

GIS 컴포넌트 추출기법

(GIS Component Extraction Method)

박태옥(朴泰玉),* 김계현(金桂炫)**

Taeog Park, Kye Hyun Kim

요약 오늘날 정보시스템을 구축하는 모든 분야에서 컴포넌트에 기반한 개발(CBD : component based development) 방법이 주 흐름으로 등장하였다. GIS 영역에서도 컴포넌트의 개발 및 조립을 통한 시스템 구축을 권장하고 있다. GIS 영역은 다른 영역에 비하여 컴포넌트의 식별이 뚜렷한 업무영역(business domain)을 비교적 많이 가지고 있다. 이 영역은 비교적 산술적 기능을 많이 요구하는 부분이다. GIS 영역에서나 전사적(enterprise) 영역에서나 비산술적 기능 부분이 많이 존재함을 인식하여야 한다.

컴포넌트는 재사용성을 보장하여야 하는 특징을 가지고 있다. 재사용성이란 가능한 기능이 단순하고 최소의 크기를 가질 때 효용가치가 상승한다. 몇 연구에서 GIS 컴포넌트를 추출하는 기법을 제시한 바가 있다. 본 연구에서는 친화력분석(affinity analysis)이란 기법을 통하여 GIS 컴포넌트를 추출하고자 한다. CBD에서는 UML을 활용하는 것이 기본으로 쓰임새(use case)와 클래스는 UML의 핵심을 이루는 요소이다. 이 쓰임새와 클래스가 해당 업무영역 내에서 얼마나 친화력을 가지는가를 살펴보는 것이 친화력 분석의 목적인데, 이 분석 결과를 가지고 컴포넌트를 식별하게 된다.

이 친화력 분석은 GIS 영역뿐만 아니라 모든 업무영역에 활용이 가능하고 분석의 수행 절차가 복잡하지 않음으로 널리 활용할 수 있을 것으로 사료된다.

ABSTRACT The CBD(component based development) method has recently been on the rise as the main current among all fields of developing Information Systems.

The developing system by building and integrating of components is encouraged in GIS sphere. Business domain which considerably identifies components occupies GIS sphere rather than any other sphere. Arithmetic function is quite needed in GIS sphere but non-arithmetic functions are also used in GIS sphere as well as in enterprise sphere.

Component is characterized by ensuring the reuse of itself. The efficiency of reusing component is promoted as the component is functionally simple and is packed to a minimum. The GIS Component Extraction Method has already been presented in several studies but this study is about the GIS Component Extraction by the Affinity Analysis Method. CBD uses UML on a basis and the core of the UML is consisted of the use case and the class. The purpose of the Affinity Analysis is a study how relate between the use case and the class in a certain business domain and then the results identify the component.

The Affinity Analysis is useful not only in GIS but also in every business domain and is considered to be popular as the procedure of this method is not complex at all.

키워드 : 컴포넌트, 개발방법론, GIS, CBD, UML, 소프트웨어 개발, 컴포넌트 추출

1. 서론

GIS 구축시 컴포넌트에 기반한 개발(CBD)을 정부에서도 장려할 뿐만 아니라 여러 기관에서도 적극 수용하고 있다[1][2][3]. 이는 오늘날 시스템 개발 환

영인 CBD의 대세에 따르는 고무적인 일이다. CBD란 시스템이 컴포넌트들의 조합에 의하여 구축된다는 의미로서 시스템 내에 존재하는 컴포넌트들의 식별이 CBD의 핵심적인 요소가 된다. 식별된 컴포넌트들은 재사용성과 상호운용성의 특징을 지녀야 반드시 한다.

* 한국정보통신대학교 정보통신교육원 팀장

topark@aiit.or.kr

** 인하대학교 지리정보공학과 부교수

kyehyun@inha.ac.kr

재사용성이란 식별된 컴포넌트가 다른 시스템에서도 사용될 수 있는 매우 높은 확률을 가지는 것이며, 상호운용성은 어떠한 플랫폼(마이크로소프트사의 COM/DCOM, OMG의 Corba, 선마이크로 시스템사의 EJB 등) 하에서도 운영될 수 있어야 함을 의미한다. 이런 조건을 충분히 만족하려면 CBD를 지원하는 방법론의 도입이 필수적이다. 그러나 GIS를 제외한 다른 IT 분야에서는 CBD 방법론의 도입이 이루어지고 있으나, GIS에서는 거의 부재한 실정이다[4]. 특히 방법론의 프로세스 중에서 컴포넌트를 식별하는 프로세스는 재사용을 높이는 컴포넌트를 추출하여야 하는바 구체적인 식별 절차가 제시되어야 한다. 추출된 컴포넌트가 하나의 응용시스템과 유사한 사례가 되지 말아야 할 것이다.

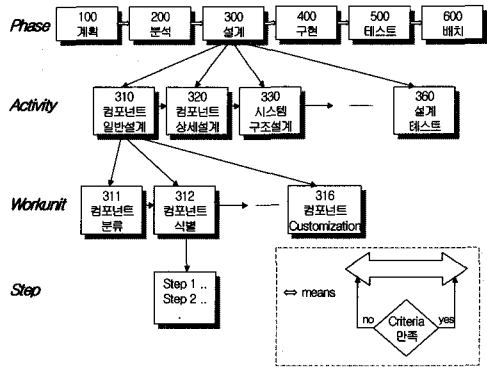
따라서 본 연구에서는 CBD 방법론의 프로세스를 따라 컴포넌트 식별 단계에서 가능한 기능이 단순하고, 크기는 최소를 유지하며, 최대의 재사용 확률을 지닌 컴포넌트 추출 기법을 제시하고자 한다. 선행 연구가 수행된바 있으나 이들과는 차별화된 방법으로서 친화력 분석(affinity analysis) 기법을 적용한다. 이는 쓰임새(use case)와 클래스(class)간의 관계에 따라 클래스들간의 친화력을 파악하는 것으로 컴포넌트 추출의 뛰어난 기법으로 여겨진다.

본 연구는 제2장에서는 본 연구의 선행연구로서, CBD 방법론의 프로세스 및 추출 기법들을 살펴 볼 것이다. 컴포넌트를 추출하는 프로세스 및 기법을 제3장에서 설명하며, 4장에서는 사례를 가지고 적용하며, 컴포넌트를 추출하는 프로세스는 실제로 구현된 모습을 제공할 것이다. 마지막 5장에서는 결론 및 향후 연구 방향에 대하여 언급할 것이다.

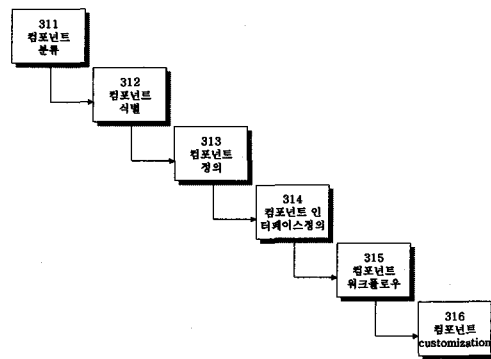
2. 관련 연구

2.1 CBD 방법론의 프로세스

GIS을 지원하는 CBD 방법론은 국내외를 통틀어서 거의 없는 실정이다. 그러나 본 연구에서는 선행 연구의 컴포넌트 기반 GIS 개발방법론(ATOM)을 활용한다. ATOM의 아키텍처가 <그림 1>에 나타나 있다. 이 프로세스에서 컴포넌트 식별과 관련된 프로세스를 추출하여 제3장에서 컴포넌트 추출 절차를 구성할 것이다. 본 연구는 설계 단계(300)의 활동 중 컴포넌트 일반 설계 프로세스(310)의 “컴포넌트 식별(312)”을 해결하는 추출 기법을 제시 및 적용하는 과정을 설명할 것이다<그림 2>. 그리고 GIS 컴포넌트 분류 체계 역시 ATOM의 분류 체계를 따른다<표 1>[4].



<그림 1> ATOM의 아키텍처



<그림 2> 컴포넌트 일반설계 프로세스

2.2 컴포넌트 추출 기법

컴포넌트 추출 방법은 여러 연구에 의하여 제시되었다. 쓰임새와 클래스 사이의 가중치(weight)를 분석하여 컴포넌트를 추출하는 방법[5], 응집도(cohesion)와 결합도(coupling)를 가지고 컴포넌트를 식별해 내는 방법[6], 쓰임새와 클래스 사이를 상호 교환에 의한 클러스터링 기법에 의하여 컴포넌트를 분리해내는 방법 등이 있다[7]. 이들은 분석 기법을 적용하여 컴포넌트를 식별해 내지만, 많은 방법론들이 나름의 논리적인 근거를 바탕으로 컴포넌트 식별 방법을 제시하고 있으나 설득력이 부족한 편이다. 그러나 이들 컴포넌트 추출 방법은 모든 시스템 영역에서 활용 가능하나, GIS 분야에 적용한 사례는 단지 쓰임새와 클래스 사이의 교환 법칙에 의한 클러스터링 기법을 GIS 도메인의 특성을 고려하여 컴포넌트 추출 방법을 제시한 적이 있다[4].

〈표 1〉 GIS 컴포넌트 설명

분류	설 명	비 고
기반	다양한 GIS 응용시스템에서 필수적으로 필요한 기능을 표준화된 인터페이스를 통하여 제공하는 컴포넌트	· 공간 데이터 액세스, 조작, 매핑 등
고유	GIS 기반 컴포넌트 외에 GIS 고유의 요소 기술을 표준화된 인터페이스를 통하여 제공하는 컴포넌트	· 네트워크 분석 · GRID 분석 · TIN 분석 기능등
공통	GIS를 포함하여 다른 정보시스템에서 공통적으로 사용 가능한 컴포넌트	· 통계, 차트, 문서 작성 등
레이어	GIS만이 가지는 feature의 공간정보(위치 및 위상)를 제공하기 위한 컴포넌트	· 배수구역도 · 하천
GIS 비즈니스	GIS 도메인 내에서 업무 로직을 처리하는 최소의 기능과 크기를 가진 컴포넌트	· 통수능분석 · 우수관망분석
응용	특정 업무를 위하여 위의 컴포넌트들을 조립하여 만들어진 어플리케이션	· 도로시설물관리응용 · 상하수도관리응용

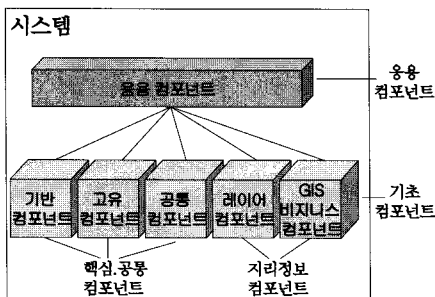
GIS 도메인에 적용한 클러스터링 기법은 쓰임새와 클래스간의 원시 CRUD 매트릭스(표 2)를 클러스터링된 CRUD 매트릭스(표 3)로 변환하는 것으로 굵은 선으로 표시된 네모 상자가 각각 컴포넌트로 인식되는 방법이다. 이렇게 식별된 컴포넌트는 〈그림 3〉의 컴포넌트 아키텍처에 기술되어 있듯이 기반, 고유, 공통, 레이어, GIS 비즈니스 및 응용의 6개의 컴포넌트로 분류한다. 본 연구에서도 〈그림 3〉의 컴포넌트 아키텍처와 〈표 1〉의 컴포넌트 설명을 적용하여 설명할 것이다.

〈표 2〉 원시 CRUD 매트릭스

	C1	C2	C3	C4
U1	C	R	C	R
U2	R	R	C	C
U3	R	C	R	R
U4	R	C	R	R

〈표 3〉 클러스터링된 CRUD 매트릭스

	C1	C3	C4	C2
U1	C	C	R	R
U2	R	C	C	R
U3	R	R	R	C
U4	R	R	R	C

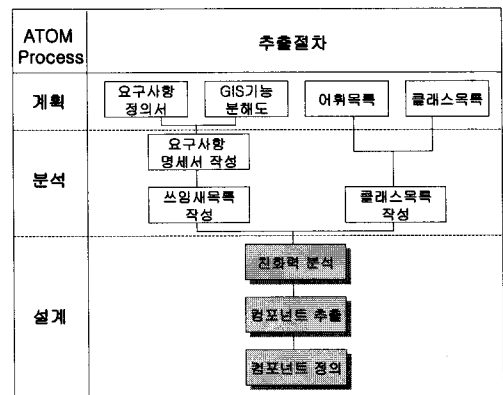


〈그림 3〉 컴포넌트 아키텍처

3. 컴포넌트 추출 절차

3.1 추출절차

본 연구의 컴포넌트를 추출하기 위한 절차는 〈그림 4〉와 같다. 왼쪽은 ATOM 프로세스의 단계(phase)에 해당하고, 오른쪽은 단계별 주요 산출물 중의 일부로서 컴포넌트 추출에 직접 관여하는 것들만 도식하였다. 추출 절차 중에서 계획 및 설계에 해당하는 산출물에 대한 설명은 선행 연구를 참조하면 되고[4], 본 연구에서는 컴포넌트 추출의 핵심 기법인 친화력 분석과 분석 결과를 가지고 컴포넌트를 추출 및 정의하는 방법을 자세히 설명할 것이다.



〈그림 4〉 컴포넌트 추출절차

3.2 친화력 분석(affinity analysis)

친화력 분석은 쓰임새들과 클래스들간의 관계를 이용하여 클래스들간의 친화력을 파악하는 것으로서 분

식 절차는 다음과 같다[8].

첫째, 쓰임새 각각에 대하여 번호(혹은 이름)를 부여하여 구분될 수 있도록 한다. U1, U2, 등.

둘째, 클래스 각각에 대하여 쓰임새와 같이 C1, C2,.... 등으로 부여한다. 여기서 친화력의 의미는, 만일 두 클래스가 동일 쓰임새와 전혀 관련이 없다면, 그들의 친화력은 제로가 되며, 만일 두 클래스가 어떤 쓰임새들에 대하여 같이 관련이 있다면, 그들의 친화력은 1이 된다. 여기서 클래스와 쓰임새가 관련이 있다는 의미는 쓰임새를 수행함에 있어서 어떤 클래스와 create, read, update, delete의 관계를 가지는 것을 말한다.

셋째, 각 클래스에 대하여 이 클래스를 활용하는 쓰임새의 번호를 모두 기재한다. C1={U1, U4, U13} 등으로 표시한다. C1 이 활용하는 쓰임새의 개수는 "(C1)=3"으로 표시한다.

넷째, 모든 클래스에 대하여 한 쌍(C1과 C2, C5와 C9 등)을 모두 활용하는 쓰임새들을 구하며, 개수는 "(C1, C2)=개수"로 표시한다.

다섯째, 이상의 내용들을 참조하여 두 클래스들에 대한 친화력 인자(factor)을 계산한다. 친화력 인자를 계산하는 한가지 방법은 다음의 식을 사용한다.

$$C1과 C2의 친화력 = \frac{a(C1, C2)}{a(C1)} \quad (1)$$

그리고, 친화력 인자는 <그림 5>의 매트릭스 형태로 나타나며, 만일 두 클래스들이 매우 높은 친화력을 가진다면 동일 컴포넌트에 있게되어 컴포넌트의 특성

을 설명하게 될 것이다.

여섯째, 클래스들을 컴포넌트 그룹을 만들기 위하여 클러스터링 알고리즘과 친화력 인자를 이용하여 클러스터링 한다(그림 6). 클러스터링 높은 값의 친화력을 가지는 클러스터들은 각각 의미 있는 컴포넌트에 속하는 클래스들로서 분명히 정의할 수 있으나, 어떤 기준 값에 미달하는 낮은 친화력을 가진 클래스 쌍들은 새로운 컴포넌트로서 정의하여야 할 부분이다. GIS 응용 도메인에서 이러한 낮은 값에 해당하는 업무 영역에 대한 분석가의 의사결정 방향이 정립되어야 한다.

3.3 컴포넌트 추출 및 정의

GIS 도메인에 친화력 분석 방법을 적용하기 위해서는 다음의 조건들이 사전에 준수되어야 한다. 첫째, 쓰임새도 작성의 입력이 되는 요구사항명세서에 지도 조작성을 위한 각각의 기능 및 GIS의 핵심 기능인 공간 분석에 대한 요구들이 기술되어야 한다. 만일 누락하면 기반 및 고유 컴포넌트의 누락을 가져올 수 있다. 이는 기존의 도구에서 제공하는 기능만을 활용함을 의미하므로, 각 프로젝트에서 요구되는 새로운 핵심·공동 컴포넌트의 창출을 차단하는 결과를 가져온다. 둘째, 특정 도메인에서 사용되는 레이어의 요구를 기술한다. GIS의 응용이 기본적으로 지도를 활용하는 작업이기 때문이며, 만일 누락하면 레이어를 정의하는 클래스가 누락되고, 또한 컴포넌트에서 반영될 수가 없다. 셋째, 클래스 정의할 때 레이어를 하나의 클래스로 간주한다. 물론 앞 조건에서 레이어를 요구조건으로 기술하는 것과 같은 의미이다.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11	C12
C1		0	0	0.92	0	0	0.01	0	0	0	0	0
C2	0		0	0	0	0	0	0.85	0	0	0.34	0.17
C3	0.01	0		0	0	0.12	0	0	0.07	0	0.38	0
C4	0.01	0	0		0.20	0.11	0	0	0.43	0.01	0	0
C5	0	0.02	0	0		0	0.21	0	0.08	0	0	0
C6	0.21	0	0	0.73	0		0.88	0	0	0.12	0.08	0.01
C7	0.35	0	0	0.76	0	0.30		0	0.01	0	0	0
C8	0	0	0	0	0	0	0		0	0.74	0	0
C9	0	0	0.21	0	0.09	0	0	0		0	0	0
C10	0	0	0	0.06	0	0.17	0	0	0		0	0.87
C11	0.01	0	0	0	0	0	0	0.90	0	0		0
C12	0	0.01	0	0	0	0	0	0	0	0.21	0	

<그림5> 친화력 인자 매트릭스(예)

클러스터링 알고리즘

1. 클래스 쌍을 친화력이 높은 값부터 정렬한다. 친화력이 가장 높은 것부터 시작한다. 가장 높은 친화력을 가진 클래스 쌍이 클러스터의 핵을 형성한다.
 C1, C4 (친화력 = 0.92) : 최초 핵심 클래스 쌍, 최초 핵심 클러스터
 C11, C8 (친화력 = 0.90)
 C6, C7 (친화력 = 0.88)
 C10, C12 (친화력 = 0.87)

2. 다음으로 높은 친화력을 가진 클래스 쌍의 하나의 클래스가 이미 형성된 핵심 클래스 쌍에 존재하는가 살펴보는 것이다. 따라서 C11과 C8, C6과 C7, C10과 C12는 서로 관련이 없으므로 각각 핵심 클러스터를 형성한다. 다음의
 C2, C8 (친화력 = 0.85)를 살펴보자
 C8이 이미 클래스 쌍 (C11, C8)에 할당되어 있으므로 C2를 (C11, C8)의 클러스터에 결합시킬 것인가를 결정하기 위하여 가중치 친화력(weighted affinity)을 계산한다.

$$\text{가중치 친화력} = \frac{(\text{affinity of C2 to C11}) * a(\text{C11}) + (\text{affinity of C2 to C8}) * a(\text{C8})}{a(\text{C11}) + a(\text{C8})}$$
 여기서, 만일 C11이 3개의 쓰임새에 의해서 활용되고, C8이 48쓰임새에 의해 활용된다면

$$\text{Composite affinity of C2 to cluster C11, C8} = \frac{0.34 * 3 + 0.85 * 48}{3 + 48} = 0.82$$
 0.82는 <그림 5>의 나머지 어떤 값보다 크므로 C2, C11, C8은 동일 클러스터에 속한다.
3. 다음으로 큰 친화력을 가진 것은 C7과 C4의 0.76이다. C7과 C4는 이미 (C1, C4)와 (C6, C7)에 있으므로 C7을 C1과 C4에 가중치 친화력을 구하고, 결합된 클러스터 C1, C4, C6에 복합 가중치 친화력 (composite weighted affinity)을 구한다. 이들 친화력이 각각 0.55와 0.37이라면, 이것은 그 다음 클래스 쌍의 친화력 C8과 C10(affinity = 0.74)보다 작기 때문에 결합에서 배제된다. 이러한 방법으로 가장 낮은 친화력을 가진 값까지 수행한다.
4. 이렇게 계산한 결과 친화력이 낮은(0에 가까운) 클래스 쌍들을 처리하는 문제가 대두되는데, 이것은 업무의 성격을 파악하여 이 작업을 수행하는 분석가 나름의 의사결정을 하여야 한다.

<그림 6> 클러스터링 알고리즘 설명

서술된 조건을 기반으로 친화력 분석을 실시하면 클러스터가 형성되고, 각각의 클러스터가 컴포넌트로 간주된다. 많은 클래스들로 이루어진 클러스터가 생성되는 것도 있는 반면, 한 두개의 클래스로 형성된 클러스터도 존재하는데, 후자의 경우를 하나의 컴포넌트로 간주할 것인가의 여부는 분석가의 업무 성격을 고려한 수작업에 의한 의사결정 부분이다.

추출된 컴포넌트는 전술한바와 같이 기초, 고유, 공통, 레이어, GIS 비즈니스 및 응용컴포넌트로 분류한다. 컴포넌트 인터페이스 정의 등을 포함하는 상세한 컴포넌트의 정의는 선행 연구의 절차 및 표기법을 따른다[4]. 다만 상위 계층의 응용컴포넌트는 기초컴포넌트들을 분석한 결과 응용컴포넌트의 정의에 부합되면 결합하여 응용컴포넌트로 등재한다.

4. 적용사례

본 친화력 분석에 의한 GIS 컴포넌트 추출은 하수

관망관리 업무를 대상으로 실시하였다. 하수관망관리 는 하수도 관망분석 및 우·오수 부족판거 검색 등에 초점을 맞춘다.

4.1 요구사항 명세서, 쓰임새 목록 및 클래스 목록 작성

요구사항 명세서는 기능분해도의 최하위 기능을 뽑아내어 상세 설명하는 것으로서, 해당 도메인 내에서 수행하여야 하는 모든 기능이 모두 포함되어야 한다(표 4). 쓰임새 목록은 요구사항 명세서를 입력으로 한다. 요구사항 명세서의 상세기능 각각을 쓰임새(1차)로 정의한다. 이 1차 쓰임새를 공통성·가변성 분석을 통하여 2차 쓰임새 목록을 창출한다. 이렇게 만들어진 2차 쓰임새들을 최종적인 쓰임새로서 목록에 기재된다, 26개의 쓰임새가 U1, U2, ...로 표기된다(표 5). 전술한 바와 같이 레이어를 포함하는 11개의 클래스 목록은 C1, C2, ...로 표기한다 [4]. 본 연구에서는 기반, 고유 및 공통에 해당하는 클래스는 생략하였다(표 6).

〈표 4〉 요구사항 명세서

기능번호	기능명	상세기능
1.1.1	기초환경변수설정	1.기초환경변수설정 메뉴 띄움 2.기초환경변수입력
1.1.2	강우강도설정	1.기초환경변수설정 메뉴 띄움 2.강우강도변수 입력
1.1.3	용도지역별계수설정	1.기초환경변수설정 메뉴 띄움 2.용도지역별계수 입력
1.2.1	우수배출용량 산정	1.환경변수 설정 2.관거추적을 통한 누계연장 산출 3.유하시간, 도달시간(t), 강우강도 산출 4.도시계획도 조회 5.우수배수구역도 조회 6.공종별 배수면적 산출 7.관거별 배수면적 및 누계면적 산출 8.관거별 유량 산출
1.2.2	우수배출용량 현황조회	1.하수관거 조회 메뉴 띄움 2.하수관 고유번호 선택 3.우수 배출용량 현황 조회
1.2.3	우수통수능부족관거검색	1.여유율 설정 2.우수관거에 대한 총하수 발생량 등을 산출 3.통수능 부족관거 검색
1.2.4	우수통수능부족관거 현황조회	1.우수통수능 부족관거 조회 2.해당 하수관 고유번호 조회
1.2.5	우수최소유속부족관거검색	1.총 발생하수량 산출 2.최대 처리하수량 산출 3.관거유속 산출 4.초기 관경 설정 5.관거유속 및 처리 유량 산출 6.적정 관경
1.2.6	우수최소유속부족관거현황조회	1.우수 최소유속 부족관거 검색 2.우수 최소유속 부족관거 조회 3.우수 최소유속 부족관거 출력
1.3.1 ~ 1.3.6	오수배출용량산정	우수와 내용이 동일함
1.4.1 ~ 1.4.6	합류배출용량산정	상 동
1.5	지도조작기능	1.지도조작관련 기능 수행
1.6.1	하수시설물도	1.하수시설물도 검색 2.하수시설물도 조회 3.하수시설물도 출력
1.6.2	도시계획도	1.공종별 배수면적 산출 2.도시계획별 배수면적 산출 3.도시계획도 검색 4.도시계획도 조회 5.도시계획도 출력
1.6.3	우수배수구역도	1.공종별 배수면적 산출 2.우수배수구역도 검색, 조회, 출력
1.6.4	오수배수구역도	1.도시계획별 배수면적 산출 2.오수배수구역도 검색, 조회, 출력

〈표 5〉 쓰임새 목록

쓰임새 번호	쓰임새명	비 고	쓰임새분류
U1	기초 환경변수 설정 메뉴 띄움		include
U2	환경변수 설정		include
U3	도시계획도 조회	검색, 출력 기능 포함	include
U4	관거별 배수면적 및 누계면적 산출		include
U5	하수관망해석	관거별 유량산출 별명	main
U6	우수관거 배출용량 산정	상위:하수관망해석	include
U7	오수관거 배출용량 산정	상위:하수관망해석	include
U8	합류관거 배출용량 산정	상위:하수관망해석	include
U9	하수관거조회 메뉴 띄움		include
U10	적정관경 제시		include
U11	환경변수 선택	가변성의 환경변수 설정에서 발생	extend
U12	우수배수구역도 조회	검색, 출력 기능 포함	main
U13	공종별 배수면적 산출		include
U14	우수배출용량 현황조회	상위:배출용량 현황조회	include
U15	우수통수능부족관거조회	상위:통수능부족관거조회	include
U16	우수최소유속부족관거조회 및 출력		include
U17	우수최소유속부족관거검색		include
U18	오수배수구역도 조회	검색, 출력 기능 포함	main
U19	도시계획별 배수면적 산출		include
U20	오수배출용량 현황조회	상위:배출용량 현황조회	include
U21	오수통수능부족관거조회	상위:통수능부족관거조회	include
U22	오수최소유속부족관거조회 및 출력		include
U23	오수최소유속부족관거검색		include
U24	여유율 설정		include
U25	지도조작관리기능 수행		main
U26	하수시설물도 조회	검색, 출력 기능 포함	main

〈표 6〉 클래스 목록

클래스 번호	클래스명	실 명
C1	관망분석	하수관망분석을 수행하기 위한 클래스
C2	환경변수	환경변수를 설정
C3	배수면적	오수, 우수, 합류 배수면적을 접근을 설정
C4	하수관망분석메뉴	하수관망을 수행하는 선택 메뉴 설정
C5	우수유량분석	우수관망분석 결과를 설정
C6	도시계획도	레이어
C7	우수배수구역도	"
C8	오수배수구역도	"
C9	하수시설물도	"
C10	배출용량현황조회메뉴	관거별 배출용량 현황을 조회하는 선택메뉴
C11	하수관거	하수관거 제원을 관리하기 위한 클래스

4.4 친화력 분석

4.4.1 클래스 활용하는 쓰임새 정의

각 클래스에 대하여 이 클래스를 활용하는 쓰임새의 번호를 모두 기재한다(표 7). 아울러 쓰임새 개수도 함께 기술한다.

〈표 7〉 클래스 활용하는 쓰임새들

클래스	쓰임새개수	쓰임새	클래스	쓰임새개수	쓰임새
C1	5	U6, U7, U8, U10, U24	C7	2	U18, U19
C2	4	U1, U2, U5, U11	C8	4	U4, U5, U12, U13
C3	2	U4, U13	C9	1	U26
C4	4	U5, U6, U7, U8	C10	6	U9, U10, U14, U15, U20, U21
C5	12	U6, U7, U8, U10, U14, U15, U16, U17, U20, U21, U22, U23	C11	7	U10, U14, U15, U16, U17, U22, U23
C6	5	U3, U4, U5, U13, U19			

4.4.2 클래스 쌍을 활용하는 쓰임새 개수 정의

모든 클래스에 대하여 한 쌍을 모두 활용하는 쓰임새들의 개수를 구한다. 대각선을 중심으로 대칭이 되므로 한쪽만 구한다.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1		0	0	3	4	0	0	0	0	1	1
C2			0	1	0	1	0	1	0	0	0
C3				0	0	2	0	2	0	0	0
C4					3	1	0	1	0	0	0
C5						0	0	0	0	5	7
C6							0	3	0	0	0
C7								0	0	0	0
C8									0	0	0
C9										0	0
C10											3
C11											

〈그림7〉 클래스 쌍을 활용하는 쓰임새 개수 매트릭스

4.4.3 친화력 인자 매트릭스

클래스 쌍의 친화력을 3.2절의 방정식 (1)을 이용하여 계산한다. 결과가 〈그림 8〉에 서술되어 있다.

4.4.4 클러스터링 및 전문가 분석

〈그림 6〉의 클러스터링 알고리즘을 적용하여 클러스터링을 수행한다. 먼저 수행 결과 클래스 쌍을 친화력이 높은 값부터 정렬하면 다음과 같다.

	C1	C2	C3	C4	C5	C6	C7	C8	C9	C10	C11
C1		0	0	0.6	0.8	0	0	0	0	0.2	0.2
C2	0		0	0.25	0	0.25	0	0.25	0	0	0
C3	0	0		0	0	1	0	1	0	0	0
C4	0.75	0.25	0		0.75	0.25	0	0.25	0	0	0
C5	0.33	0	0	0.75		0	0	0	0	0.41	0.58
C6	0	0.2	0.4	0.2	0		0	0.6	0	0	0
C7	0	0	0	0	0	0		0	0	0	0
C8	0	0.25	0.5	0.5	0	0.75	0		0	0	0
C9	0	0	0	0	0	0	0	0		0	0
C10	0.16	0	0	0	0.83	0	0	0	0		0.5
C11		0	0	0	1	0	0	0	0	0.42	

〈그림8〉 친화력 인자 매트릭스

- 1 = [C3, C6]
- 1 = [C3, C8]
- 1 = [C11, C5]
- 0.83 = [C10, C5]
- 0.8 = [C1, C5]
- 0.75 = [C4, C1], [C4, C5], [C5, C4], [C8, C6]
- 0.6 = [C6, C8], [C1, C4]
- 0.58 = [C5, C11]
- 0.5 = [C10, C11]

다음으로 높은 친화력을 가진 클래스 쌍의 하나가 클래스가 형성된 핵심 클래스 쌍에 존재하는가를 살펴보는가를 포함하여 〈그림 6〉의 3번, 4번을 수행한다. 특히 4번에서 분석가의 의사결정이 필요한 부분들이 발생하는데, 친화력 1을 가진 두쌍, [C3, C6]과 [C3, C8]을 살펴보면 C3이 공통으로 있으므로 클러스터링 될 가능성을 가지고 있으나 가중치 친화력은 0.67로서 C6과 C8은 동일 클러스터가 될 수 없다. 그러나 C3을 각각 포함하고 있다. 이럴 경우 C3의 업무적 성격을 한번 더 분석하여 독립된 클래스(클러스터)로 남겨두거나, C6이나 C8의 한쪽으로 남겨둔다. 또한 낮은 값을 가지는 C2, C7 및 C9는 독립된 클러스터로 간주한다. 이렇게 수행된 결과의 클러스터는 다음과 같다.

[C1, C4, C5, C10], [C3, C8], [C6], [C11], [C2], [C7], [C9].

4.5 컴포넌트 추출 및 정의

4.4.4 절의 클러스터링 및 전문가 분석을 통하여 7

개의 클러스터가 생성되었다. 각각의 클러스터를 컴포넌트라고 부르며 컴포넌트명과 기능을 정의할 것이다(표 7).

5. 결론 및 향후 연구방향

친화력 분석에 의한 컴포넌트 추출 및 정의는 사례 연구를 통하여 쉽게 적용될 수 있음을 확인하였다. 이 분석법의 개념은 단순하다. 그러나 쓰임새와 클래스와의 관계를 정확히 정의해야 하는 것을 기본으로 한다. 따라서 친화력의 분석 방법에 대한 이해도 중요하지만 해당 도메인에 대한 업무적 이해가 우선됨을 또한 이해하여야 한다. 업무적 이해도가 떨어지면 클러스터링 작업을 수행할 때 의사결정에 애로 사항을 겪게 된다. 산술적 계산에 의하여 단순히 클러스터(차후 컴포넌트가 됨)가 형성되지 않고 분석가의 의사결정 부분을 포함하는 번거로운 부분이 있으나 몇 번 이 분석 방법을 적용하여 보면 쉽게 클러스터를 도출할 수 있을 것으로 사료된다.

향후 보완해야 할 부분으로는 친화력 분석에 의한 컴포넌트 추출 과정을 비주얼 프로그래밍을 통하여 쉽게 적용할 수 있도록 시스템화하여야 할 것으로 생각된다. 전문가적 의사결정 부분은 전문가 시스템의 도입 또한 요망된다. 그리고 컴포넌트 구성에 대해서는 발견적해결방법(heuristic approach)이 다소 포함되어 있어 이에 대한 귀납적인 증명이 어려워 향후 귀납적 결과의 증명 방안에 대한 연구가 병행되어야 할 필요가 있다.

〈표 7〉 컴포넌트 추출 및 정의

컴포넌트명	포함되는 클래스	기능	분류
관망분석	C1, C4, C5, C10	하수배출용양 등 하수관거관리의 핵심 기능	비즈니스
우수배수구역도	C3, C8	도시계획도, 우수배수구역도, 오수배수구역도 등 레이어의 중첩을 통한 정보 추출을 위한 컴포넌트	레이어
도시계획도	C6		레이어
오수배수구역도	C7		비즈니스
환경변수설정	C2	관망분석을 위한 기초환경변수를 설정하는 기능	비즈니스
하수시설물도	C9	하수시설물의 위치정보를 파악	레이어
하수관거	C11	오수, 우수 및 합류 관거에 대한 속성 파악을 하는 기능	비즈니스

참고문헌

- [1] 이종훈, "개방형 컴포넌트 기술관점에서의 GIS산업 활성화 전략", 2000년 지식기반사회를 대비한 국가GIS 정책 및 기술방향에 관한 국제세미나, 2000, pp 175-190.
- [2] 김계현, 이우철, 김준철, "지자체 업무지원을 위한 하수도 관망해석 컴포넌트 설계", 개방형 GIS연구회지, 제2권 제1호, 2000, pp 27-36.
- [3] 윤태권, 정한일, "컴포넌트 소프트웨어 산업 동향", 2000년 정보처리학회 논문집, 2000, pp 79-84.
- [4] 박태욱, "UML을 활용한 컴포넌트 기반의 GIS 개발방법론에 관한 연구", 인하대학교 박사학위논문, 2002.
- [5] 유영란, "Use Case 및 클래스의 가중치 분석에 의한 컴포넌트 추출 기법", 숭실대학교 석사학위논문, 2001.
- [6] 강민수, "메트릭을 이용한 재사용 컴포넌트 추출", 서울대학교 석사학위논문, 2000.
- [7] 김수동, "Quick Overview of FOCUS Methodology", 발표자료, 2002.
- [8] James Martin, "Strategic Data-Planning Methodology", Prentice-Hall, 1982, pp 128-138.



김계현

1982년 한양대학교 자원공학과 졸업(공학사)
 1989년 아리조나 주립대 수문학과 졸업(공학석사)
 1993년 위스콘신주립대 토목환경공학과 졸업(공학박사)
 1995년 ~ 인하대학교 지리정보공

학과 부교수

관심분야: 환경, 수자원, 상하수도 및 지하시설물 관리분야의 GIS활용, 주제도 제작 및 GIS 표준화 등



박태욱

1981년 부산대학교 지구과학과 졸업(이학사)
 1985년 부산대학교 지구과학과 졸업(이학석사)
 2002년 인하대학교 지리정보공학과 졸업(공학박사)
 1982년~1996년 한국전자통신연

구원 선임연구원

1997년~현재 한국정보통신대학교 정보통신교육원 GIS팀장

관심분야: UIS, 개발방법론, 컴포넌트 GIS, Mobile GIS, 공간데이터베이스 등