

동시출현단어 분석을 활용한 비탈면 붕괴 예측 및 분석 연구에 관한 지적구조 분석

김선겸^{1*} · 김승현²

¹한국건설기술연구원 수석연구원, ²한국건설기술연구원 연구위원

Domain Analysis of Research on Prediction and Analysis of Slope Failure by Co-Word Analysis

Sun-Kyum Kim^{1*} · Seung-Hyun Kim²

¹Senior Researcher, Korea Institute of Civil Engineering and Building Technology

²Research Fellow, Korea Institute of Civil Engineering and Building Technology

Abstract

Although it is currently conducting slope management and research using digital technologies such as drones, big data, and artificial intelligence, it is still somewhat insufficient and is still vulnerable to slope failure. For this reason, it is inevitable to present the development direction for research on prediction and analysis of slope failure using the digital technologies to effectively deal with slope failure, which requires a preemptive understanding of prediction and analysis of slope failure. In this paper, we collected literature data based on the Web of Science for five years from January 1, 2016 to December 31, 2020 and analyzed by co-word analysis to identify the domain structure of research on prediction and analysis of slope failure. Detailed subject areas were identified through network analysis, and the domain relationships between keywords were visualized to derive global and regionally oriented keywords through relationship, centrality analysis. In addition, the clusters formed by performing cluster analysis were displayed on the multidimensional scaling map, and the domain structure according to the correlation between each keyword was presented. The results of this study reveal the domain structure of research on prediction and analysis of slope failure, and are expected to be usefully used to find future research directions.

Keywords: slope failure, analysis, domain structure, co-word analysis, network analysis, cluster analysis

OPEN ACCESS

*Corresponding author: Sun-Kyum Kim
E-mail: sunkyumkim@kict.re.kr

Received: 27 July, 2021
Revised: 17 August, 2021
Accepted: 19 August, 2021

© 2021 The Korean Society of Engineering Geology



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

초 록

최근 드론 및 빅데이터, 인공지능 등 디지털 기술을 활용한 비탈면 연구를 수행하고 있으나 다소 미흡한 실정이며, 여전히 비탈면 붕괴 대비에 취약하다. 이러한 이유로 비탈면 붕괴에 효과적으로 대처하기 위해 디지털 기술을 활용한 비탈면 붕괴 예측 및 분석 연구에 대한 발전방향을 제시하는 것이 필연적이며, 이를 위해 비탈면 붕괴 예측 및 분석에 관한 이해가 선제되어야 한다. 본 연구는 비탈면 붕괴 예측 및 분석 연구의 지적구조를 규명하여 연구방향을 제시하기 위해 2016년 1월 1일부터 2020년 12월 31일까지의 5년간의 Web of Science 기반으로 문헌 데이터를 수집하고 이를 동시출현단어를 활용하여 분석하였다. 네트워크 분석을 통하여 세부 주제 영역을 밝히고, 키워드 간의 지적 관계를 시각화하여 관계, 중심성 분석을 통한 전역 및 지역 중심성이 높은 키워드를 도출하였다. 또한 군집분석을 실시하여 형성된 군집을 다차원축적지도에 표시하였으며, 각 키워드들 간의 상관관계에 따른 지적구조를 제시하였다. 이러한

연구의 결과는 비탈면 붕괴 분석 및 예측 연구의 지적구조를 밝히고, 향후 연구 방향을 찾는 데 도움이 될 것으로 기대한다.

주요어: 비탈면 붕괴, 분석, 지적구조, 동시출현단어 분석, 네트워크 분석, 군집분석

서론

비탈면은 도로, 철도 선로, 건설재료 및 부지 등의 사회기반시설(SOC) 확보를 위한 목적으로 조성된다(Kim et al., 2011). 이 비탈면은 자연적으로 존재하는 자연비탈면과 깎거나 쌓아서 인공적으로 조성된 인공비탈면으로 구분할 수 있다(Lee et al., 2019). 자연비탈면과 인공비탈면은 일반적으로 불균질한 토양과 암반으로 구성되어 있고, 암반 불연속면의 상태와 방향성에 민감하며, 강우 등의 기상상태와 지진 등의 지반의 영향으로 항시 ‘붕괴’ 위험이 존재한다(Park et al., 2006). 따라서 비탈면 안정성 평가를 통해 사전에 붕괴위험 요인을 파악하고 이에 적절한 대응책을 수립해야 하며, 비탈면을 시공 한 이후에는 꾸준히 모니터링 및 유지보수를 통해 비탈면을 관리해야 한다(MLTMA, 2011). 최근 드론 및 빅데이터, 인공지능 등 디지털 기술을 활용한 비탈면 연구를 수행하고 있으나 다소 미흡한 실정이며, 여전히 비탈면 붕괴 대비에 취약하다. 이러한 이유로 비탈면 붕괴에 효과적으로 대처하기 위해 디지털 기술을 활용한 비탈면 붕괴 예측 및 분석 연구에 대한 발전방향을 제시하는 것이 필연적이며, 이를 위해 비탈면 붕괴 예측 및 분석에 관한 이해가 선제되어야 한다.

비탈면 붕괴 예측 및 분석의 이해를 위해서 관련 문헌에 대한 지적구조 분석을 통해 이루어질 수 있다. 지적구조 분석은 문헌이나 저자의 동시출현단어 분석, 서지결합법, 동시인용 분석 등의 기법이 활용된다. 이 가운데 동시출현단어 분석은 텍스트 전문, 제목, 초록, 키워드를 이용하여 주제영역으로부터 하위분야 및 분야의 패턴 등의 분석이 가능하며, 해당 분야의 발전 방향을 예측할 수 있는 근거가 된다(Seo and Chung, 2013).

본 연구의 목적은 이러한 동시출현단어 분석을 활용하여 비탈면 붕괴 예측 및 분석에 관한 연구 경향을 반영하는 지적구조를 도출하고 주제영역 구성을 밝히는 것이다. 이를 위하여 비탈면 붕괴 예측 및 분석에 관한 문헌 데이터를 수집하고 이를 분석하였다. Web of Science 데이터베이스로부터 수집한 2016년부터 2020년까지 5년간 발간된 비탈면 붕괴 예측 및 분석 주제의 총 396건의 문헌으로부터 제목과 초록에서 추출된 키워드는 총 2,678개였다. 추출된 키워드를 네트워크 분석과 시각화를 통해 중심 및 세부 주제영역을 파악하고, 중심성 분석을 통해 해당 키워드들이 해당 분야에 얼마만큼 영향을 주는지 도출하였다. 마지막으로 네트워크 분석을 보완하기 위해 군집분석을 활용하여 이를 다차원축적지도로 나타내어 비탈면 붕괴 분석 및 예측 영역의 전체적인 주제영역의 흐름 및 구성을 제시하였다. 본 연구를 통해 최근 비탈면 붕괴에 관한 분석 및 예측에 대한 학문적 구조 및 주제 분야에 대한 정보를 제공할 수 있으며, 이를 토대로 연구방향성 모색에 도움이 될 수 있을 것으로 기대한다.

본 논문의 구성은 다음과 같다. 먼저 관련된 지적구조를 위한 데이터 추출 및 행렬 작성에 대해서 살펴본다. 또한 비탈면 붕괴 예측 및 분석 연구의 지적구조를 위한 데이터 추출 및 행렬 작성을 설명하며, 지적구조 분석결과를 보여준다. 최종적으로는 연구를 수행한 결론과 의의를 논하는 것으로 마무리한다.

지적구조를 위한 데이터 추출 및 행렬 작성

지적구조 분석 과정

비탈면 붕괴 예측 및 분석 연구의 지적구조 분석을 위하여 문헌 데이터로부터의 키워드 추출과 분석이 필요하다. 이를 위한 전 과정을 Fig. 1에 나타내었다. 우선, 자료 수집과 키워드 선정을 위해 Web of Science에 등재된 저널에 게재된 문헌을

수집하였다. Web of Science 데이터베이스에서 제공하는 저널 범주를 지구과학 다학제(Geo sciences multidisciplinary), 지질공학(Engineering geological), 지질학(Geology)으로 제한하였고, 기간은 2016년 1월 1일부터 2020년 12월 31일까지의 데이터를 활용하였다. 분석 대상이 되는 비탈면 붕괴 예측 및 분석 관련 문헌들을 추출하기 위해 키워드 ‘slope failure’를 첫 번째 검색어로, 두 번째 키워드(and)로 ‘prediction’과 ‘analysis’를 ‘or’로 입력하여 주제(Topic) 검색을 실시하였다. 검색 결과, 이 기간 동안 발표된 비탈면 붕괴 분석 및 예측에 관한 순수 연구 논문 380건과 미리보기 논문 4건, 편집자료 3건, 프로시딩 논문 3건, 리뷰 6건으로 총 396건이 수집되었다. 이를 공개 소프트웨어인 CiteSpace 5.7을 통해 수집된 기초 데이터에서 의미 있는 단어들을 추출하기 위해 단어 정제 과정인 전처리(preprocess) 분석을 수행하였다. 기초 데이터 수집과 전처리 과정을 통해 최종적으로 분석에 필요한 데이터 셋을 제작하였다. 이후 COOC ver 0.4 프로그램(Lee, 2006a)을 통해 동시출현단어를 추출하고 이를 코사인과 피어슨 상관계수를 활용하여 단어별 유사도를 구하여 단어 간의 관계를 파악하였다. 병렬 최근접 이웃 클러스터링 알고리즘(PNNC)을 활용하여 WNET ver 0.4(Lee, 2013)를 통해 중심성 분석을 통한 네트워크 분석을 진행하였고, 이를 Nodexl(Hansen et al., 2011)을 통해 시각화 하였다. 마지막으로 SPSS ver. 26을 통해 군집분석을 하였고, 비탈면 붕괴 예측 및 분석 연구 영역에서의 흐름을 확인하였다.

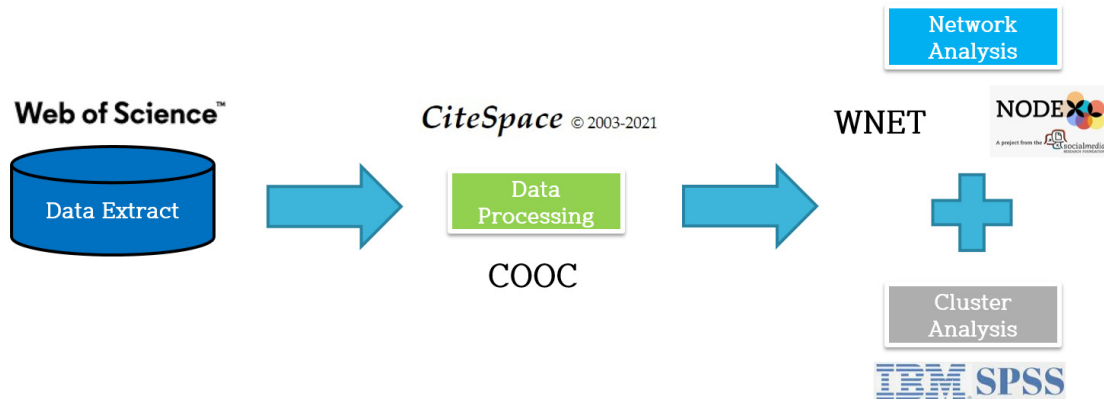


Fig. 1. Cadastral structure analysis process in the field of prediction and analysis of slope failure.

키워드 선정

단어 추출을 위해 본 연구에서는 수집된 396건의 문헌들을 계량분석이 가능한 소프트웨어인 CiteSpace 5.7을 활용하여 제목과 초록에서 키워드와 키워드의 빈도수를 추출하였다. 키워드 ‘slope failure’는 명사구이기 때문에 명사구의 추출을 허용하였으며, 상세한 주제 표현을 위해서 하나의 명사구당 명사 범위 개수를 2개에서 4개 사이로 지정하였다. 초록이 없는 문헌에 대해서는 제목에서만 명사구를 추출하였다. 이 396건의 문헌들에서 추출된 명사구는 총 2,678개(문헌 당 4.62개)이었다. 이 중 CiteSpace를 이용, 분석의 용이성을 위해 중복단어 나 복수의 단수화 등 불필요한 불용어를 제거하여 자주 출현한 명사구 311개의 리스트를 추출하였다. 예를 들어 ‘slope stability’와 ‘stability’의 중복단어는 ‘slope stability’로, ‘slope’와 ‘slopes’는 단수형태인 ‘slope’ 등으로 보다 많은 빈도수의 대표 단어로 변경하였다. 단, ‘rainfall’, ‘rainfall infiltration’과 ‘rock slope’, ‘rock slope failure’ 등과 같이 명사 또는 명사구가 의미하는 바가 다른 경우는 모두 포함시켰다.

이와 같은 키워드 정규화 과정 후에 복잡성을 고려하여 빈도수가 낮은 다수의 키워드를 제외시키기 위하여 상위 약 80% 이상을 차지하는 빈도수 7회 이상의 키워드 89개의 리스트를 완성하였다. 이 리스트는 하나의 문헌에서 중복 출현된 단어의 빈도수가 모두 포함하여 합산되어 있다. 본 연구의 분석에서 필요한 키워드는 각 문헌에서 한 번씩 출현한 횟수이기

때문에 단어 빈도와 문헌 빈도의 값의 치우침에 따른 키워드 선정을 방지하기 위하여 89개의 리스트에서 어떤 단어가 특정 문서 내에서 얼마나 중요한 것인지를 나타내는 통계적 수치인 단어 빈도에 대한 문헌 빈도 값을 나타내는 TF/IDF를 사용하였다. 이 TF/IDF 값이 전체에서 상위 53%를 차지하는 22.77 이상인 키워드를 필터링하여 최종적으로 키워드 리스트 총 36개를 재선정하였다. 추출된 단어 리스트에서 해당 분야와 관련이 없는 ‘American society’, ‘good agreement’ 등 단어는 제외시켰다. 이렇게 완성된 분석 대상이 되는 최종 키워드 36개 리스트는 Table 1과 같다.

비탈면 붕괴 예측 및 분석을 대표로 하는 키워드 slope failure(1), slope stability(2), stability analysis(3), debris flow (4), rainfall(5)은 빈도수가 매우 높아 현재까지 매우 중요하게 연구되고 있다. 또한 numerical analysis(10), numerical simulation(18), safety factor(24), numerical modeling(29), failure probability(31)의 키워드를 통해 수학적인 증명과 시뮬레이션을 통해 데이터 분석이 이루어지고 있음을 알 수 있다.

Table 1. List of final 36 keywords with frequency of 7 or more

No.	Keyword	Frequency	No.	Keyword	Frequency
1	slope failure	267	19	slip surface	22
2	slope stability	132	20	reliability analysis	20
3	stability analysis	111	21	shallow landslide	19
4	debris flow	52	22	failure mode	16
5	rainfall	50	23	soil slope	16
6	failure mechanism	46	24	safety factor	15
7	shear strength	39	25	unsaturated soil	15
8	limit equilibrium method	35	26	landslide susceptibility	14
9	slope instability	35	27	rainfall infiltration	13
10	numerical analysis	27	28	failure surface	13
11	rock slope	27	29	numerical modeling	10
12	early warning	26	30	kinematic analysis	10
13	risk assessment	24	31	failure probability	8
14	hydraulic conductivity	24	32	mass movement	8
15	finite element method	24	33	case study	8
16	spatial variability	23	34	rock slope failure	8
17	progressive failure	23	35	pore water pressure	8
18	numerical simulation	22	36	rainfall-induced landslide	7

동시출현단어 행렬 작성

본 연구는 키워드 간 연관성을 측정하는 방법 중, 키워드 간 직접적인 연관성을 측정하고 네트워크 분석에 유용한 코사인 유사도와 키워드 간 간접적인 연관성을 측정하고 주로 시각화에 사용하는 피어슨 상관계수를 활용하여 연관성 분석에 활용하였다. Fig. 1과 같이 36개의 키워드가 선정된 문헌 396건의 정보를 엑셀에 모두 입력하여 COOC ver 0.4를 이용하여 동시출현단어 행렬을 작성하였다. CiteSpace를 통해 제목과 초록에서 추출된 키워드를 입력하고 [문헌번호, 키워드] 형태의 쌍을 만들어 입력하여 키워드 전처리 작업을 끝낸 후, 총 1,146행의 최종 분석 대상 출현 단어 리스트를 완성하였다. 이를 COOC를 통해 최종 키워드 리스트, 키워드 출현빈도 리스트, 동시출현단어 행렬을 생성한다. 이 동시출현단어 행렬은 정방대칭행렬이며, 36×36 의 동시출현빈도와 대각선 칸의 출현빈도로 구성된다. 이 동시출현단어 행렬을 다시

COOC를 통해 코사인 유사도계수로 정규화한 행렬을 획득한다. 이 코사인 유사도 행렬의 값은 1에 가까울수록 두 단어는 유사도가 높고, 0에 가까울수록 두 단어의 유사도는 낮음을 나타낸다. 유사도가 높은 키워드 쌍은 주제 분야가 높다는 것을 의미한다. 이 연관성 행렬을 다시 COOC ver 0.4에 입력하면 마지막으로 피어슨 상관계수에 의한 2차 연관성 행렬이 추출된다. 이와 같은 2차 연관성 행렬을 이용(White and Griffith, 1981)하여 두 키워드와 제3의 키워드 간의 동시출현 패턴의 유사함을 측정할 수 있다. 이 피어슨 상관계수의 값은 관계의 크기와 방향의 동시에 나타내며 -1에서 1사이를 가진다. 관계의 크기는 절대값으로 표현하며, 이 절대값이 클수록 두 키워드가 사이의 관계성이 높음을 나타내며, 작을수록 관계성이 낮음을 나타낸다. 또한 +1에 가까울수록 양의 상관관계, -1에 가까울수록 음의 상관관계를 나타내며, 0은 두 키워드 간의 관련성이 없음을 뜻한다.

분석결과, 1차 연관성 행렬인 코사인 유사도 행렬에서 유사도가 가장 높은 키워드 쌍은 ‘spatial variability’-‘risk assessment’(0.41039)로 나타났으며, 그 다음으로 ‘spatial variability’-‘reliability analysis’(0.40415), ‘soil slope’-‘rainfall infiltration’(0.37796) 순으로 나타났다. 유사도가 0인 키워드 쌍을 제외하고, 유사도가 가장 낮은 것으로 나타난 키워드 쌍은 ‘stability analysis’-‘rainfall’(0.02579)로 나타났으며, 그 다음으로 ‘slope stability’-‘hydraulic conductivity’(0.03066), ‘shear strength’-‘rainfall’(0.03341) 순으로 매우 낮은 유사도를 보여주었다.

피어슨 상관계수의 정도에 대하여 상관계수의 절댓값이 0.9~1.0은 아주 높은 상관관계, 0.7~0.9는 높은 상관관계, 0.4~0.7은 비교적 높은 상관관계, 0.2~0.4는 낮은 상관관계, 0.2 이하는 매우 경미한 상관관계를 나타낸다. 이 피어슨 상관계수의 값을 분석한 결과, 상관관계가 가장 높은 키워드 쌍은 코사인 유사도의 두 번째로 높은 키워드 쌍이었던 ‘spatial variability’-‘reliability analysis’(0.72404)이며, 그 다음으로는 ‘shear strength’-‘safety factor’(0.70591), ‘spatial variability’-‘risk assessment’(0.70086), ‘risk assessment’-‘reliability analysis’(0.68888) 순으로 나타났으며, 양의 높은 상관관계를 갖는다. 음의 상관관계가 가장 큰 키워드 쌍은 ‘rainfall infiltration’-‘numerical modeling’(-0.16988), ‘safety factor’-‘numerical modeling’(-0.16903), ‘numerical modeling’-‘hydraulic conductivity’(-0.08472), ‘spatial variability’-‘finite element method’(-0.15961) 순으로 나타났다. 이들은 음의 경미한 상관관계를 가지고 있다. 이 값들을 제외하고 음의 상관관계를 지닌 키워드 쌍은 총 343개였다.

지적구조 분석결과

분석을 위한 네트워크 생성

네트워크 관계에 따른 여러 군집의 주제영역으로 구분하기 위해 동시출현단어 분석을 사용하며 네트워크를 구성하고 이를 시각화 한다. 본 연구에서는 비탈면 붕괴 예측 및 분석 연구의 지적구조를 분석하기 위해 피어슨 상관계수를 이용하여 패스 파인더 네트워크 알고리즘(Lee, 2006a)을 적용하여 네트워크를 생성하였다. 그리고 패스파인더 네트워크상에서 주제와 세부주제들을 명확하게 하기위해 병렬 최근접 이웃 클러스터링 알고리즘(PNNC)(Lee, 2006b)을 이용하였다. 이를 위하여 네트워크 생성과 클러스터링 알고리즘을 이용하기 위해 WNET을 활용하였고, 시각화를 위한 NodeXL을 사용하였다. 키워드 36개에 관한 동시출현단어의 네트워크 분석 결과, PNNC 알고리즘으로 9개의 최적의 군집이 생성되었다. Fig. 2는 9개의 군집의 각 세부 영역을 병합하여 나타내었으며, 각 군집의 번호와 각 군집에서 빈도수가 가장 높은 키워드인 slope failure(1), slope stability(2), stability analysis(3), shear strength(5), progressive failure(17), hydraulic conductivity(14), shallow landslide(21), unsaturated soil(25)이 대표 키워드로 해당 군집의 주제명으로 부여하였으며, 노드의 빈도수를 노드의 크기로 표현하였다.

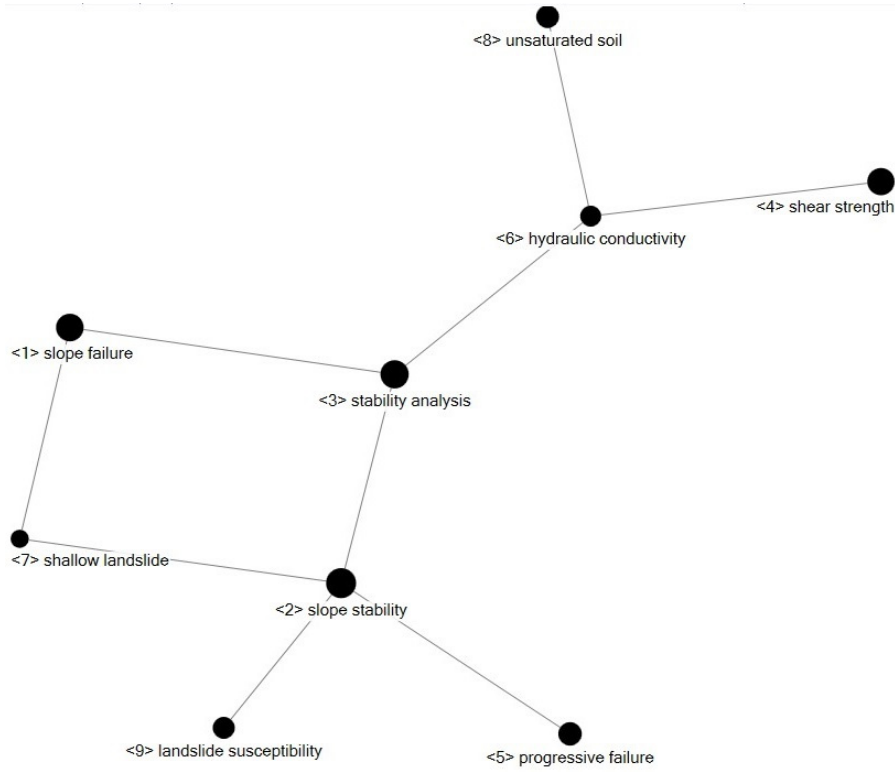


Fig. 2. 9 clusters appearing in PFNET.

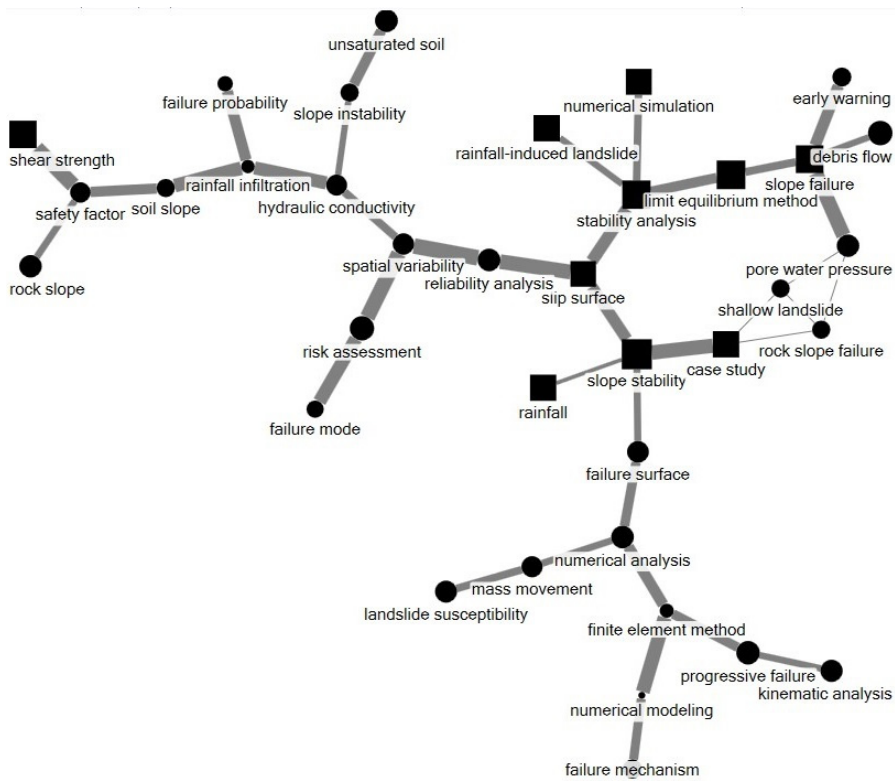


Fig. 3. Keyword relations network by closeness centrality.

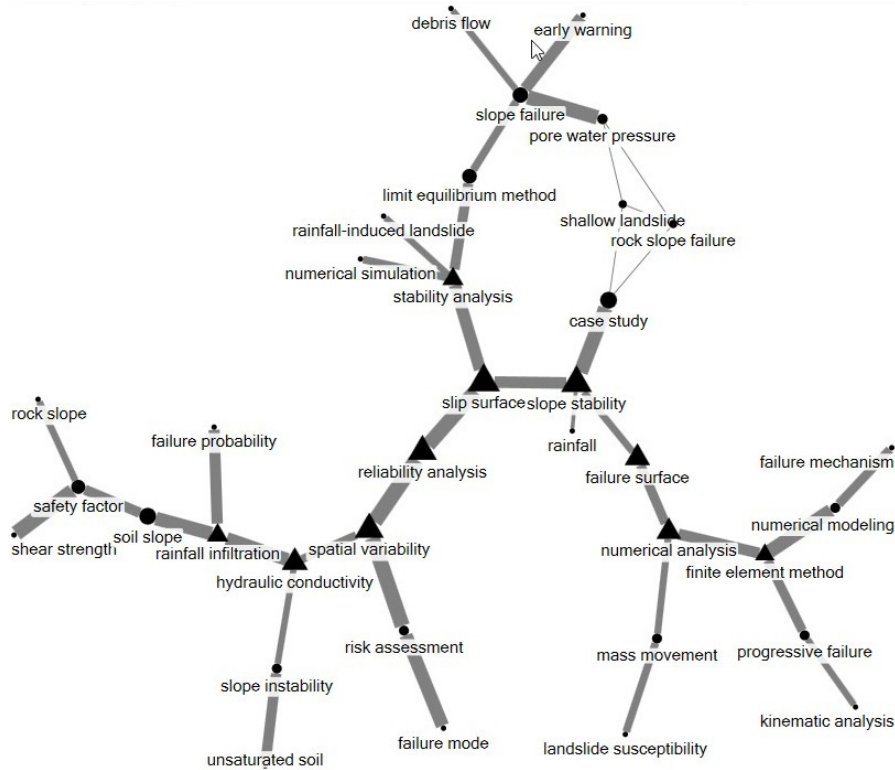


Fig. 4. Keyword relationship network by degree and betweenness centrality.

Fig. 2의 병합을 해제하여, Fig. 3, Fig. 4와 같이 각 9개의 군집에 속한 하위 주제영역을 나타내고 네트워크 중심성 분석을 활용하여 전역중심성, 지역중심성, 매개중심성이 높은 키워드를 확인하였다. 전역중심성이 높은 키워드는 비탈면 붕괴 예측 및 분석 영역에 전반적으로 타 키워드들과 다양하게 연계되어 있는 키워드이고, 지역중심성이 높은 키워드는 군집 내의 세부 영역에서 영향력이 있는 키워드이며, 매개중심성은 각 노드들을 연결해주는 중개자 역할을 한다.

네트워크 분석에 의한 지적구조 분석

Fig. 2에 최근접 이웃 클러스터링 알고리즘을 통해 구성된 네트워크는 빈도수를 통해 얼마나 높은 연관성을 가지고 연결이 되었는지에 대한 연결 강도를 갖는 가중 네트워크이므로 전체 네트워크에서 키워드들이 구조적으로 어떻게 연결되었는지에 대해 중심성 분석을 하기 위해 가중 네트워크에 활용되는 중심성 분석 척도(Lee, 2006c)를 활용하였다. 전역중심성을 높은 키워드를 측정하기 위해 중심성 분석 척도 중에서 상대적 삼각매개중심성(Relative triangle betweenness centrality)과 평균연관성(Average mean association)을 활용하였으며, 이를 반영하여 Fig. 3과 같이 네트워크를 작성하였다.

중심성은 네트워크를 구성하는 동시출현단어의 강도, 영향력 범위, 구조적인 입지를 나타낸다. 상대적 삼각매개중심성은 이 동시출현단어의 영향력 범위를 많이 반영하고 구조적인 입지를 적게 반영한 값이며 평균연관성의 경우 동시출현단어의 강도를 반영하므로 동시출현단어에 따른 중심성 분석에 활용되었다. 노드 간의 연결된 간선에 키워드 빈도에 따른 가중치에 따라 굵기를 표현하였으며, 상대적 삼각매개중심성 값은 노드의 크기로 표현하였다. 또한 평균연관성을 확인하기 위해 중심성 값이 0.04 이상 상위 10개 노드들을 모두 사각형으로 바꾸어 나타내었다. 이 상위 10위까지의 키워드를 활용하여 상대적 삼각매개중심성지수와 평균연관성 지수를 비교하였으며, 이를 통해 비탈면 붕괴 예측 및 분석에 관한 전반

에 걸친 중심이 되는 키워드를 구별하여 확인할 수 있었다. Table 2와 같이 상대적 삼각매개중심성과 평균연관성을 비교하였을 때, ‘slope stability(2)’, ‘limit equilibrium method(3)’, ‘stability analysis(3)’, ‘slope failure(1)’, ‘shear strength(4)’, ‘slip surface(3)’는 모두 10위 안에 포함되는 높은 전역중심성을 가진 키워드라고 할 수 있으며, 이는 비탈면 붕괴 예측 및 분석에 광범위하게 활용되는 키워드라고 할 수 있다. 또한 상대적 삼각매개중심성에서의 ‘rainfall-induced landslide(3)’, ‘case study(2)’, ‘rainfall(5)’, ‘numerical simulation(3)’, 평균연관성에서의 ‘spatial variability(3)’, ‘risk assessment(13)’, ‘hydraulic conductivity(14)’, ‘reliability analysis(20)’, 키워드 역시도 비탈면 붕괴 예측 및 분석 연구에 자주 활용되는 키워드라고 볼 수 있다.

Table 2. Top 10 in two global centrality indices for weighted networks

Rank	Keyword (No.)	Relative triangle betweenness centrality (rTBC, 0~1)	Keyword (No.)	Average mean association (AVGSIM)
1	slope stability (2)	0.59664	slope stability (2)	0.16171
2	limit equilibrium method (8)	0.54118	stability analysis (3)	0.14917
3	stability analysis (3)	0.52773	shear strength (5)	0.12224
4	slope failure (1)	0.5042	spatial variability (16)	0.11149
5	shear strength (5)	0.49244	reliability analysis (20)	0.11084
6	rainfall-induced landslide (36)	0.45882	limit equilibrium method (8)	0.11065
7	case study (33)	0.45714	risk assessment (13)	0.10044
8	rainfall (7)	0.4521	slip surface (22)	0.09965
9	numerical simulation (18)	0.44202	slope failure (1)	0.09918
10	slip surface (19)	0.42857	hydraulic conductivity (14)	0.08254

다음으로 9개 군집의 세부 주제영역을 분석하기 위해 지역중심성 지수를 측정하고, 세부 키워드들을 이해하기 위해 군집에서 영향력이 가장 높은 키워드를 확인할 수 있는 상대적 최근접이웃중심성값을 이용하였다. 각 군집에서 지역중심성이 가장 높은 중심 주제어는 제1군집 slope failure(1), 제2군집 slope stability(2), 제3군집 stability analysis(3), 제4군집 shear strength(5), 제5군집 progressive failure(17), 제6군집 hydraulic conductivity(14), 제7군집 shallow landslide(21), 제8군집 unsaturated soil(25), 제9군집 landslide susceptibility(26)이었으며, 제7, 8, 9군집은 군집에 속한 2개에 키워드로만 구성되었다.

또한 비탈면 붕괴 예측 및 분석 영역에서 군집 간 중개자 역할을 하는 키워드들을 확인하기 위하여 NodeXL을 이용하여 매개중심성 지수를 측정하였다. 이 매개중심성 값 상위 11위까지를 기준으로 slip surface(19), slope stability(2), spatial variability(16), reliability analysis(20), hydraulic conductivity(14), failure surface(28), numerical analysis(10), rainfall infiltration(27), stability analysis(3), finite element method(15) 순이었다. Fig. 4와 같이 노드의 크기는 지역중심성에 따라, 간선의 크기는 키워드 간 빈도에 비례하도록 하여, 매개중심성 값이 3.65 이상인 상위 10개의 노드들을 삼각형 형태로 표시하여 네트워크 지도를 작성하였으며, 이를 통해 비탈면 붕괴 예측 및 분석에 관한 키워드 간 주요 매개역할을 하는 키워드를 구별하여 확인할 수 있었다.

제1군집은 비탈면 붕괴(slope failure)와 가장 관련이 깊은 키워드로 형성되었다. 이 군집에 속해 있는 키워드들은 비탈면 붕괴와의 유사 명칭, 원인, 대응과 관련한 연구 영역으로 세부 주제를 요약할 수 있다. slope failure(1), debris flow(6),

early warning(12), pore water pressure(35)로 구성되어 있으며, debris flow(6)은 일반적인 비탈면 붕괴와는 달리 활동면을 가지지 않는 토석류를 나타내며, early warning(6)은 모니터링에 기반한 조기경보시스템, pore water pressure(35)은 간극수압을 나타낸다. 각 키워드의 빈도수는 매우 많은 키워드부터 낮은 키워드까지 다양하며, 비탈면 붕괴의 원인과 대응에 대한 연구가 활발히 진행 중임을 알 수 있다.

제2군집의 키워드들은 비탈면 안정성(slope stability)과 그와 관련된 세부 주제들이며, 제3군집, 제5군집, 제7군집과 연결되어 있어 광범위하게 연구가 진행 중이다. slope stability(2), rainfall(7), failure surface(28)로 구성되어 있으며, 비탈면 안정성과 관련하여 rainfall(7)은 강우의 영향과 그 특성을, failure surface(28)는 붕괴 활동면의 형태와 그 특성을 나타낸다. 비탈면 붕괴의 원인과 이로 인해 영향을 받는 비탈면 활동면에 관한 연구가 이루어지고 있음을 알 수 있다.

제3군집은 안정성 분석(stability analysis)과 관련된 키워드로 구성되어 있으며, 특히 본 연구인 비탈면 붕괴 예측 및 분석과 관련해 가장 많은 키워드와 핵심 키워드로 이루어져 있으며, 제1군집, 제2군집, 제6군집과 연결되어 있어 구조적으로도 매우 중요한 군집이라고 할 수 있다. 또한 이 군집은 stability analysis(3), spatial variability(16), risk assessment(13), reliability analysis(20), numerical simulation(18), limit equilibrium method(8), rainfall-induced landslide(36), failure mode(22), slip surface(19)가 포함되며, 안정성 분석 방법과 안정성 분석과 관련한 활동면(또는 파괴면) 연구로 이루어져 있다. spatial variability(16)은 안정성 분석과 관련된 공간적 분포 변화, risk assessment(13)은 위험성 평가, reliability analysis(20)는 신뢰성 분석, numerical simulation(18)은 수치해석, limit equilibrium method(8)은 한계평형해석, rainfall-induced landslide(36)는 강우로 인한 산사태, failure mode(22)는 파괴유형, slip surface(19)는 활동면을 나타낸다. 각 키워드들은 상대적 삼각매개중심성, 평균연관성 모두 높은 값들을 보유하고 있으며, 높은 유사도를 가진 키워드로 구성되어 비탈면 붕괴 예측 및 분석과 관련하여 가장 폭넓게 연구 중인 것을 알 수 있다. 제2군집과 제3군집은 다소 유사한 측면이 있으나, 본 연구의 연구방법을 통해 연관되는 주요 용어들이 다소 상이한 측면이 있어 이를 구분하였다.

제4군집은 전단강도(shear strength)와 관련된 키워드들로 구성되어 있다. shear strength(5), rock slope(11), safety factor(24)로 이루어져 있으며, rock slope(11)은 암반 내 불연속면의 전단강도에 가장 영향을 크게 받는 암반 비탈면을, 안정성 해석결과로부터 도출된 안전율인 safety factor(24)을 나타낸다.

제5군집도 비탈면의 붕괴 중 상태인 진행성 붕괴(progressive failure)와 관련된 키워드와 그와 관련된 세부 주제들이다. progressive failure(17), failure mechanism(4), numerical analysis(10), finite element method(15), failure surface(28)로 구성되어 있으며, 제3군집 다음으로 많은 키워드로 구성되는 것으로 보아, 붕괴 가능성과 더불어 붕괴가 앞으로 어떻게 진행될 것인가에 대한 연구도 활발히 이루어지고 있음을 확인할 수 있었다. failure mechanism(4)은 붕괴에 대한 메커니즘, numerical analysis(10)의 경우 이러한 진행성 붕괴 예측 및 분석과 관련해 수치해석을 통해 이를 보다 과학적으로 규명하기 위한 연구가 진행되고 있음을 알 수 있으며, finite element method(15)는 비탈면 구성요소의 거동을 설명하기 위한 유한요소법에 대한 연구도 진행되고 있고, 실제로 확인할 수 있는 파괴면(failure surface(28))에 관한 연구 역시 진행성 붕괴와 관련하여 이루어지고 있음을 확인할 수 있었다.

제6군집은 토양과 같은 다공성 매질의 투수성을 나타내는 수리전도도(hydraulic conductivity)와 이와 관련된 키워드들로 구성되어 있으며, 제3군집, 제4군집, 제8군집과 연결되어 있다. hydraulic conductivity(14), soil slope(23), rainfall infiltration(27), failure probability(31)가 군집을 이루고 있으며, soil slope(23)는 수리전도도와 연관이 깊은 토양 비탈면을, rainfall infiltration(27)은 강우에 의한 침투 영향, failure probability(31)은 붕괴확률이며, 주제들이 물과 영향이 깊기 때문에 수리전도도 기반으로 구성된 본 군집으로 묶여있는 연관성을 확인할 수 있었다.

제7군집은 얇은 산사태(shallow landslide)와 관련된 키워드들로 구성되어 있으며, shallow landslide(21)과 rock slope

failure(34)로 2개의 키워드로만 구성되어 있지만 비탈면 붕괴의 핵심 군집인 제1군집과 제2군집과 연결로 인해 매우 중요한 군집 중 하나라고 할 수 있다. rock slope failure(34)는 암반 비탈면 붕괴로 붕괴의 얕은 깊이 특성상 얕은 산사태로 같이 연구가 되고 있다고 확인할 수 있다.

제8군집은 불포화토(unsaturated soil)와 관련된 unsaturated soil(25), slope instability(9)로 이루어져 있으며, slope instability(9)는 비탈면 불안정성을 나타낸다.

마지막으로 제9군집은 산사태 민감성(landslide susceptibility)과 관련된 키워드인 landslide susceptibility(26)과 mass movement(32)로 이루어져 있다. mass movement(32)는 중력에 의해 일어나는 지표면의 현상으로 지진, 폭우, 해빙으로 인해 토양의 내부마찰이 감소하여 일어나는데 이와 관련된 산사태 민감성이 연구되고 있는 것으로 확인된다.

군집분석에 의한 지적구조

네트워크 분석을 활용한 군집 분석을 보완하기 위하여 통계프로그램인 SPSS ver 26.0을 활용해 피어슨 상관계수를 활용한 2차 연관성 행렬로부터 군집분석을 하였다. 이를 위해 Ward 기법을 활용한 클러스터링, z점수 표준화 설정 및 제곱 유클리디안 거리를 이용하여 덴드로그램으로 산출하였다. 본 연구에서 네트워크 분석과 군집 분류를 통해 비탈면 붕괴 예측 및 분석에 관한 연구 경향을 가장 잘 나타내는 군집의 수는 2개라고 판단하였고 군집명의 경우 군집별 개별 키워드를 종합하는 가장 적절한 이를 통한 Table 3은 SPSS를 통해 도출된 군집의 순서와 관계없이 도출된 군집분류이며 비탈면 붕괴 예측 및 분석 연구 영역에서의 흐름을 확인할 수 있었다.

Table 3. Domain structure by cluster analysis

Name of cluster	Keyword (No.)	Name of cluster	Keyword (No.)
Cluster #1 (16 keywords) slope failure	slope failure (1)	Cluster #2 (20 keywords) stability analysis	slope stability (2)
	failure mechanism (4)		stability analysis (3)
	shear strength (5)		debris flow (6)
	limit equilibrium method (8)		rainfall (7)
	rock slope (11)		slope instability (9)
	hydraulic conductivity (14)		numerical analysis (10)
	progressive failure (17)		early warning (12)
	numerical simulation (18)		risk assessment (13)
	failure mode (22)		finite element method (15)
	unsaturated soil (25)		spatial variability (16)
	landslide susceptibility (26)		slip surface (19)
	rainfall infiltration (27)		reliability analysis (20)
	kinematic analysis (30)		shallow landslide (21)
	mass movement (32)		soil slope (23)
	rock slope failure (34)		safety factor (24)
	pore water pressure (35)		failure surface (28)
	numerical modeling (29)		
	failure probability (31)		
	case study (33)		
	rainfall-induced landslide (36)		

제1군집 slope failure는 비탈면 붕괴와 관련된 키워드들로 구성되어 있다. 비탈면 붕괴 원인 분석을 위한 failure mechanism(4), rock slope(11), progressive failure(17), failure mode(22), rock slope failure(34), 원인 요소인 shear strength(5),

hydraulic conductivity(14), unsaturated soil(25), landslide susceptibility(26), rainfall infiltration(27), mass movement (32), pore water pressure(35), 분석방법인 limit equilibrium method(8), numerical simulation(18), kinematic analysis(30)를 포함한다.

제2군집은 stability analysis(3)가 대표 키워드로서, 안정성 분석과 관련된 키워드들로 구성되어 있다. 안정성과 그 활동면을 나타내는 slope stability(2), slope instability(9), rainfall(7), spatial variability(16), slip surface(19), failure surface (28), 분석을 위한 방법인 numerical analysis(10), risk assessment(13), finite element method(15), reliability analysis(20), safety factor(24), numerical modeling(29), failure probability(31), case study(33), 형태 및 붕괴 종류와 관련된 debris flow(6), shallow landslide(21), soil slope(23), rainfall-induced landslide(36), 조기 알람인 early warning(12)을 포함한다.

이러한 실질적인 군집분석을 통하여 비탈면 붕괴 예측 및 분석 영역에 걸쳐 비탈면 붕괴와 안정성 분석이 핵심적인 연구영역이며, 비탈면 붕괴와 관련하여 원인요소, 분석방법에 관한 연구를 수행하고 있으며, 안정성 분석을 위해 안정성과 활동면, 분석 방법, 형태 및 붕괴 종류에 관하여 연구가 이루어지고 있음을 알 수 있다.

다차원축적지도에 의한 지적구조

군집분석으로 작성된 다차원축적지도에 의한 키워드들의 지적구조를 분석하기 위해 SPSS ver 26.0로부터 PROXSCAL 알고리즘을 이용하였다. 변수는 z점수로 표준화 하였고 스트레스 값은 0.001로 설정하였다. 이를 Fig. 5와 같이 키워드지도상에 군집분석의 결과를 2개의 클러스터영역의 경계를 점선으로 표시하고 군집명을 표기하였으며 상관관계에 따라 키워드의 위치를 확인할 수 있다.

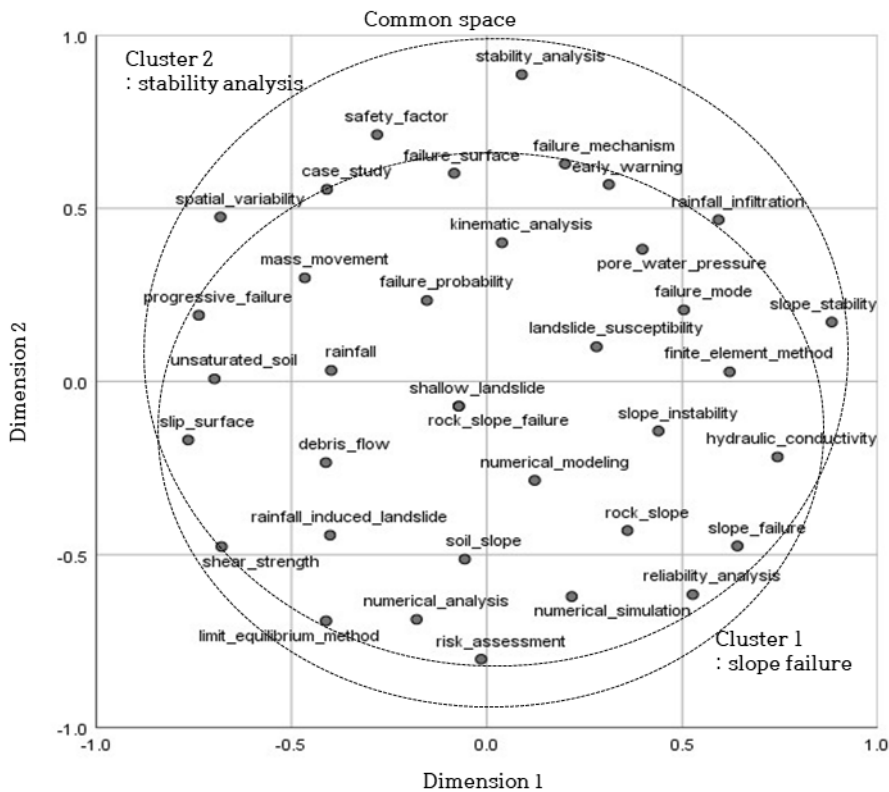


Fig. 5. Domain structure by multidimensional accumulation map.

지도의 y축을 기준으로 하단에 위치한 키워드들은 제1군집인 slope failure의 세부 키워드들과 유사하게 분포되어 있는 것을 확인할 수 있다. 그리고 상단에 위치한 키워드들은 제2군집인 stability analysis의 세부 키워드들과 유사하게 나타났다. 제1군집 slope failure와 제2군집 stability analysis의 약 80%는 교집합으로 겹치는 것을 확인할 수 있으며, 이는 비탈면 붕괴 예측 및 분석 연구영역에 있어서 slope failure와 stability analysis와 이와 관련된 키워드들 모두 상관관계가 높으며, 중심축에 위치하고 있음을 다차원축적지도에 의한 지적구조 분석을 통해 보여주는 결과이다. 다시 말하면, slope failure와 slope analysis와 관련한 모든 키워드들은 비탈면 붕괴 분석 및 예측 연구에서 중심이 되어 사용되고 있음을 직관적으로 알 수 있다.

결론 및 토론

본 연구는 최근 5년간 Web of Science에서 비탈면 붕괴 예측 및 분석을 주제로 하는 데이터들을 수집하고 동시출현단어 행렬을 통한 네트워크 분석, 군집 분석, 다차원축적지도를 통해 비탈면 붕괴 예측 및 분석에 관한 연구 경향을 반영하는 지적구조를 분석하였다. 네트워크 분석을 위해 유사도 및 피어슨 상관계수를 통한 연관성 행렬을 생성하고 이들의 관계를 패스파인더 네트워크를 통한 시각화와 최근접 이웃 클러스터링 군집을 통해 비탈면 붕괴 예측 및 분석의 세부 주제영역을 분석할 수 있었다.

전역 및 지역 중심성 분석을 통해 비탈면 붕괴 예측 및 분석에서 가장 핵심이 되는 키워드와 군집별로 영향력이 있거나 매개가 되는 키워드들을 파악하였다. 네트워크 분석을 보완하기 위한 군집분석을 실시하였고 이 군집분석의 결과를 다차원축적지도에 반영하여 이를 통해 비탈면 붕괴 예측 및 분석 연구의 지적구조와 세부 주제영역을 분석하였다. 네트워크 분석의 결과, 9개의 군집이, 군집분석의 결과로 2개의 군집이 추출되었다. 네트워크 분석을 통해, slope failure, slope stability, stability analysis, shear strength, progressive failure, hydraulic conductivity, shallow landslide, unsaturated soil, landslide susceptibility가 대표 키워드로서 군집이 생성됨으로써 이와 관련된 세부 주제로 연구가 진행되고 있음을 알 수 있었다. 또한 군집분석 결과를 다차원축적지도에 표시한 결과 제1군집, 제2군집이 매우 많이 교집합을 이루는 결과를 보여주었다. 이것은 현재의 연구결과 키워드 중심 분석을 통해 실질적으로도 제1군집의 비탈면 붕괴-slope failure와 제2군집의 안정성 분석-stability analysis의 상관관계가 높으며, 비탈면 붕괴 예측 및 분석 연구의 중심이 되는 키워드들로 연구되고 있음을 보여주었다. 전역중심성이 높은 키워드는 slope stability, limit equilibrium method, stability analysis, shear strength 순으로 나타났다. 지역중심성이 높은 키워드는 slip surface, slope stability, spatial variability, reliability analysis, hydraulic conductivity 순으로 나타났다. 이러한 결과는 비탈면 붕괴의 예측 및 분석 연구에서 붕괴원인, 안정성, 안정성 분석 방법, 분석방법에 포함되는 요소들(특히 전단강도)과 비탈면 활동면에 대한 많은 연구가 이루어지고 있음을 알 수 있었다. 이상의 결과를 종합해 보면 데이터를 추출했던 2016년부터 2020년까지의 다학제 지구과학, 지질공학, 지질학 범주에서 비탈면 붕괴 예측 및 분석의 핵심적인 연구 영역은 비탈면 붕괴와 안정성 분석 영역이다.

본 연구의 목적은 최근 비탈면 붕괴 예측 및 분석 관련 문헌들의 지적구조를 분석함으로써 현재 주요하게 다루어지고 있는 그 주제와 상호 관련성을 규명하는데 있다. 문헌의 특성상, 비탈면 붕괴와 그 안정성 분석이라는 것이 관련 문헌의 핵심주제가 되고, 이를 밝히고 과학적 증거로 뒷받침하기 위한 관련 키워드들이 함께 다루어지는 것은 당연한 결과라고 치부할 수 있다. 그럼에도 불구하고 비탈면에 대한 연구를 처음 시작하는 학생들이나 드론, 빅데이터, 인공지능 등 최신기술을 다루면서도 비탈면에 대한 지식이 없는 전문가들 입장에서는 본 논문이 제시하는 연관 키워드나 그 내용이 도움이 될 수 있을 것이라고 판단한다.

본 연구는 비탈면 붕괴 예측 및 분석 연구의 동향을 파악하는 연구였기 때문에 수집 범주를 다학제 지구과학, 지질공학, 지질학으로 한정하여 결과를 도출한 것이므로 수리학, 지형학 등에서 다루고 있는 비탈면 붕괴의 분석과 예측에 관한 주제와 관련지을 수 없는 제한적 측면을 포함하고 있기는 하다. 그러나 본 연구는 비탈면 붕괴 예측 및 분석 연구의 지적구조를 분석할 수 있었다는 것에 의의가 있으며, 향후 최근 인기있는 미래 기술을 활용하여 비탈면 붕괴 예측 및 분석 연구를 접목하고자 하는 연구자에게는 도움이 될 수 있을 것으로 기대한다.

사사

본 연구는 2020년도 정부(과학기술정보통신부)의 재원으로 한국연구재단의 지원을 받아 수행된 기초연구사업입니다 (2020R1F1A107471).

References

- Hansen, D., Shneiderman, B., Smith, M.A., 2011, Analyzing social media networks with NodeXL: Insights from a connected world, Morgan Kaufmann, 283p.
- Kim, S.H., Koo, H.B., Hwang, J.H., Son, M., 2011, Case study on the cause of failure and characteristics of soil at a collapsed cut-slope at the ○○ Detour, Jeonranam-Do, The Journal of Engineering Geology, 21(4), 313-322 (in Korean with English abstract).
- Lee, J.Y., 2006a, A study on the network generation methods for examining the intellectual structure of knowledge domains, Journal of the Korean Society for Library and Information Science, 40(2), 333-355 (in Korean with English abstract).
- Lee, J.Y., 2006b, A novel clustering method for examining and analyzing the intellectual structure of a scholarly field, Journal of the Korean Society for Information Management, 23(4), 215-231 (in Korean with English abstract).
- Lee, J.Y., 2006c, Centrality measures for bibliometric network analysis, Journal of the Korean Society for Library and Information Science, 40(3), 191-214 (in Korean with English abstract).
- Lee, J.Y., 2013, A comparison study on the weighted network centrality measures of tnet and WNET, Journal of the Korean Society for Information Management, 30(4), 241-264 (in Korean with English abstract).
- Lee, K.B., Shin, H.S., Kim, S.H., Ha, D.M., Choi, I., 2019, A study on automatic classification of characterized ground regions on slopes by a deep learning based image segmentation, Tunnel and Underground Space, 29(6), 508-522 (in Korean with English abstract).
- MLTM (Ministry of Land, Transport and Maritime Affairs), 2011, Roadside slope maintenance manual, 495p.
- Park, D.G., Kim, T.H., Park, J.H., 2006, Status and countermeasures of slope damage in Korea, Geotechnical Engineering, 22(6), 6-18.
- Seo, S.K., Chung, E.K., 2013, Domain analysis on the field of open access by co-word analysis, Journal of the Korean BIBLIA Society for Library and Information Science, 24(1), 207-228 (in Korean with English abstract).
- White, H.D., Griffith, B.C., 1981, Author cocitation: A literature measure of intellectual structure, Journal of the American Society for Information Science, 32(3), 163-171.