전자 발굴, 신약 재창출 등 다양한 영역에서 네트워크 분석 기술이 사회와 인류에게 가치 있는 정보를 제공할 수 있는 가능성을 제시하면서 대용량 네트워크 데이터를 효율적으로 분석할 수 있는 네트워크 분석 시스템의 중요성이 증가하고 있다.

관세청은 국가경제의 발전을 위하여 급증하는 수출입 물품과 여행자 관리와 밀수 단속, 유해식품 및 유해화학물질 불법반입 단속, 불법외환거래 단속 및 자금세탁 단속 등 수출입통관질서를 시스템적으로 관리할 수 있도록 IRM-PASS를 구축했고 지속적으로 시스템의 안전과 속도 고도화를 위한 연구를 수행하고 있다. 하지만, 세계화와 정보통신기술의 급격한 발전으로 국제 통상 및 경제 활동과 기업 활동 환경이 국제 무역 환경에서 자유 무역 환경으로 급변하고 있고, 이러한 무역 환경의 변화는 무역 규모의 증대와 위험 요소의 증가가 동시에 수반되어 신속하고 정확한 무역 거래의 안전 관리에 대한 요구가 급증하고 있다[2].

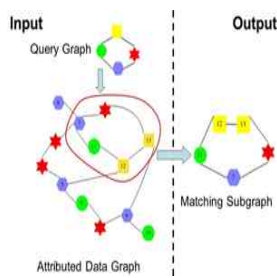
본 논문에서는 다양하고 변화무쌍한 무역 거래와 관련된 각종 위험에 선제적으로 대응할 수 있는 환경 마련을 위해, 네트워크 탐색 기술을 기반으로 한 무역 거래 위험 요소 적발 시스템을 개발했다. 본 논문에서 개발한 시스템은 무역 거래 정보의 연계를 통한 연관 관계 중심의 위험 정보 분석 및 변화 추이 모니터링 기능을 제공한다.

II. 관련연구

네트워크 분석 연구가 최근의 주요 연구주제로 급부상하면서 2013년 이후 다양한 네트워크 분석 기능을 제공하는 그래프 데이터베이스에 대한 수요가 급성장하고 있다. 일례로, 그림 1에서와 같이 그래프 데이터베이스에 대한 수요가 급증하면서 전체 데이터베이스 시장 내에서 그래프 데이터베이스가 가장 높은 성장률을 보이며, 연평균 40.8%의 성장률을 보이고 있다. 또한, 그림 2와 같이 네트워크 탐색(subgraph-matching) 기술은 사용자가 원하는 일부 네트워크를 전체 네트워크에서 모두 찾는 기술로, 많은 데이터 내에서 사용자가 원하는 데이터를 빠르게 파악하는데 폭넓게 활용되고 있다. 본 논문에서는 무역 거래와 관련된 수출입 공급망 및 공급망간 연계 분석을 통해 수출입 무역 거래 패턴을 파악하고 우범 거래 패턴 분석을 통해 관련 공급망 및 유사 거래 패턴에 대한 선제적인 대응 수단을 마련하기 위해서 네트워크 탐색 기술을 사용했다.



▶▶ 그림 1. 그래프 데이터베이스 시장동향

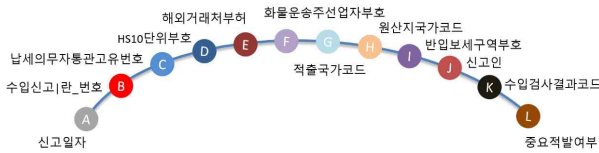


▶▶ 그림 2. 네트워크 탐색 기술 개념도

III. 무역 거래 위험 요소 적발 시스템

1. 무역 거래 위험 요소 네트워크 구축

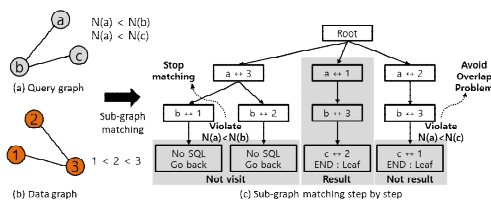
IRM-PASS 내 관계형 데이터베이스에는 수출입 공급망 및 공급망간 연계 정보가 구축되어 있고 무역 거래와 관련해 관련 정보가 지속적으로 갱신되고 있다. 본 논문에서는 네트워크 탐색 기술을 기반으로 한 무역 거래 위험 요소 적발 기능을 개발하기 위해서, 무역 거래 위험 요소 적발과 관련된 중요 정보인 (A)신고일자, (B)수입신고/관번호, (C)납세의무자동관고유번호, (D)KS 10단위부호, (E)해외거래처부호, (F)화물운송주선업자부호, (G)적출국가코드, (H)원산지국가코드, (I)반입보세구역부호, (J)신고인, (K)수입검사결과코드, (L)중요적발여부를 그림 3과 같은 정보간 관계를 가지는 네트워크로 구축했다.



▶▶ 그림 3. 위험 요소 적발을 위해 구축한 네트워크 구조

2. 위험 요소 적발을 위한 네트워크 탐색 기술

제안하는 네트워크 탐색 기술은 데이터 그래프에 대한 정점 번호 부여, 질의 그래프에 대한 정점 번호 부여 그리고 정점 번호를 기반으로 한 데이터 그래프 내 질의 그래프 탐색을 수행한다. 그리고 그래프 탐색 성능 향상을 위해 데이터 그래프를 구성하는 각 정점에 대한 질의 그래프 탐색은 최대 가용 스레드 수를 고려해서 병렬 수행된다. 데이터 그래프에 대한 정점 번호 부여는 그림 4(b)에서와 같이, 차수가 가장 작은 정점에 가장 큰 번호가 부여되며, 동일 차수를 가지는 정점들에 대해서는 임의의 번호로 부여된다. 질의 그래프에 대한 정점 번호 부여는 그림 4(a)에서와 같이 질의 그래프 탐색시, 먼저 방문된 정점에 가장 작은 번호가 부여되며, 나중에 방문된 정점일수록 큰 번호가 부여된다.

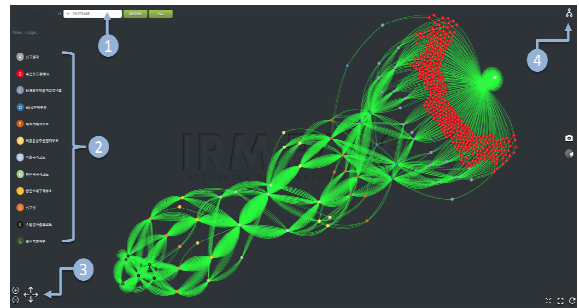


▶▶ 그림 4. 위험 요소 적발을 위한 네트워크 탐색 기술 개념도 [3]

정점 번호는 데이터 그래프 내 질의 그래프 탐색 시 발생할 수 있는 중복 탐색을 회피해 그래프 탐색 성능을 향상시키기 위해서 사용된다. 일례로, 그림 4(a)의 1<2와 그림 4(b)의 N(a)<N(b)를 통해, 그림 4(c)에서 정점 a와 정점 2가 매칭후 정점 b와 정점 3이 매칭될 수 있다. 하지만, 그림 4(a)의 1<3 때문에 정점 c와 정점 1이 매칭될 수 없다는 것을 사전에 판단할 수 있기 때문에 불필요한 질의 탐색을 회피할 수 있게 된다.

3. 무역 거래 위험 요소 적발 시스템 가시화 도구

그림 5는 개발한 무역 거래 위험 요소 적발 시스템에 대한 네트워크 가시화 도구를 통해 가시화된 무역 거래 위험 요소 네트워크를 보여주고 있다. 개발한 가시화 도구는 HTML5, CSS3, Canvas 등의 최신 웹 기술을 적용하여, ActiveX 또는 Flash 등의 별도 플러그인 없이 사용 가능하기 때문에 웹 브라우저의 제한 없이 사용이 가능하다. 그리고 그림 5의 주요 아이콘들의 기능은 다음과 같다. ①은 전체 네트워크에서 사용자가 관심을 갖는 특정 정점을 키워드와 ②의 유형을 통해 검색할 수 있는 기능을 제공한다. ③은 네트워크에 대한 Zoom In/Out과 이동 기능을 제공하며, ④는 사용자가 정점에 대한 키워드, 유형 그리고 와일드카드를 통해 패턴 질의를 생성할 수 있는 기능을 제공한다.



▶▶ 그림 5. 개발한 무역 거래 위험 요소 적발 시스템 UI

감사의 글

본 연구는 한국과학기술정보연구원의 「연구데이터 공유·확산체제 구축(K-18-L11-C03-S01)」 사업으로부터 지원을 받아 수행된 연구임.

참고 문헌

- [1] 서동민, 유석중, 이민호, “대용량 네트워크 압축 기반 클러스터링 알고리즘 개발”, 한국콘텐츠학회 2016 춘계 종합학술대회, 제14권, 제1호, pp.53-54, 2016.
- [2] Pyung K., Dongmin S., Hanmin J., Kyungsun K. and In-Chae Y., “The Relation Tracking Service of Customs Service based on Semantic Web”, International Conf. on EEE, pp.554-555, 2011.
- [3] Dongmin Seo, Ha-Myung Park, Seok Jong Yu, Min-Ho Lee, U Kang, “Efficient Subgraph Matching on Large Networks”, ICCV, pp.275-276, 2016.