

## 국가 연구개발 정보체계 분석을 위한 정보생성행렬 분석

김종우\* · 주영진\*\* · 이성용\*\*\* · 정현수\*\*\*\*

### Information Creation Matrix Analysis for Analyzing National R&D Information System

Jong Woo Kim\* · Young Jin Joo\*\* · Seong Young Lee\*\*\* ·  
Hyun Soo Jung\*\*\*\*

#### Abstract

In this paper, we propose matrix analysis methods for analyzing national R&D information creation. In order to analyze R&D information creation at national level, it is necessary to analyze whether information is created systematically for each technical category and for each information type. In this paper, 'uniformity' and 'concentration' criterions are proposed to check national R&D information creation and we provide formulas to measure the criterions. The criterions are applied to domestic information creation in information and communication domain to show the utilization of the proposed method.

- 
- \* 충남대학교 통계학과 조교수
  - \*\* 배재대학교 전자상거래학과 전임강사
  - \*\*\* 한국전자통신연구원 연구원
  - \*\*\*\* 한국전자통신연구원 책임연구원

## 1. 서론

기업에 있어서 정보 및 지식 관리 체계가 중요한 것과 마찬가지로 국가 수준에서도 정보 및 지식 관리 체계의 중요성이 강조되고 있다. 이러한 국가 정보 및 지식 체계를 발전시키기 위한 노력들은 선진국을 중심으로 다양한 접근방법을 통해서 활발하게 진행되고 있다. 예를 들어, 기존의 도서관시스템을 발전시켜 국가적인 수준의 지식 경영 체계로 확장하려는 노력으로는 캐나다의 CISTI (Canada Institute for Science and Technical Information), 미국 CNI (Coalition for Networked Information) 등을 손꼽을 수 있다[4,6,9]. 다른 접근법으로는 공식적인 지식 공유 네트워크의 구축과 활성화를 통해서 국가 지식 체계를 강화하려는 시도인데, 협회 중심의 지식 공유 네트워크인 일본의 OITDA (Optoelectronic Industry and Technology Development Association), 국가주도의 공식 지식 네트워크 구축 프로그램인 캐나다의 NCE(Networks of Centres of Excellence) 프로그램 등이 그 대표적인 예이다[5,11]. 이외에도 국가적인 태스크포스 구성을 통해서 국가정보체계의 활성화를 도모하려는 시도로는 미국의 KM-IP (National Knowledge and Intellectual Property Management) Task Force 등을 들 수 있다 [8].

본 논문에서는 국가 수준에서 정보 생성 현황을 파악하기 위한 기법을 제시하고자 한다. 기존의 정보생성 소재를 분석하기 위한 대표적인 기법 중의 하나로 정보공학 (Information Engineering)의 CRUD(Create, Read, Update, Delete) 행렬 기법이 있다. CRUD 행렬 기법을 활용하면, 정보의 생성, 읽기, 갱신, 삭제의 소재를 알 수 있으며 균

집화 알고리즘을 활용하여 관련성 높은 데이터들의 군집을 생성하는 것이 가능하다 [7,10]. 하지만 이러한 CRUD 행렬은 기업 내의 정보 소재의 파악과 데이터베이스 분산을 위한 목적으로 활용되는 것이고, 국가 차원의 정보 생성 소재를 분석하기에 충분한 정보를 제공하지 못한다. 본 논문에서는 국가 수준에서 정보 생성 현황을 체계적으로 분석하기 위한 개념으로 균질도 (Uniformity)와 집중도(Concentration) 개념을 제시하고, 이를 정보 생성 기관 현황을 조사한 정보생성행렬(ICM, Information Creation Matrix)로부터 측정하는 방법을 제시한다. 제시된 개념 및 방법은 정보통신 분야의 정보 생성 및 유통을 주요 목적으로 정보통신부와 정보통신연구진흥원이 주도하고 있는 정보통신 연구기반조성 사업에서의 정보 생성 현황 분석에 적용하여 그 유용성을 검토한다. 본 논문의 구성은 다음과 같다. 2장에서는 본 연구를 수행하게 한 연구 배경을 설명한다. 3장에서는 정보생성행렬의 정의를 제시하고, 4장에서는 정보균질도, 정보집중도 개념을 제시하고, 예제를 통해서 유용성을 설명한다. 5장에서는 결론을 제시한다.

## 2. 연구 배경

국가 차원의 특정 분야의 연구개발 정보 생성을 분석하기 위해서는 다양한 수준의 기술 분류별, 정보 유형별 정보 생성 현황에 대한 조사가 필요하다. 예를 들어 정보통신 분야의 정보 생성 현황을 검토하기 위해서는 정보 통신 기술 범주(유선통신, 무선기술, 방송/TV, 인터넷 등)의 세부 기술 분류별로 정보 생성 기관을 파악하는 것이 필요하다. 예를 들어, <표 1>은 정보통신분야의 기술 분류 체계의 한 예이다.

〈표 1〉 정보통신 기술분류체계

기술범주	기술세부분류
유선통신	서비스, 교환/전송기술, 가입자망, 구매망, 광통신기술, 단말기술
무선기술	서비스, 이동교환/전송, 위성교환/전송, 단말기술
방송/TV	방송서비스, 방송전송기술, 방송신호처리, 스튜디오기기, 방송수신/단말
인터넷	인터넷기반, 인터넷지원, 인터넷활용
컴퓨터/주변장치	하드웨어, 컴퓨터통신기술, 주변장치
소프트웨어	패키지소프트웨어, 컴퓨팅서비스, 지식정보처리 소프트웨어
멀티미디어콘텐츠	디지털콘텐츠제작, 교육용콘텐츠, 오락게임용콘텐츠
반도체 및 부품	반도체, 부품
정보가전	홈네트워킹, 홈서버/정보가전단말, 정보가전기반소프트웨어
정보보호	공통기반기술, 시스템/네트워크보호

주) 정보통신 중장기기술개발 계획(정보통신부) 참조

한 세부 기술 분류에 대해서도 국가적인 차  
원에서는 다양한 유형의 정보가 정보 수요  
자의 목적에 따라 필요하게 된다. 예를 들  
어 정보통신 분야의 경우는 기술정보, 시장

정보, 정책정보, 기타정보 등으로 구분할 수  
있으며, 이 각각의 내용은 <표 2>와 같이  
정리된다[2].

〈표 2〉 정보유형

정보 유형	주요 내용
기술정보	기술개요, 기술동향, 기술전망, 기술의 상용화
시장정보	시장개요, 시장동향, 시장전망, 기타(과급효과 분석, 요소기술 수급 동향 등)
정책정보	중장기 기술 개발, 산업육성 및 인력양성 정책, 국제 기구 현안 분석 및 대응 전략, 표준화 동향 및 대응 전략, 정책 프로그램 효과 분석
기타정보	주요 전시회 및 포럼 정보, 기술-시장간의 match-making, 인적 네트워킹을 위한 전문인력 DB

전체 국가 정보 체계 관점에서는 <표 1>과 <표 2>에서와 같은 다양한 기술 분야별로 다양한 유형의 정보가 어떤 정보 생성 단위(주체, 기관 또는 과제)에 의해서 어떻게 생성되고 있는지를 파악하는 것이 필요하다. 실제로 본 연구의 동기가 된 정보통신 분야 국가 정보 생성에 대한 주요한 질문들은 다음과 같다.

- 특정 세부 기술 분류에 해당하는 정보를 생성하는 과제 및 기관은 어디인가?
- 정보통신 분야 내의 세부 기술 분류의 정보들이 빠짐없이 생성되고 있는가?
- 특정 범주 내에 포함된 세부 기술 분류에 해당하는 정보들이 빠짐없이 균일하게 생성되고 있는가?
- 특정 세부 기술 분류에 대하여 기술 정보, 시장 정보, 정책 정보, 기타 정보가 모두 생성되고 있는가?
- 정보 생성 과제 및 기관들은 자신의 정보 생성 역할과 범위에 대하여 명확한 이해를 가지고 있는가?
- 어떤 정보 생성 과제 또는 기관이 특정 기술 범주 또는 특정 정보 유형에 특화된 정보 생성 단위인가?

위와 같은 질문을 답하기 위해서는 기존의 정보공학의 CRUD 행렬[7,10]이나 정보연관성분석방법(Association Rules Analysis)[1,3]만으로는 부족하고 보다 체계적인 정보 생성 현황 파악과 분석 방법이 필요하였다.

### 3. 정보생성행렬

#### 3.1 정보생성행렬 정의

국가적 수준에서 정보 생성 현황을 파악하기 위해서는 정보 생성 단위(과제 또는 기관)별 상세한 조사가 필요하다. 상세한 조사 이전에 정보 생성 단위별로 어떤 기술 분류의 정보를 생성하는지, 어떤 정보 유형의 정보를 생성하는지에 대한 개관적인 파악이 필요하다. 이러한 정보 생성 파악 조사 도구로 정보생성행렬(Information Creation Matrix, ICM)을 다음과 같이 정의한다.

#### Definition 1. 정보생성행렬(ICM)

$I$  = 정보생성단위 집합,  $J$  = 기술세부분류 집합,  $K$  = 정보유형집합이라 하면,

정보생성행렬  $A = [a_{ijk}]$ 는  $N(I) \times N(J) \times N(K)$  행렬로, 여기서  $i \in I, j \in J, k \in K, N(I) = I$ 의 원소수,  $N(J) = J$ 의 원소수,  $N(K) = K$ 의 원소수이고, 각 원소값  $a_{ijk}$ 는 다음과 같이 정의된다.

$$a_{ijk} = \begin{cases} 1, & \text{정보생성단위 } i \text{가 기술분류 } j \text{에 대한 정보유} \\ 0, & \text{생성하지 않는 경우} \end{cases}$$

특정 기술세부분류는 하나의 기술범주에 속하게 되는데, 이를 다음과 같이 정의한다.

#### Definition 2. 세부기술분류의 기술범주 포함관계

임의의 세부기술분류  $j \in J$ 에 대하여, 세부기술분류  $j$ 가 속하는 기술범주  $c \in C$ 가 정확히 하나 존재한다. 이를  $j \rightarrow c$ 로 표기한다.

#### 3.2 정보생성 행렬 예제

본 논문에서는 국내 정보통신분야의 정보 생성 현황 분석 자료를 활용하여 정보생성행렬 분석법을 설명하도록 한다. 정보통신 분야의 정보체계를 검토하는 일환으로 2000년, 2001년 정보통신 분야의 19개 연구기관 조성 사업에서의 정보 생성 및 배분에 대한

현황이 조사되었다[2]. 이 중 정보 생성의 역할을 담당하는 16개 과제에 대하여 정보 생성행렬이 작성되었다. 조사양식에서는 [부록 1]에서와 같이 기관별로 세로축에는 세부기술분류 축을 가로축에는 정보유형을 표시하여 작성되었다.

## 4. 정보생성의 균질도와 집중도

### 4.1 균질도와 집중도 정의

정보 생성의 균질도(Uniformity)는 기술범주 균질도(technology Category Uniformity, CU)와 정보유형 균질도(information Type Uniformity, TU)로 구분된다. 기술범주 균질도는 특정 기술범주에 속한 하위 세부 기술 분류들이 빠짐없이 생성되는지에 대한 정도를 표현하며, 정보유형 균질도는 특정 정보 유형에 대하여 다양한 기술 세부 분류가 빠짐없이 균일하게 생성하는 정도에 대한 척도로 다음과 같이 정의된다.

#### Definition 3 (정보균질도1)

정보생성 단위  $i$ '의 기술범주  $c$ '에 대한 균질도

$$CU(i, c) = \frac{N(a_{ijk}=1, i=i, j \rightarrow c)}{N(j|c) \times N(K)},$$

정보생성 단위  $i$ '의 정보유형  $k$ '에 대한 균질도

$$TU(i, k) = \frac{N(a_{ijk}=1, i=i, k=k)}{N(J)},$$

단,  $N(a_{ijk}=1, i=i, j \rightarrow c)$ 는 A의 원소 중  $a_{ijk}=1, i=i, j \rightarrow c$ 인 원소의 수,  $N(j|c)$ 는 기술범주  $c$ '에 속한 기술 세부 분류 수,  $N(a_{ijk}=1, i=i, k=k)$ 는 A의 원소 중  $a_{ijk}=1, i=i, k=k$ 인 원소의 수이다.

예를 들어, 특정 정보생성단위  $i$ '가 기술 범

주  $c$ '의 모든 하위 세부 기술 분류의 모든 정보 유형을 생성하는 경우는  $CU(i, c)$ 가 1이고, 전혀 생성을 하지 않는 경우는 0을 가지게 된다. 특정 정보생성단위가 특정 기술범주에 대해서 정보생성을 균질적으로 생성하는 것은, 만일 해당 기술 범주의 정보 생성을 하는 경우에는( $CU \neq 0$ 인 경우에는), 1에 가까운 값을 갖는다는 것을 의미한다. 따라서, 정보생성단위별 기술범주 균질도는 다음과 같이 정의된다. 마찬가지로, 정보생성단위별 정보유형 균질도, 기술범주별 기술범주 균질도, 정보유형별 정보유형 균질도가 다음과 같이 정의된다.

#### Definition 4 (정보균질도2)

정보생성 단위  $i$ '의 기술범주 균질도

$$ICU(i) = \text{Average}(CU(i, c) | CU(i, c) \neq 0)$$

정보생성 단위  $i$ '의 정보유형 균질도

$$ITU(i) = \text{Average}(TU(i, k) | TU(i, k) \neq 0)$$

기술범주  $c$ '의 기술범주 균질도

$$CCU(c) = \text{Average}(CU(i, c) | CU(i, c) \neq 0)$$

정보유형  $k$ '의 정보유형 균질도

$$KTU(k) = \text{Average}(TU(i, k) | TU(i, k) \neq 0)$$

단,  $\text{Average}(CU(i, c) | CU(i, c) \neq 0)$ 는 특정 정보생성 단위  $i$ '의  $CU(i, c) \neq 0$ 인 것들 중  $CU(i, c)$ 들의 평균을 의미하며,  $\text{Average}(TU(i, k) | TU(i, k) \neq 0)$ ,  $\text{Average}(CU(i, c) | CU(i, c) \neq 0)$ ,  $\text{Average}(TU(i, k) | TU(i, k) \neq 0)$ 도 유사하게 정의된다.

정보생성단위에 따라서는 특정 기술 범주에 집중해서 정보를 생성하는 기관과 폭넓은

기술 범주에 대해서 정보를 생성하는 기관이 존재한다. 이를 정보 생성의 집중도(Concentration)로 정의한다. 정보 생성의 균질도는 균질도가 높을수록 바람직한 반면에, 정보 생성의 집중도는 높은 기관(특정 범주의 정보를 집중적으로 생성하는 기관)도 필요하고, 집중도가 낮은 기관(폭넓은 범주의 정보를 포괄적으로 생성하는 기관)도 함께 필요하다.

#### Definition 5 (정보집중도1)

정보생성 단위  $i'$ 의 기술범주  $c'$ 에 대한 집중도

$$CC(i', c') = \frac{N(a_{ijk}=1, i=i', j \rightarrow c')}{N(a_{ijk}=1, i=i')}$$

정보생성 단위  $i'$ 의 정보유형  $k'$ 에 대한 집중도

$$TC(i', k') = \frac{N(a_{ijk}=1, i=i', k=k')}{N(a_{ijk}=1, i=i')}$$

여기서,  $N(a_{ijk}=1, i=i')$ 는  $A$ 에서  $a_{ijk}=1$ 이고,  $i=i'$ 인 원소의 수이다.

정보 집중도 역시 정보 균질도와 마찬가지로 0과 1사이의 값을 갖는다. 만일 특정 정보생성단위에서 특정 기술범주 내의 정보만을 생성하는 경우는 해당 정보집중도가 1이 된다. 또한 정보 균질도와 유사하게 정보 집중도에 대해서도 정보생성단위별 기술범주 집중도, 정보생성단위별 정보유형 집중도, 기술범주별 기술범주 집중도, 정보유형별 정보유형 집중도가 다음과 같이 정의된다.

#### Definition 6 (정보집중도2)

정보생성 단위  $i'$ 의 기술범주 집중도

$$ICC(i') = \text{Average}(CC(i', c) | CC(i', c) \neq 0),$$

정보생성 단위  $i'$ 의 정보유형 집중도

$$ITC(i') = \text{Average}(TC(i', k) | TC(i', k) \neq 0)$$

기술범주  $c'$ 의 기술범주 집중도

$$CCC(c') = \text{Average}(CC(i, c') | CC(i, c') \neq 0)$$

정보유형  $k'$ 의 정보유형 집중도

$$KTC(k') = \text{Average}(TC(i, k') | TC(i, k') \neq 0)$$

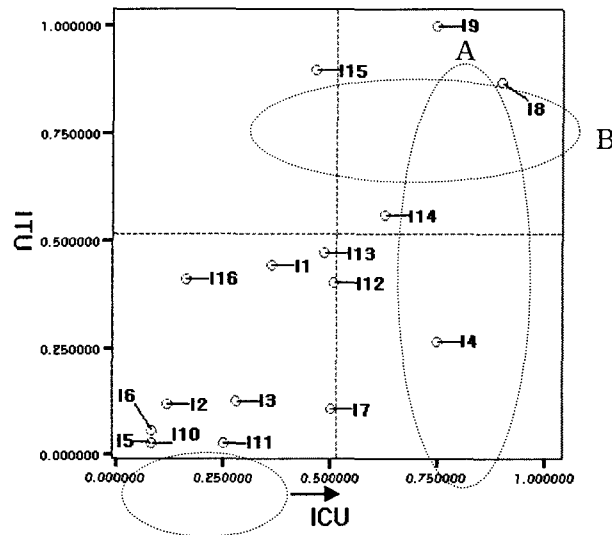
## 4.2 균질도와 집중도 예제

앞에서 언급한 연구기반조성 사업에 대한 정보생성행렬에 대하여 정보균질도와 정보 집중도를 계산하였다. (그림 1)은 정보생성 단위별 기술범주 균질도(ICU), 정보유형 균질도(ITU)를 2차원에 도식적으로 표현한 것이다. 그림에서 타원 A 내부에 존재하는 점들(I8, I9, I4)은 기술범주 균집도가 상대적으로 높은 정보생성단위들이고, 타원 B 내부에 존재하는 점들(I9, I15, I8)은 정보유형 균집도가 상대적으로 높은 정보생성단위들이다. 타원 C 내부에 존재하는 점(I5, I6, I10, I11, I12, I13)은 두 가지 균질도가 모두 상대적으로 낮은 정보생성단위들이므로, 균질도를 높일 수 있도록 정보 생성을 유도하는 것이 필요하다.

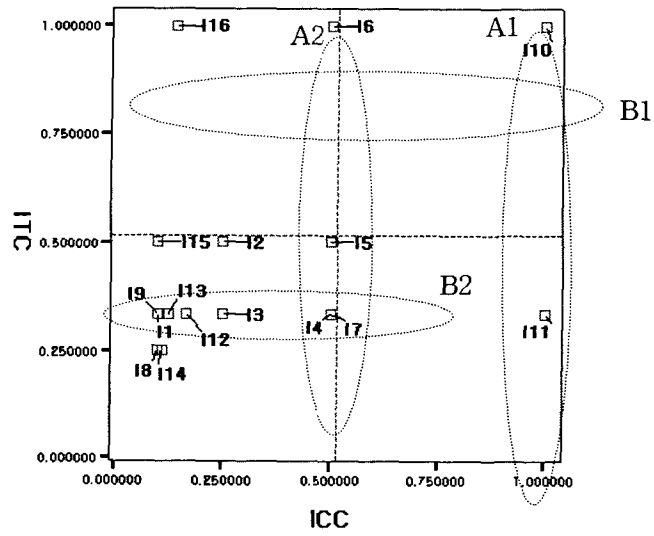
(그림 2)는 정보생성단위별 두 가지 집중도(ICC, ITC)를 2차원 평면에 표시한 것이다. 타원 A1, A2에 포함된 점들이 상대적으로 기술범주 집중도(ITC)가 높은 정보생성단위로, A1에 속한 I10, I11의 경우는 한 가지 기술범주의 정보만을 전문적으로 생성하는 정보생성단위이고, I6, I5, I4, I7은 두개의 범주의 정보만을 전문적으로 생성하는 정보생성단위이다. 마찬가지로 타원 B1, B2에 속한 점들에 대한 해석이 가능하다. (그림 2)의 좌측하단에 위치한 I8, I14 등은 포괄적인 정보생성단위로, 다양한 기술범주와

정보유형을 생성하는 정보생성단위이다. (그림 3)은 기술범주별 균질도(CCU), 집중도(CTU)를 2차원 평면에 표시한 것이다. 그림에서 J4, J8에 해당하는 기술범주가 상대적으로 균질도가 높은 기술범주이다. J4, J8은 각각 인터넷, 반도체 및 부품에 해당하므로, 인터넷, 반도체 관련 정보들은 다른 기술범주에 비해서 상대적으로 균질도 높게(빠짐없이) 생성되고 있음을 알 수 있다. (그림 4)는 정보유형별 균질도(KTU), 집중

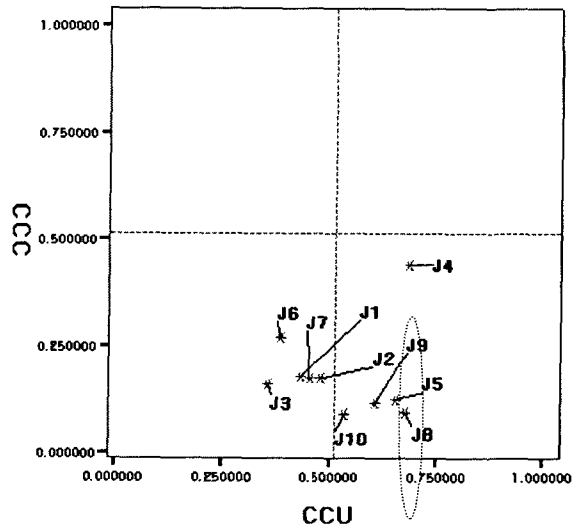
도(KTC)를 표현한 것이다. K2(시장정보)가 정보생성 집중도가 가장 높음을 알 수 있고, K4(기타정보)는 균질도는 가장 높으나, 집중도는 가장 낮음을 알 수 있다. 즉, 기타 정보를 생성하는 정보생성단위의 경우는 모든 세부분류를 상대적으로 고루 생성하고 있으나, 기타정보만을 전문적으로 생성하는 정보생성단위는 상대적으로 적음을 알 수 있다.



(그림 1) 정보생성단위별 균질도

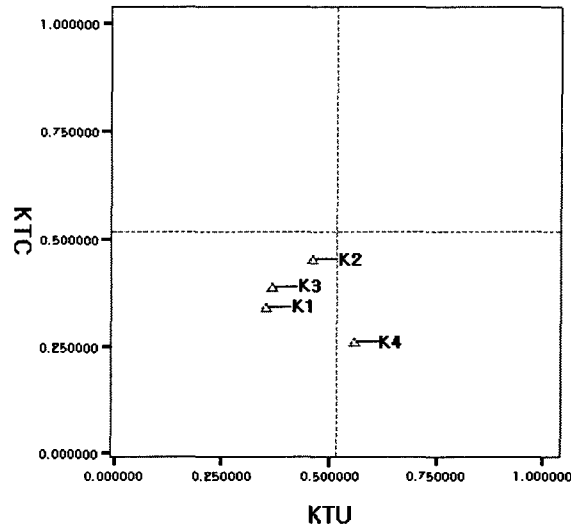


(그림 2) 정보생성단위별 집중도



(그림 3) 기술범주별 균질도와 집중도





(그림 4) 정보유형별 균질도와 집중도

### 4.3 정보균질도와 연관성 분석

본 논문에서 제시한 기술범주 정보균질도는 특정 기술범주에 속한 하위 세부 기술 분류의 정보들이 빠짐없이 생성되는지에 대한 척도이다. 이것은 동일한 기술 범주에 속한 세부 기술 분류 정보들의 어느 정도 함께 생성되는지와 상당히 관련이 있는 개념이다. 동일한 세부 기술 분류 정보들이 얼마나 자주 함께 생성되는지를 검토할 수 있는 방법으로는 데이터 마이닝에서 자주 사용되는 연관성 분석(Association Rules Analysis)이 있다[1,3]. 연관성 분석 방법은 장바구니 분석(Market Basket Analysis)이라고도 하는데, 하나의 장바구니에 함께 구매되는 제품을 분석함으로써 제품간의 연관성을 찾아내는 기법이다. 연관성 분석 기법을 본 연구에 적용하기 위해서는 동일한 정보생성단위에서 생성되는 정보들을 하나의 제품 항목들로 보고, 동일한 정보생성단위를 하나의 장바구니로 취급함으로써 가능하다. 연관성 분석에서는 크게 리프트(Lift),

지지도(Support), 신뢰도(Confidence) 3개의 척도를 가지고 연관성 규칙의 유용성을 평가한다. 이중 신뢰도는 연관성의 강도를 나타내는 척도로, 예를 들어 “만일 상품 A를 구매하면, 상품 B를 구매한다”라는 규칙이 있다면 이 규칙의 신뢰도 값은 조건부확률 값인  $\text{Prob}(\text{상품 B 구매} | \text{상품 A 구매})$ 로 계산된다.

본 연구에서는 기술범주들간의 연관성(즉, 하부 세부기술분류들이 얼마나 함께 정보가 생성되는지)을 검토하기 위해서 SAS Enterprise Miner를 사용하여 연관성 분석을 수행하였다[1,12]. [부록 2]는 연관성 분석 결과의 일부로 기술범주 J4(인터넷)의 정보들간의 신뢰도 값들을 보여준다. 즉, 행렬의 첫번째 줄의 처음에 있는 37.5은 인터넷기반에 대한 기술정보가 생성되면 인터넷기반의 시장정보가 생성된다는 규칙의 신뢰도 값으로, 인터넷기반에 대한 기술정보를 생성하는 정보생성단위 중 37.5%가 인터넷기반에 대한 시장정보를 함께 생성한다는 것을 의미한다.

〈표 3〉 기술범주별 균질도와 연관성 분석의 연결강도와의 비교

순위	기술범주별 균질도		연관성 분석	
	기술범주	CCU	기술범주	평균연결강도(%)
1*	J4	0.685185	J4	24.14773
2*	J8	0.678571	J8	22.09821
3*	J5	0.654762	J5	19.31818
4	J9	0.607143	J2	17.39583
5	J10	0.535714	J9	16.28788
6	J2	0.488542	J7	15.15152
7	J7	0.453704	J10	14.73214
8*	J1	0.432292	J1	12.88496
9*	J6	0.386364	J6	12.12121
10*	J3	0.355556	J3	11.01974

〈표 3〉의 5번째 열의 값들은 앞에서 구한 신뢰도 값들은 각 기술범주별로 평균한 값이다. 즉, 기술범주 J4의 값 24.14773(%)는 [부록 2]의 행렬의 값들의 평균값이다. 〈표 3〉은 기술범주별 균질도와 연관성 분석의 연관도와 비교하기 위해서 작성되었다. 〈표 3〉에서 볼 수 있듯이 본 논문에서 제시한 기술범주별 균질도(CCU)의 순위가 연관성 분석의 결과로 얻어진 순위가 상당히 관련성을 가짐을 알 수 있다. 실제로 상위 3개 순위와 하위 3개 순위에 해당하는 기술범주가 동일함을 알 수 있다. 따라서 본 연구에서 제시한 기술범주별 균질도가 정보 생성의 연관도와 밀접한 관계를 가지고 있음을 알 수 있다. 하지만 본 논문에서 다루는 문제에 연관성 분석만을 적용하면 발생하는 문제점은, 생성되는 정보 항목간의 연관도를 파악할 수 있지만, 정보생성단위에 관한 정보(예를 들어, 어떤 정보생성단위가 동일 기술범주 내의 세부기술분류를 빠짐없이 생성하는지)를 제공할 수 없다는 것이다.

## 5. 결론 및 추후 연구 과제

국가 정보 및 지식 체계의 중요성이 강조되고 있으며, 실제적으로 필요함에도 불구하고 국가 수준의 정보 및 지식 생성, 배분 체계의 분석을 위한 이론적인 연구들은 기업 수준의 지식 관리에 대한 연구들에 비해서 부족한 형편이다. 본 논문에서는 국가 수준의 연구개발 정보의 생성 현황을 파악하고 분석하기 위한 개념으로 균질도(Uniformity)와 집중도(Concentration)를 제시하였다. 균질도는 특정 기술범주 내에 하부 세부기술분류의 정보들이 빠짐없이 생성되는지를 나타내는 척도이고, 집중도는 정보생성단위(기관, 과제, 또는 주체)가 특정 정보를 전문적으로 생산하는지, 아니면 다양한 정보를 포괄적으로 생산하는지를 검토할 수 있는 척도이다. 균질도와 집중도는 정보생성의 소재를 조사하는 정보생성행렬(Information Creation Matrix)로 통해서 생성될 수 있으며, 정보생성단위별, 기술범주

별, 정보유형별로 2차원 평면에 도식적으로 표현할 수 있다. 이러한 가시화를 통해서 정보생성단위의 상대적인 균질도와 집중도를 검토할 수 있으며 정보생성단위의 군집(cluster)을 생성하는데도 활용이 가능하다. 추후 연구과제로는 균질도와 집중도 개념을 보다 형식화하기 위한 노력으로 기술분류 계층도가 일반적으로 다층 구조를 갖는 경우에 대한 이론적인 확장이 필요하다. 또한 다양한 분야에 실제적인 적용을 통해서 두 척도의 유용성을 심도있게 검토하고, 정보생성 현황을 분석하는데 유용한 추가적인 척도의 도출이 필요하다.

## 참 고 문 헌

- [1] 강현철, 한상태, 최종후, 김은석, 김미경, 데이터마이닝-방법론 및 활용, 자유아카데미, 1999.
- [2] 김종우 외, 국가 정보 생성/제공기관 현황 조사, 한국전자통신연구원, 2001.
- [3] Berry, M.J.A. and Linoff Gordon, Data Mining Techniques For Marketing, Sales, and Customer Support, John Wiley & Sons, Inc., 1997.
- [4] CISTI (Canada Institute for Science and Technical Information), <http://www.nrc.ca/cisti>, 2002.
- [5] Clark, Howard C., The Growth of Canadian Knowledge Networks, Journal of Knowledge Management, Vol. 3, No. 4, 1999, pp.304-307.
- [6] CNI (Coalition for Networked Information), <http://www.cni.org>, 2002.
- [7] Finkelstein, Clive, An Introduction to Information Engineering: From Strategic Planning to Information Systems, Addition-Wesley Publishing Company, 1989.
- [8] KM-IP (National Knowledge and Intellectual Property Management) Task Force, <http://www.km-iptask.org>, 2002.
- [9] Lucier R.E., Knowledge Management: Refining Roles in Science Communication, EDUCOM Review, Fall, 1990.
- [10] Martin, James, Information Engineering, Book II: Planning & Analysis, Prentice-Hall, NJ, 1990.
- [11] OITDA (Optoelectronic Industry and Technology Development Association), <http://www.oitda.or.jp>, 2002.
- [12] SAS Institute Inc, Data Mining Using SAS Enterprise Miner Software. SAS Institute Inc. 1997.

## [부록]

## [부록-1] 특정기관의 정보생성행렬 예

대분류	중분류	기술정보	시장정보	정책정보	기타정보
유선통신	서비스				
	교환/전송기술	1	1	1	
	가입자망				
	구내망				
	광통신기술	1	1	1	
	단말기술		1	1	
무선기술	서비스				
	이동교환/전송	1	1	1	
	위성교환/전송	1	1	1	
	단말기술				
방송/TV	방송서비스				
	방송전송기술		1	1	
	방송신호처리		1	1	
	스튜디오기기				
	방송수신/단말				
인터넷	인터넷기반		1	1	
	인터넷지원				
	인터넷활용				
컴퓨터/ 주변장치	하드웨어		1	1	
	컴퓨터통신기술		1	1	
	주변장치		1	1	
소프트웨어	패키지소프트웨어		1	1	
	컴퓨팅서비스				
	지식정보처리소프트웨어				
멀티미디어 컨텐츠	디지털컨텐츠제작		1	1	
	교육용컨텐츠				
	오락게임용컨텐츠				
반도체 및 부품	반도체	1	1	1	
	부품	1	1	1	
정보가전	홈네트워킹		1	1	
	홈서버/정보가전단말		1	1	
	정보가전기반소프트웨어				
정보보호	공통기반기술		1	1	
	시스템/네트워크보호	1	1	1	
기 타					

참고) 기타정보는 주요 전시회 및 포럼 정보, 전문인력 DB 등을 포함함

표의 각 셀 중 공백은 0을 의미함

## [부록-2] 연관성 분석의 예 (J4: 인터넷 기술 범주)

		인터넷기반				인터넷지원				인터넷활용			
		기술 정보	시장 정보	정책 정보	기타 정보	기술 정보	시장 정보	정책 정보	기타 정보	기술 정보	시장 정보	정책 정보	기타 정보
인 터 넷 기 반	기술 정보	-	37.5	25	12.5	37.5	37.5	25	12.5	37.5	37.5	25	12.5
	시장 정보	37.5	-	50	18.75	37.5	50	37.5	18.75	37.5	18.75	50	31.25
	정책 정보	25	50	-	12.5	25	37.5	37.5	12.5	25	31.25	31.25	0
	기타 정보	12.5	18.75	12.5	-	12.5	18.75	12.5	18.75	12.5	12.5	0	12.5
인 터 넷 지 원	기술 정보	37.5	37.5	25	12.5	-	37.5	25	12.5	37.5	37.5	25	12.5
	시장 정보	37.5	50	37.5	18.75	37.5	-	37.5	18.75	37.5	43.75	31.25	12.5
	정책 정보	25	37.5	37.5	12.5	25	37.5	-	12.5	25	31.25	31.25	6.25
	기타 정보	12.5	18.75	12.5	18.75	12.5	18.75	12.5	-	12.5	12.5	0	12.5
인 터 넷 활 용	기술 정보	37.5	37.5	25	12.5	37.5	37.5	25	12.5	-	37.5	25	12.5
	시장 정보	37.5	50	31.25	12.5	37.5	43.75	31.25	12.5	37.5	-	31.25	12.25
	정책 정보	25	31.25	31.25	0	25	31.25	31.25	0	25	31.25	-	6.25
	기타 정보	12.5	12.5	0	12.5	12.5	12.5	6.25	12.5	12.5	12.5	6.25	-

## ■ 저자소개



### 김종우

현재 충남대학교 통계학과에서 조교수로 재직중이다. 서울대학교 수학과에서 이학사(1989), 한국과학기술원 경영과학과에서 공학석사(1991)와 산업경영학과에서 공학박사(1995)를 취득하였다. 주요 관심분야는 전자상거래, 데이터마이닝 응용, 지능형 정보시스템, 정보시스템 분석 및 설계 등이다.



### 이성용

영남대학교 컴퓨터공학과에서 공학사(1997), 고려대학교 전산학과에서 이학석사(1999)를 취득하였으며, 현재 한국전자통신연구원 IT정보센터에 연구원으로 재직중이다. 주요 관심분야는 정보보호, 전자상거래, 정보시스템 설계 등이다.



### 주영진

연세대학교 경제학과에서 학사, 한국과학기술원 경영과학과에서 석·박사를 수여받고, 한국전자통신연구원 선임연구원/팀장을 거쳐 2001년 3월부터 배재대학교 경영정보학부에 재직중이다.



### 정현수

숭실대학교 전산학과에서 석·박사(1995)를 수여받고, 1989년에 정보처리기술사를 취득하였으며, 현재 한국전자통신연구원 정보체계연구팀 팀장으로 재직중이다. 주요 관심분야는 정보보호, IR, 전자상거래, Emerging Tech 등이다.