

# 동적 다변량 그래프의 연속적 분석을 위한 질의 모델 설계 및 구현

배예찬<sup>†</sup> · 함도영<sup>†</sup> · 김태양<sup>†</sup> · 정혜진<sup>††</sup> · 김동윤<sup>†††</sup>

## 요 약

본 연구에서는 동적 다변량 그래프 데이터의 연속적 분석이 가능한 질의 모델을 설계 및 구현하였다. 먼저, 질의 모델을 판별함수 설정과 시간에 따른 통합 방법 선택의 두 단계로 설계하고, 질의 패널, 그래프 시각화 패널, 속성 패널로 구성된 질의 시스템으로 구현하였다. 또한, 그래프 표현에는 노드-링크 다이어그램과 Force-Directed Graph Drawing 알고리즘을 이용하였으며, 질의 결과로 선택된 대상들에 효과를 적용하여 사용자가 시각적으로 구분할 수 있도록 처리하였다. 마지막으로, 세계 소형 무기 거래량 데이터를 이용하여, 본 연구에서 설계한 동적 다변량 그래프 질의 모델을 검증하였다. 본 연구는 동적 그래프의 연속적 분석이 가능한 새로운 질의 모델을 설계하는 것을 통해, 기존 모델이 동적 그래프를 시점별로 이산적으로만 분석할 수 있는 한계를 개선하였다는데 의의가 있다. 본 연구는 추세 분석이나, 복잡하게 네트워크 해석 등 동적 그래프를 사용하는 연구에 기여할 수 있을 것으로 기대된다.

**주제어** : 동적 데이터, 동적 그래프, 다변량 그래프, 질의 모델

## A Query Model for Consecutive Analyses of Dynamic Multivariate Graphs

Yechan Bae<sup>†</sup> · Doyoung Ham<sup>†</sup> · Taeyang Kim<sup>†</sup> · Hayjin Jeong<sup>††</sup> · Dongyoon Kim<sup>†††</sup>

### ABSTRACT

This study designed and implemented a query model for consecutive analyses of dynamic multivariate graph data. First, the query model consists of two procedures; setting the discriminant function, and determining an alteration method. Second, the query model was implemented as a query system that consists of a query panel, a graph visualization panel, and a property panel. A Node-Link Diagram and the Force-Directed Graph Drawing algorithm were used for the visualization of the graph. The results of the queries are visually presented through the graph visualization panel. Finally, this study used the data of worldwide import & export data of small arms to verify our model. The significance of this research is in the fact that, through the model which is able to conduct consecutive analyses on dynamic graph data, it helps overcome the limitations of previous models which can only perform discrete analysis on dynamic data. This research is expected to contribute to future studies such as online decision making and complex network analysis, that use dynamic graph models.

**Keywords** : Dynamic Data, Dynamic Graph, Multivariate Graph, Query Model

---

† 준회원: 경기과학고등학교  
 †† 정회원: 경기과학고등학교 교사  
 ††† 정회원: 아주대학교 정보컴퓨터공학과 교수(교신저자)  
 논문접수: 2014년 8월 14일, 심사완료: 2014년 9월 17일, 게재확정: 2014년 10월 8일

## 1. 서론

그래프에 대한 질의문 처리는 하이퍼텍스트, 월드 와이드 웹 검색, 지리학적 정보 시스템, 이중 정보 통합, RDF 데이터, 분자생물학 등 다양한 분야에서 연구, 활용되고 있다[1][2][3][4][5][6]. 이러한 그래프 질의문 처리는 그래프 데이터베이스에 활용되며, 관련 연구들은 그래프에서 대상간의 연관관계를 파악해 정확하고 빠르게 조건을 만족하는 대상들을 찾기 위한 그래프 질의 방법을 중요한 연구로 제시하고 있다[7][8].

최근에는 한 시점에 대한 그래프 정보만을 가지는 정적 그래프 뿐 아니라, 연속적으로 변화하는 동적 그래프 데이터를 처리하는 것이 중요한 연구 분야로 부각되고 있다. 즉, 데이터의 연속적인 분석을 통한 추세 예측에 대한 다양한 연구들이 이루어지고 있다[9][10][11]. 특히, 그래프 분야에서는 복잡계 네트워크를 동적 그래프를 사용하여 해석하려는 새로운 시도가 나타남에 따라 동적 그래프 처리의 중요성이 부각되고 있다[12][13][14].

연속적으로 변화하는 그래프 데이터에 대한 질의를 처리하기 위해서는 새로운 동적 그래프 질의 모델이 필요하다. 동적 그래프는 정점과 간선의 속성들이나 연결 상태가 시간에 따라 변화하며, 정점과 간선이 삭제되거나 새로운 정점 또는 간선이 생성된다. 따라서 동적 그래프 질의를 위해서는 이러한 그래프의 변화를 처리하는 기능을 가지고 있어야 한다. 그러나 기존 그래프 질의 연구들은 정적 그래프 질의 모델을 사용해 각 시점의 정보를 이산적으로 처리하는 것은 가능하지만, 시간 변화에 따른 그래프의 추이를 연속적으로 분석하는 데는 한계를 보여주고 있다.

이에 본 연구에서는 다양한 그래프 형태들 중 정점과 간선의 속성을 통해 유연한 질의 구성이 가능한 다변량 그래프(multivariate graph)를 질의 대상으로 하여, 다양한 시간대의 다변량 그래프를 연속적으로 분석할 수 있도록 동적 다변량 그래프 질의 모델 및 시스템을 설계 및 구현하였다. 또한, 구현된 시스템을 이용하여 세계의 소형 무기 거래량 데이터에 대한 질의를 수행함으로써 제시한 동적 다변량 그래프 질의 모델을 검증하였다.

## 2. 관련 연구

그래프 질의문 처리는 최근 빅데이터 및 실시간 웹 어플리케이션의 발달과 함께 빠른 속도로 성장중인 NoSQL 데이터베이스의 한 종류인 그래프 데이터베이스의 중요 요소이다[7][8][15]. 그래프 질의 처리에 대한 연구로 그래프 데이터베이스의 양적 분석을 위한 질의 언어, 객체 기반 질의 언어 OQL을 그래프 검색용으로 확장한 GOQL, 그래프 패턴을 정규화된 문자열로 나타내어 검색을 수행하는 GraphGrep 등 다양한 방식의 접근이 진행되어 왔다[16][1][17].

이러한 그래프 질의문 처리와 관련된 연구들에서 사용된 다양한 그래프 종류들 중, 본 연구에서 질의 대상으로 하는 그래프는 다변량 그래프이다. 다변량 그래프는 정점과 간선에 사용자 정의 속성들이 존재하는 그래프로, 그래프의 연결 구조 뿐 아니라 정점과 간선의 속성들을 이용하여 질의문을 구성할 수 있어 사용자가 원하는 다양한 질의를 구성할 수 있다는 점에서 유연성이 높다. 이러한 다변량 그래프 질의 모델들은 SQL과 유사한 형태의 쿼리문 작성, 시각화 프로그램과의 통합 등 다양한 형태로 개발되고 있으며, 그 활용 범위가 점점 늘어나고 있다. 다변량 그래프 질의에 관한 연구로는 대화식 그래프 시각화 프로그램 CGV, 그래프 탐색 및 조정 프로그램 Gephi 등이 있다[18][19].

최근 데이터의 연속적인 분석을 통한 추세 예측의 중요성이 부각됨에 따라 시간에 따라 변화하는 데이터 처리에 대한 다양한 연구들이 이루어지고 있다[9][10][11]. 특히, 그래프 분야에서는 복잡계 네트워크를 동적 그래프를 사용하여 해석하려는 새로운 시도가 나타남에 따라 동적 그래프 처리의 중요성이 상승하고 있다[12][13][14].

기존의 다변량 그래프 질의 모델들은 정적 그래프를 대상으로 하며, 연속적인 그래프 데이터를 분석하는 데는 한계가 있다. 시간의 추이에 따라 변화하는 동적 그래프를 기존의 다변량 그래프 질의 모델을 통하여 질의할 때, 동적 그래프의 각 시점의 그래프 상태를 하나의 정적 그래프처럼 해석하여 질의하는 것은 가능하다. 하지만 기존의 질의 모델은 수학의 그래프 이론에 기반하여 정

적 그래프 자체를 하나의 맥락(context)으로 해석하기 때문에, 정적 그래프 질의를 통해서도 각 시점들 간에 연관성을 나타내지 못하며 이산적인 분석만이 가능하다. 따라서 동적 그래프에서 정점과 간선에 대해 다양한 조건을 만족하도록 질의하기 위해서는 단순히 여러 개의 정적 그래프 질의를 통한 다중 질의가 아닌 확장된 질의 모델이 필요하다[20][21].

### 3. 연구 방법

본 연구는 동적 다변량 그래프를 위한 질의 모델을 설계하고, 질의 모델과 통합된 동적 그래프 시각화 프로그램을 개발해 동적 그래프 질의와 관련된 시나리오를 수행함으로써 질의 모델의 효과성을 검증하였다. 본 연구 과정은 <표 1>과 같다.

<표 1> 연구과정

[1] 동적 다변량 그래프 데이터 입력/관리 환경 설정
1) 동적 다변량 데이터 입력 환경 설정
2) 동적 다변량 데이터 관리 환경 설정
↓
[2] 동적 다변량 그래프 질의 모델 설계
1) 질의 모델 판별함수 설계
2) 시간 변화에 따른 판별함수 통합 설계
3) 질의 결과 시각화 설계
↓
[3] 동적 다변량 그래프 질의 모델 구현
1) 동적 다변량 그래프 질의 모델의 입력 및 데이터 관리 환경 구현
2) 동적 다변량 그래프 데이터 질의 패널 구현
3) 동적 다변량 그래프 시각화 패널 구현
↓
[4] 동적 다변량 그래프 질의 시스템 검증

#### 3.1 동적 다변량 그래프 데이터의 입력 및 관리 환경 설정

##### 3.1.1 동적 다변량 그래프 데이터의 입력 환경 설정

동적 다변량 그래프 입력 환경은 정점 및 간선의 생성, 삭제 및 속성값 변화를 포함한다. 이들을 통하여 특정 시점에서 그래프의 상태를 계산

하도록 설정하였다.

먼저, 동적 다변량 그래프의 특정 시점에서의 상태는 다음과 같이 정의하였다. 정점과 간선의 속성 개수가 각각  $n, m$ 이고  $A_V^i: 1 \leq i \leq n, A_E^i: 1 \leq i \leq m$ 이 각각 정점과 간선의  $i$ 번 속성의 정의역일 때, 동적 그래프 각 시점에 따른 상태  $G$  는 다음과 같이 정의할 수 있다.

$$G = (V, E)$$

$$\text{where } V \subseteq \{A_V^1 \times \dots \times A_V^n\}$$

$$\text{and } E \subseteq (V \times V) \times \{A_E^1 \times \dots \times A_E^m\}$$

그래프는 유향 그래프로 설정하였다. 만약, 정점  $A, B$  사이에 있는 무향 간선을 나타내고 싶은 경우  $A-B$  간선과  $B-A$  간선을 동일한 속성을 가지도록 설정하면 무향 간선이 존재하는 것과 동일한 효과를 얻을 수 있도록 하였다.

두 번째로, 각 정점들의 첫 번째 속성은 항상 문자열이라는 제약 조건을 추가하였고, 이 값을 서로 다른 정점을 구분하는 식별자로서 사용하였다. 또한, 간선은 간선의 양 끝에 있는 정점의 식별자 쌍을 식별자로서 사용하였다. 따라서 같은 정점 쌍 사이에 여러 개의 간선이 존재하는 경우는 표현할 수 없다.

세 번째로, 시각화 프로그램에서는 동적 그래프를 프로그램과 동일한 폴더에 존재하는 세 개의 텍스트 파일(frame.txt, node.txt, edge.txt)을 통해 입력 받도록 하였다. frame.txt 파일은 정점과 간선들이 어떤 속성들을 가지는지에 대한 정보를 저장, node.txt 파일은 정점의 생성/삭제 및 속성의 변화에 대한 정보를 저장, edge.txt 파일은 간선의 생성/삭제 및 속성의 변화에 대한 정보를 저장하도록 설계하였다.

##### 3.1.2 동적 다변량 그래프 데이터의 관리 환경 설정

본 연구에서는 입력된 동적 다변량 그래프 데이터를 다음과 같이 관리하도록 설정하였다.

먼저, 입력된 데이터에 최초로 이루어지는 처리는 정점과 간선 두 종류의 변화 단위로 분류하여 두 개의 배열에 시간 순서로 정렬하여 보관하는

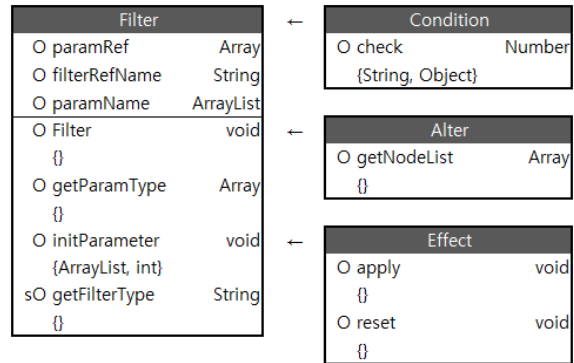
것이다. 정렬 후 같은 시점에 발생한 동일한 정점 또는 간선에 대한 변화단위들을 병합하여 하나의 변화 단위로 관리하였으며, 변화 정보들을 연관 배열 형태로 관리하도록 설계하였다. 효율적인 처리를 위하여 정점 또는 간선이 다양한 속성들 가운데 변화하지 않은 속성들은 병합하는 과정에서 연관 배열에 삽입하지 않아 실제로 변화가 일어난 속성들만을 고려하도록 설계하였다.

연관 배열 형태로 속성값의 변화를 저장한 후, 최초의 그래프 상태에서 시작하여 변화 단위를 하나씩 읽어가며 그래프를 생성하면서, 정점 또는 간선에 변화가 있을 때마다 변화 이전 시점으로 돌아가는 역방향의 변화를 두 번째 연관 배열을 만들어 저장하였다. 두 연관 배열을 이용하여 특정 시점의 그래프에서 시간이 증가하거나 감소했을 때 다음 그래프 상태를 변화 단위 개수에 따른 선형 시간에 계산 가능하다.

### 3.2 동적 다변량 그래프 질의 모델 설계

질의 과정에서 사용자가 총 세 단계의 절차를 통해 질의를 구성하도록 설계하였다. 먼저, 어떤 속성값을 이용해 판별함수를 구성할지를 결정한다. 두 번째로, 구성된 판별함수를 시간 변화에 따라 어떻게 통합할 것인지를 결정한다. 마지막으로, 질의 결과의 출력을 위해 질의를 통해 선택된 그래프 요소들에 어떤 시각적 효과를 적용할 것인지를 결정한다.

사용자가 질의를 구성할 때, 매개변수 설정 및 UI 통합 등과 관련된 부분을 Filter 클래스로 추상화하였다. 판별함수 설정은 Condition 클래스, 시간 변화에 따른 통합 방법 선택은 Alter 클래스, 시각적 효과 결정은 Effect 클래스에 각각 대응된다. <그림 1>은 동적 다변량 그래프 질의 모델 클래스 다이어그램이다.



<그림 1> 동적 다변량 그래프 질의 모델 클래스 다이어그램

먼저, 질의 모델에 사용되는 판별함수는 대상 정점과 특정 시점의 그래프 정보를 입력받아 대상으로 하는 정점의 판별함수의 값을 반환하도록 설계하였다.

두 번째로, 판별함수를 시간 변화에 따라 계산한 값을 어떻게 통합할 것인지 결정하도록 설계하였다. 시간 변화에 따른 함수값을 통합하는 방법의 예시로는 ‘판별함수가 항상 특정 값 이상’ 또는 ‘판별함수의 값이 단조 증가’ 등이 있다. 즉, 판별함수를 매개변수로 받아 여러 시점의 판별함수 값을 통합하여 특정한 정점이 조건을 만족하는지를 반환하도록 설계하였다.

마지막으로, 질의 모델의 결과 출력은 질의 결과로 선택된 정점과 간선들에 사용자의 요구사항에 맞게 시각적 효과를 적용하도록 설계하였다. 질의를 통해 선택된 대상들은 시각화 과정에서 크기나 색이 변화하는 등의 효과를 받는다.

### 3.3 동적 다변량 그래프 질의 시스템 구현

설계된 질의 모델을 바탕으로, 실제로 동적 다변량 그래프 질의를 수행할 수 있는 프로그램을 제작하였다. 구현 시에는 FlashDevelop 4 IDE에서 Apache Flex SDK 4.10.0를 사용하였다.

먼저, 동적 다변량 그래프 데이터 입력 및 관리 환경을 구현하였다. 다루고자 하는 동적 다변량 그래프의 범위를 정의하고, 그래프의 입력방식, 입력 과 일의 형식, 기억장소 저장 방식을 설정하였다. 두 번째로, 동적 다변량 그래프 질의 모델은 질의 패널, 질의 결과에 따른 그래프 시각화

패널, 속성 패널로 구성하여 구현하였다. 질의 패널은 사용자가 질의 언어를 사용하지 않고, 그래픽 컴포넌트를 이용하여 질의를 구성하도록 하였다. 질의 결과를 보여주는 그래프 시각화 패널은 그래프를 노드-링크 다이어그램을 이용해 표현하였고, 레이아웃 알고리즘으로 Force-Directed Graph Drawing 알고리즘을 사용하였다. 속성 패널에서는 그래프 시각화 패널에서 선택한 그래프 요소의 속성을 확인할 수 있도록 하였다.

### 3.4 동적 다변량 그래프 질의 시스템 검증

본 연구에서 구현한 동적 다변량 그래프 질의 시스템을 검증하고자, 구글의 세계 소형 무기 거래량 시각화 사이트 Interactive Globe에 사용된 데이터를 이용하였다. <표 2>는 시스템 검증을 위해 사용된 데이터 및 시나리오를 보여준다.

<표 2> 동적 다변량 그래프 질의 시스템 검증 데이터 및 시나리오

질의주제	소형 무기 거래량
데이터 (무기 거래량 국가명)	Malawi Zimbabwe Korea, Republic of Korea, Democratic People's Republic of Italy Switzerland United States China Unspecified Country
기간	1992~2010
시나리오	2000년대 들어 중국보다 한 번이라도 무기 수입을 많이 한 나라를 하이라이팅하고, 한국보다 항상 무기 수출량이 적었던 나라를 보이지 않게 해라.
총 수입량	i_all
총 수출량	e_all

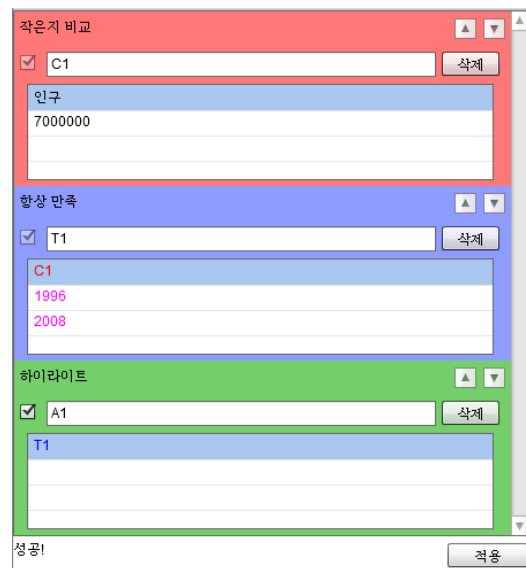
## 4. 연구 결과

### 4.1 동적 다변량 그래프 질의 시스템 구현

본 연구는 설계한 동적 다변량 그래프 질의 모델을 이용하여 다변량 그래프를 질의하고, 결과값을 출력하는 과정을 구현하였다.

#### 4.1.1 동적 다변량 그래프 데이터 질의 패널 구현

본 연구에서는 질의 모델을 사용할 때 사용자가 질의어나 스크립트를 직접 작성하는 것이 아니라, 블록 형태의 그래픽 컴포넌트들을 조합하여 질의를 구성할 수 있도록 <그림 2> 와 같이 구현하였다.



<그림 2> 동적 다변량 그래프 질의 모델의 질의 패널

컴포넌트들은 <그림 2>와 같이 세로로 나열되는 형태로 질의 패널에 추가된다. 컴포넌트의 좌측 상단에는 어떤 처리를 하는 컴포넌트인지를 나타내는 문자열이 출력된다. 우측 상단에는 화살표 버튼이 존재해 컴포넌트들의 순서를 바꾸는 기능을 제공한다. 그 하단 부분에는 체크 박스와 컴포넌트의 참조명을 입력할 수 있는 텍스트 박스가 존재한다. 체크 박스는 시각적 효과 적용과 관련된 컴포넌트인 경우 활성화되며, 효과 적용 여부를 빠르게 토글링하여 확인할 수 있다. 참조명은 다른 컴포넌트에서 매개변수로 사용할 때 참조할 수 있도록 해 주는 문자열이다.

컴포넌트들은 배경색을 통해 질의 시스템의 어떤 단계에 속하는지를 구분할 수 있도록 하였다. <그림 2>와 같이 C1과 같이 판별함수 설정과 관련된 컴포넌트는 붉은색 배경으로, T1과 같이 함숫값 통합 방법 선택과 관련된 컴포넌트는 푸른

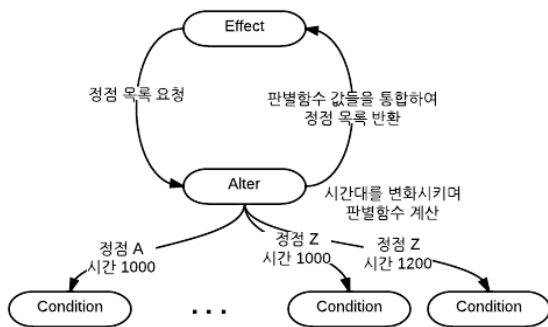
색 배경으로, A1과 같이 시각적 효과 적용과 관련된 컴포넌트는 초록색 배경으로 나타내었다.

하단에 있는 매개변수 목록을 클릭하여 값을 변경할 수 있으며, 글자 색을 통해 어떤 종류의 매개변수를 입력받는지 알 수 있도록 하였다. 붉은 색은 판별함수, 푸른색은 합숫값 통합 방법, 초록색은 시각적 효과 적용, 검은색은 값(value), 보라색은 시간대를 나타낸다. ‘값’은 상수이거나 특정한 정점의 속성일 수 있다. 문법은 <표 3>과 같다.

<표 3> ‘값’ 매개변수에 사용할 수 있는 문법

종류	문법	의미
문자열 상수	@str	문자열 상수 str
실수 상수	15321	실수 상수 15321
현재 정점의 속성	A	현재 정점의 A 속성값
특정 정점의 속성	B.C	B 정점의 C 속성값

질의 패널의 수행과정은 <그림 3>과 같다.



<그림 3> 질의 패널의 수행 과정

먼저, Effect 클래스에서 Alter 클래스에 정점 목록을 요청하면, Alter 클래스는 시간과 대상 정점을 변화시키며 Condition 클래스를 호출한다. 그 후, Alter 클래스에서는 반환된 판별함수 값을 통합하여 조건을 만족하는 정점 목록을 Effect 클래스에 반환하고, 다시 Effect 클래스에서는 이들 정점에 시각적 효과를 적용시키도록 하였다.

#### 4.1.2 동적 다변량 그래프 시각화 패널 구현

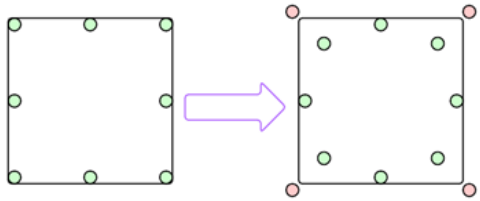
본 연구에서는 그래프의 정보를 시각화 할 때 노드-링크 다이어그램을 이용하였으며, 레이아웃 알고리즘으로는 Force-Directed Graph Drawing 알고리즘을 이용하였다. 레이아웃 알고리즘에 의해 배치된 그래프의 형태를 사용자가 커스터마이징할 수 있도록 정점들을 드래그하여 움직일 수 있는 기능을 구현하였다.

또한, 현재 관찰중인 시간이 변화함에 따라 정점이 새롭게 생성되는 경우 화면상의 임의의 위치에 생성되도록 처리하였다. 이 때, 질의의 결과로 선택된 정점들의 가중치를 변경할 수 있도록 하여 가중치가 더 클수록 화면에 더 크게 그려지며 더 강한 힘을 작용하도록 처리하였다.

그러나 Force-Directed Graph Drawing 알고리즘은 본 연구에서 그대로 적용하였을 때 일부 문제점이 발생하는데, 이 문제점을 다음과 같이 해결하였다.

첫 번째 문제는 관찰중인 시간대의 이동에 따라 정점이 생성될 때 기존의 정점에 가까운 곳에 생성되거나, 사용자가 정점들을 드래그하여 다른 정점에 가까이 가져가는 등의 과정에서 두 정점 사이의 거리가 지나치게 가까워질 수 있다. 레이아웃 알고리즘에서 임의의 두 정점이 서로에게 미치는 힘을 계산하는 과정 중 두 정점 사이의 거리가 분모에 대입되는데, 이 때문에 두 정점이 지나치게 가까운 경우 정점끼리 많은 힘을 받아 비정상적으로 튕겨 나가는 경우가 생길 수 있다. 이를 방지하기 위하여 Force-Directed Graph Drawing에서 두 정점 사이의 힘을 계산할 때 두 정점 사이의 거리가 특정한 상수 이하로 작아지는 경우 두 정점 사이의 거리를 사용하는 대신 상수를 사용하여 처리하였다.

두 번째 문제는 가상의 이차원 상자 안에 갇혀 있는 형태로 시뮬레이션하기 때문에, 정점들 사이의 척력 때문에 대부분의 정점이 상자의 벽 근처로 이동하는 현상이 생길 수 있다. 이러한 현상을 막기 위하여 상자의 네 꼭짓점에 보이지 않고 척력만 작용하도록 하는 가상의 정점을 두어 <그림 4>와 같이 정점들이 지나치게 구석으로 몰리지 않도록 처리하였다.



<그림 4> 정점이 상자의 벽 근처로 이동하지 못하도록 가상의 정점 배치

간선을 그리는 과정에서 정점의 중심과 중심을 이어서 그리는 경우 방향 그래프에서 두 정점 사이에 서로에게 향하는 간선이 있는 경우 간선이 겹쳐서 보이며, 두 간선의 구분이 잘 되지 않는다. 이러한 문제는 각 간선을 진행 방향을 기준으로 약간 오른쪽에 그려서 반대방향의 간선과 겹치지 않도록 하고, 출발 정점에서 도착 정점으로 향할수록 간선의 색깔이 붉은색에서 푸른색으로 변하도록 그래데이션을 적용하여 간선의 방향을 파악할 수 있도록 처리하였다.

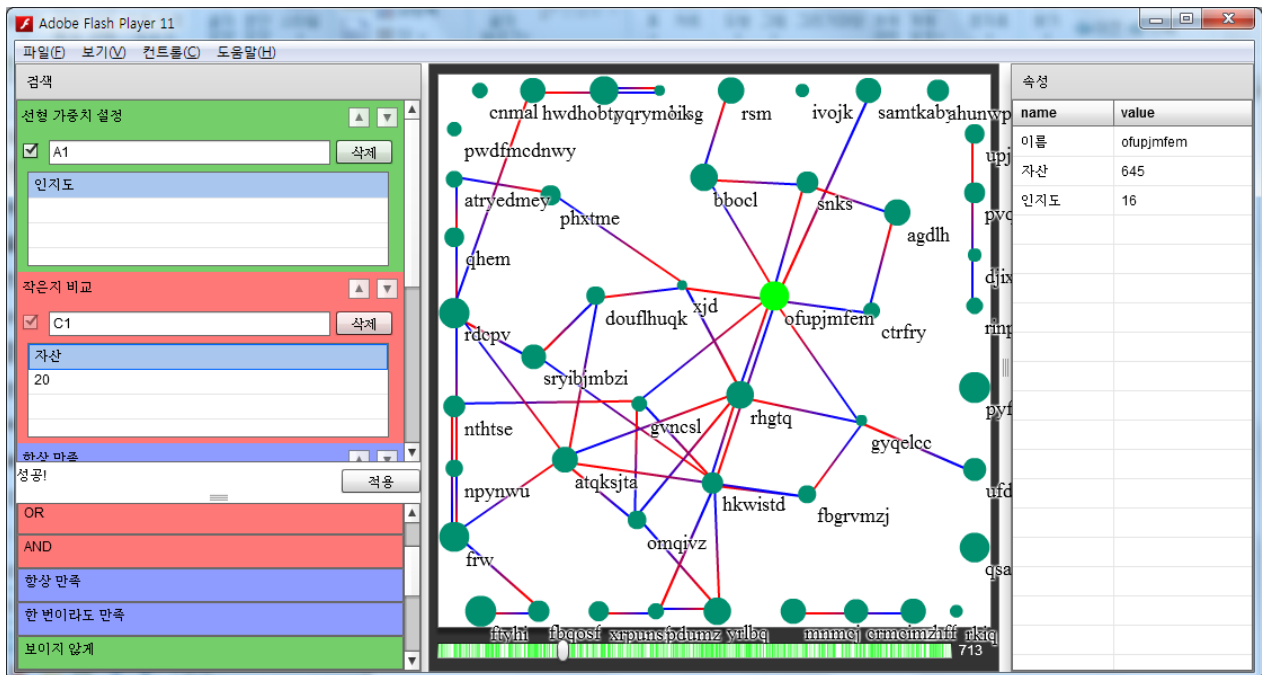
### 4.1.3 동적 다변량 그래프 질의 시스템 화면 구성

동적 다변량 그래프 질의 시스템은 <그림 5>

와 같이 질의 패널, 시각화 패널, 속성 패널의 세 부분으로 화면을 구성하였다.

먼저, 좌측의 질의 패널에서는 그래프에 적용할 질의를 그래픽 컴포넌트를 이용해 구성할 수 있다. 질의 패널 하단에는 사용할 수 있는 컴포넌트들의 목록이 있고, 상단에는 현재 활성화된 컴포넌트 리스트가 있다. 하단 영역에서 추가하고 싶은 컴포넌트를 더블클릭하여 적용중인 리스트에 추가할 수 있으며, 상단 영역에서 현재 활성화된 컴포넌트 리스트를 볼 수 있으며, 매개변수들을 조정할 수 있다.

두 번째로, 중앙의 시각화 패널에서는 현재 그래프의 상태를 파악하고, 관찰중인 시간대를 변경할 수 있다. 시각화 패널의 상단에는 그래프가 표시되어 그래프의 정점들을 드래그하여 이동시켜 레이아웃을 변경할 수 있다. 시각화 패널의 하단에는 슬라이더가 있어 관찰중인 시간대를 변화시킬 수 있다. 슬라이더는 frame.txt에서 입력되는 시작 시간과 끝 시간 사이에서 같은 방법으로 입력되는 시간 간격을 단위로 하여 움직인다. 슬라이더의 움직임에 따라 현재 화면에 표시되고 있는 동적 그래프의 시점이 변화한다. 슬라이더의 배경에는 변화가 일어난 시점에 선이 표시되어



<그림 5> 동적 다변량 그래프 질의 시스템의 화면 구성

어떤 범위에서 변화가 많이 일어났는지를 파악할 수 있으며, 슬라이더 우측에서 현재 시점 정보를 볼 수 있다.

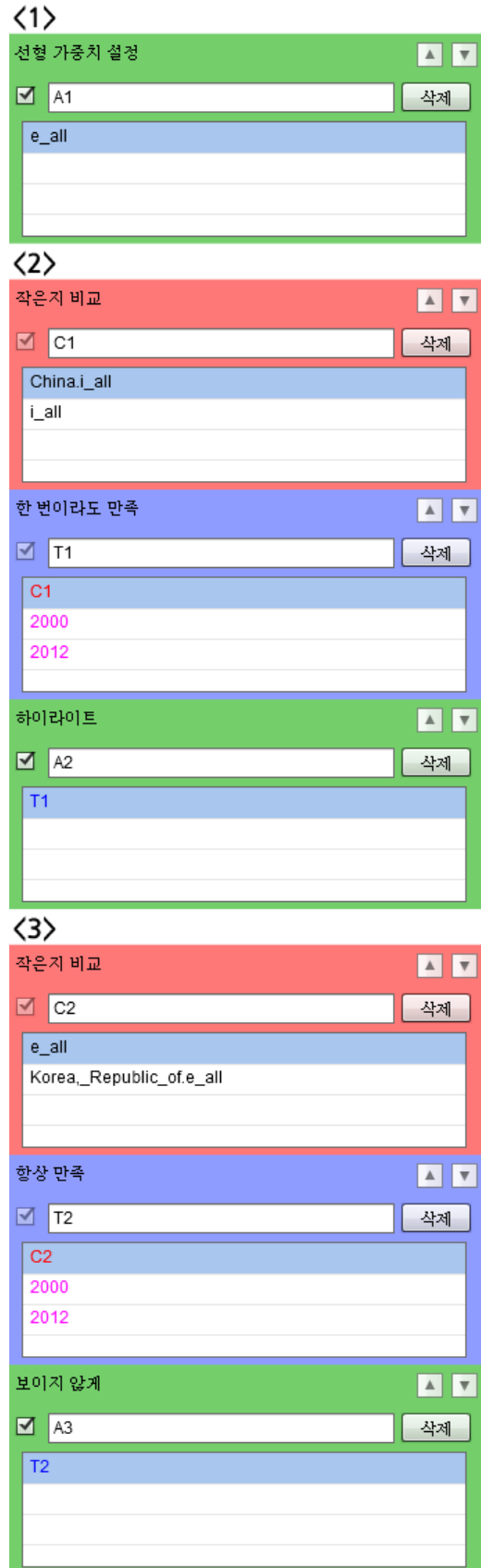
마지막으로, 우측의 속성 패널에서는 시각화 패널에서 선택한 정점 또는 간선들의 속성값을 확인할 수 있다. 속성들의 순서는 frame.txt에 입력된 것과 동일한 순서로 나타난다.

#### 4.2 동적 다변량 그래프 질의 시스템 검증

본 연구에서 구현한 동적 다변량 그래프 질의 시스템을 검증하고자, 구글의 세계 소형 무기 거래량 시각화 사이트에 사용된 자료(Interactive Globe)를 이용하였다.

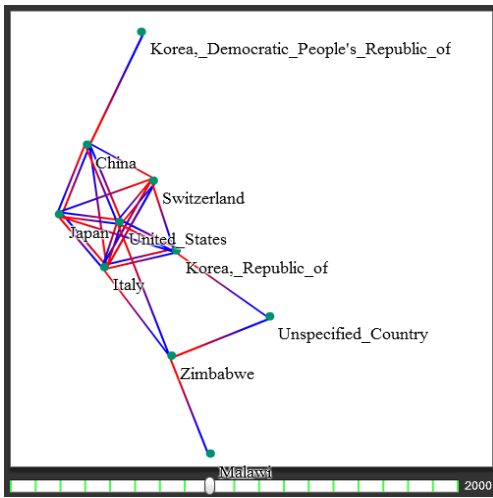
질의 패널 부분은 <그림 6>, 질의 시스템 실행 결과 나타나는 시각화 패널은 <그림 7>~<그림 10>이다.

먼저, 어떠한 효과도 적용하지 않은 상태에서의 소형 무기 거래량 그래프는 <그림 7>과 같다. 질의를 적용하지 않은 상태에서는 간선의 연결 상태를 통해 어떤 나라끼리 거래가 이루어지는 정보만을 알 수 있을 뿐, 거래량이나 거래 규모 등에 대한 정보를 파악하기 위해서는 정점과 간선을 클릭하여 속성 패널에서 일일이 확인해야 한다. 질의 모델은 이러한 원시 데이터에 질의문의 적용을 통해 사용자가 원하는 특정한 정보들을 강조해서 나타내어 시각적으로 확인할 수 있게 해 준다.



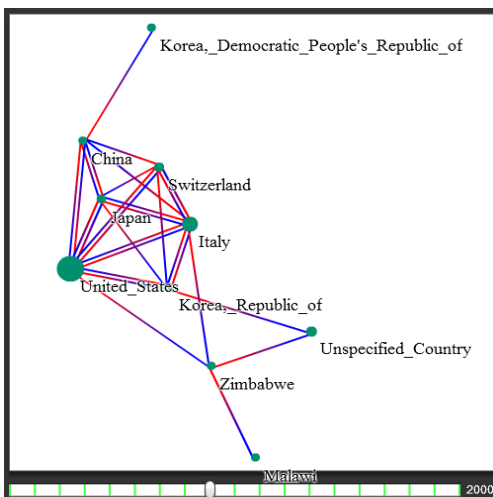
<그림 6> 적용된 질의 목록





<그림 7> 아무런 효과도 적용하지 않은 상태에서의 소형 무기 거래량 그래프

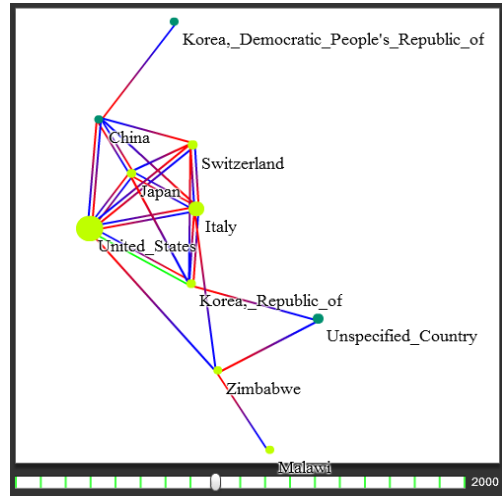
<그림 8>은 총 수출량(e\_all)을 기준으로 정점들에 가중치를 적용한 그래프 결과로, <그림 6>의 1번 컴포넌트가 적용되었다. <그림 8>에서와 같이 노드 크기를 통해 수출량을 직관적으로 파악할 수 있다.



<그림 8> 총 수출량(e\_all)을 기준으로 선형 가중치를 적용한 상태

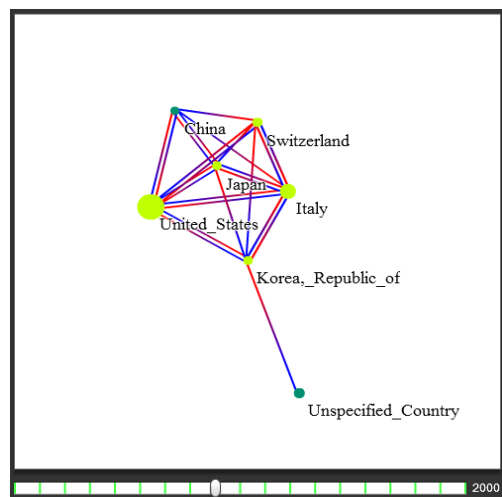
<그림 9>는 2000년대 이후로 중국보다 한 번이라도 무기 수입(i\_all)을 많이 한 나라를 강조해서 표시한 그래프 실행결과이다. <그림 8>의 질의 컴포넌트에 총 네 개의 컴포넌트가 추가적으로 적용된 상태이다. 스위스, 일본, 이탈리아, 미국, 대한민국, 짐바브웨, 말라위가 강조되어 표시

된 것을 알 수 있다.



<그림 9> <그림 8>에서 2000년대 이후 중국보다 한 번이라도 무기 수입(i\_all)을 많이 한 나라를 강조하여 표현한 모습

마지막으로, <그림 10>는 2000년대 이후로 한국보다 항상 무기 수출량(e\_all)이 적었던 나라를 보이지 않게 한 그래프 실행 결과이다. 적용된 질의 컴포넌트는 <그림 9>에 3개의 컴포넌트가 추가되어 총 7개의 질의 컴포넌트가 적용된 상태로, 말라위, 짐바브웨, 북한이 보이지 않게 된 것을 알 수 있다.



<그림 10> <그림 9>의 상태에서 2000년대 이후 한국보다 항상 무기 수출량(e\_all)이 적었던 나라를 보이지 않게 한 상태

## 5. 결론

본 연구에서는 기존의 정적 다변량 그래프 질의 모델의 연속적으로 변화하는 데이터 질의에 대한 한계점을 극복하기 위해 동적 다변량 그래프 질의 모델을 설계, 구현하였다. 또한, 제시한 질의 모델을 바탕으로 동적 다변량 그래프 질의 시스템을 구현하여 질의를 수행하고, 질의의 결과를 시각적으로 확인할 수 있도록 구현하였다. 구현한 시스템을 통해 세계의 소형 무기 거래량 데이터를 이용하여 동적 다변량 그래프 질의를 수행함으로써 제시한 질의 모델의 가능성을 검증하였다.

본 연구에서 설계한 동적 그래프 질의 모델은 기존의 그래프 질의 모델이 동적 그래프를 이산적으로만 분석할 수 있던 한계점을 개선하고, 시간에 따라 변화하는 그래프 데이터를 연속적으로 분석할 수 있도록 확장한 점에 의의를 둘 수 있다. 본 연구는 그래프 데이터에 대한 추세 분석이나 최근 이루어지고 있는 복잡계 네트워크를 동적 그래프로 해석하는 새로운 시도 등 변화하는 그래프에 대한 연구들에 기여할 수 있을 것으로 기대된다.

## 참고 문헌

- [1] Sheng, L., Ozsoyoglu, Z.M., Ozsoyoglu, G. (1999). A graph query language and its query processing, *Proceedings, IEEE ICDE*, 572-581.
- [2] Arocena, G., Mendelzon, A. (1998). WebOQL: Restructuring Documents, Databases and Webs, *Proceedings, IEEE ICDE*, 24-33.
- [3] Gyssens, M., Van Den Bussche, J., Paradaens, J. (1994). A Graph-Oriented Object Database Model, *Knowledge and Data Engineering* 6(4), 572-586.
- [4] Mendelzon, A.O., Mihaila, G.A., Milo, T. (1996). Querying the WorldWide Web, *Proceedings, IEEE ICPADS*, 80-91.
- [5] 김진하, 송인철, 김명호 (2008). RDF 데이터에 대한 효율적인 키워드 검색 기법, **정보과학회논문지 : 데이터베이스**, 35(6), 495-504.
- [6] Sealfon, R., Hibbs, M., Huttenhower, C., Myers, C., Troyanskaya, O. (2006). GOLEM: an interactive graph-based gene-ontology navigation and analysis tool, *BMC Bioinformatics* 7:443.
- [7] Holzschuher, F., Peinl, R. (2013). Performance of graph query languages: comparison of cypher, gremlin and native access in Neo4j, *Proceedings, the Joint EDBT/ICDT 2013 Workshops*, 295-204.
- [8] Vicknair, C., Macias, M., Zhao, Z., Nan, X., Chen, Y., Wilkins, D. (2010). A comparison of a graph database and a relational database: a data provenance perspective, *Proceedings, ACM the 48th Annual Southeast Regional Conference*, Article No.42.
- [9] Gupta, R., Ramamritham, K. (2007). Optimized Query Planning of Continuous Aggregation Queries in Dynamic Data Dissemination Networks, *Proceedings, WWW '07*, 321-330.
- [10] Hochheiser, H., Shneiderman, B. (2004). Dynamic query tools for time series data sets: Timebox widgets for interactive exploration, *Information Visualization* 3(1), 1-18.
- [11] Huang, Y., Yu, P. (1999). Adaptive Query Processing for Time-Series Data, *Proceedings, ACM SIGKDD KDD '99*, 282-286.
- [12] Lu, J., Chen, G. (2005). A Time-Varying Complex Dynamical Network Model And Its Controlled Synchronization Criteria, *Automatic Control* 50(6), 841-846.
- [13] Wang, X., Chen, G. (2002). Synchronization in Scale-Free Dynamical Networks: Robustness and Fragility, *Circuits and Systems I: Fundamental Theory and Applications* 49(1), 54-62.
- [14] Braha, D., Yam, Y. B. (2006). From Centrality to Temporary Fame: Dynamic Centrality in Complex Networks,

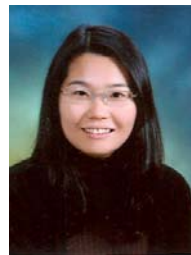
*Complexity* 12(2), 59-63.

- [15] Strauch, C., Kriha, W. *NoSQL Databases*. Lectures notes of Ultra-Large Scale Sites at Stuttgart Media University.
- [16] 박성찬, 이상구 (2011). 그래프 데이터베이스의 양적 분석을 위한 질의 언어, **한국정보과학회 2011한국컴퓨터종합학술대회 논문집** 38(1), 77-80.
- [17] Giugno, R., Shasha, D. (2002). GraphGrep: A Fast and Universal Method for Querying Graphs, *Proceedings, ICPR, 2*, 112-115.
- [18] Tominski, C., Abello, J., Schumann, H. (2009). CGV - An interactive graph visualization system, *Computers & Graphics* 33(6), 660-678.
- [19] Bastian, M., Heymann, S., Jacomy, M. (2009). Gephi : An Open Source Software for Exploring and Manipulating Networks, *Proceedings, AAAI ICWSM*, 361-362.
- [20] Angles, R., Gutierrez, C. (2008). Survey of Graph Database Models, *ACM Computing Surveys* 40(1), Article No.1.
- [21] Trudeau, R. J. (1993). *Introduction to graph theory*. Courier Dover Publications.



### 김 태 양

2015 경기과학고등학교 졸업예정  
 관심분야: 알고리즘  
 E-Mail: k5888200@gmail.com



### 정 혜 진

1993 경기대학교  
 진산학과(이학사)  
 1999 아주대학교  
 컴퓨터공학과(공학석사)  
 2008 아주대학교 정보통신전문 대학원  
 박사수료  
 2004~현재 과학영재학교 경기과학고등학교 교사  
 관심분야 : 유전 알고리즘, 생물정보학,  
 문제해결전략설계  
 E-Mail : seagujinny@gmail.com



### 배 예 찬

2015 경기과학고등학교 졸업예정  
 관심분야: 알고리즘, 그래프이론  
 E-Mail: qwazpia@gmail.com



### 김 동 운

1974 서울대학교  
 수학과(이학사)  
 1976 KAIST  
 진산학과(이학석사)  
 1985 MIT 응용수학(Ph.D)  
 1991~현재 아주대학교 정보컴퓨터공학과 교수  
 2013~현재 서울어코드 부의장  
 2014~현재 World Computer Congress 추진위원장  
 관심분야 : 컴퓨터비전, 알고리즘, 컴퓨터교육정책  
 E-Mail : dykim@ajou.ac.kr



### 함 도 영

2015 경기과학고등학교 졸업예정  
 관심분야: 알고리즘  
 E-Mail: wj2dy@naver.com