



(19) 대한민국특허청(KR)
(12) 공개특허공보(A)

(11) 공개번호 10-2019-0086937
(43) 공개일자 2019년07월24일

(51) 국제특허분류(Int. Cl.)

G06F 16/00 (2019.01)

(52) CPC특허분류

G06F 16/93 (2019.01)

G06F 2216/11 (2013.01)

(21) 출원번호 10-2018-0005085

(22) 출원일자 2018년01월15일

심사청구일자 2018년01월15일

(71) 출원인

청주대학교 산학협력단

충청북도 청주시 청원구 대성로 298 (내덕동)

고려대학교 산학협력단

서울특별시 성북구 안암로 145, 고려대학교 (안암동5가)

(72) 발명자

전성해

인천광역시 연수구 해돋이로120번길 16, 202동 1702호(송도동, 송도풍림아이원2단지아파트)

박상성

서울특별시 동대문구 한천로58길 139, 대우아파트 103-101

(74) 대리인

추혁

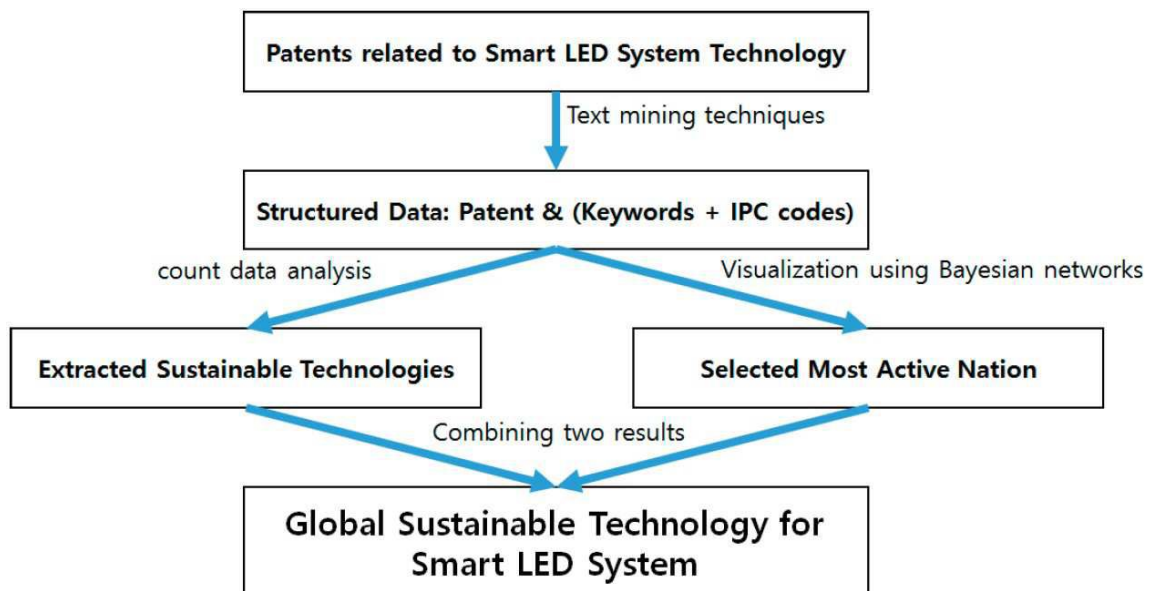
전체 청구항 수 : 총 13 항

(54) 발명의 명칭 베이지안 네트워크를 이용한 기술문서 분석방법

(57) 요약

본 발명은 특허문서의 분석을 이용하여 대상기술의 핵심기술을 분석할 수 있는 베이지안 네트워크를 이용한 기술 문서 분석방법에 관한 것으로서, 베이지안 네트워크를 기반으로 한 시각화와 대상기술과 관련된 특허문서를 분석하는 선형 카운트 모델로 구성되어, 이러한 시각화 및 모델링 결과를 종합하여 대상기술의 지속 가능성을 향상시킬 수 있는 계층적 기술구조를 제공함으로써, 세계적으로 지속 가능한 기술을 찾기 위한 기술경영에서 효율적으로 연구개발을 계획하고, 대상기술 분야에서 가장 활발한 연구개발을 수행하고 있는 국가를 식별하는데 있어서 객관적이고 안정된 결과를 얻을 수 있는 통계적 모델링을 기반으로 하는 신규한 기술분석방법론을 제공한다.

대표도 - 도3



이 발명을 지원한 국가연구개발사업

과제고유번호 2017031152

부처명 교육부

연구관리전문기관 한국연구재단

연구사업명 개인기초연구(교육부)

연구과제명 연구개발 빅데이터를 이용한 미래기술예측 전문가시스템 개발

기 여 율 1/1

주관기관 청주대학교

연구기간 2017.06.01 ~ 2020.05.31

명세서

청구범위

청구항 1

분석하고자 하는 대상기술과 관련된 특허문서를 수집하는 단계;

상기 수집된 특허문서에서 데이터를 추출하고 이를 분석할 수 있도록 구조화된 특허 데이터로 가공하는 단계;

상기 구조화된 특허 데이터에 대하여 빈도회귀분석을 실시하는 단계;

상기 구조화된 특허 데이터에서 베이지안 네트워크를 기반으로 한 시각화를 수행하는 단계; 및

상기 빈도회귀분석을 실시한 결과와 상기 수행된 베이지안 네트워크의 시각화 정보를 결합하는 단계;를 포함하는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 2

제 1항에 있어서, 상기 수집된 특허문서를 분석할 수 있도록 구조화된 특허 데이터로 가공하는 단계는,

텍스트 마이닝 기법을 통해 상기 특허문서로부터 키워드를 추출하는 과정과,

상기 특허문서에 포함된 IPC를 분리하여 저장하는 과정을 포함하는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 3

제 2항에 있어서, 상기 구조화된 특허 데이터는,

특허문서들을 나타내는 열(rows)과 키워드 또는 IPC코드에 대응하는 행(columns)의 행렬(matrix)을 포함하며, 행렬의 요소는 각 키워드 또는 IPC코드가 특정 특허문서에서 발생하는 빈도를 나타내는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 4

제 3항에 있어서, 상기 키워드 및 상기 IPC코드는,

카운트 회귀 모델 및 베이지안 네트워크를 기반으로 한 시각화를 통해 독립적으로 분석되는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 5

제 4항에 있어서, 상기 카운트 회귀 모델은,

포아송허들회귀분석(Poisson hurdle regression) 모델인 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 6

제 5항에 있어서,

상기 포아송허들회귀분석 모델은 상기 대상기술을 응답변수로 하고, 추출된 키워드 및 IPC코드를 설명변수로 하

며,

상기 응답변수 및 설명변수는 상기 베이지안 네트워크를 기반으로 한 시각화 과정에서 노드(nodes)로 표현되고, 상기 노드들 사이의 기술적 관계를 나타내는 에지(edges)로 표현되는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 7

제 6항에 있어서, 상기 가공된 특허 데이터에서 베이지안 네트워크를 기반으로 한 시각화를 수행하는 단계는, 전체 전체 네트워크에서 유의미한 서브 네트워크 추출하는 과정과,

추출된 서브 네트워크 간의 중요한 네트워크를 검정하는 과정으로 이루어지는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 8

제 7항에 있어서,

상기 전체 전체 네트워크에서 유의미한 서브 네트워크 추출하는 과정은 제약기반 접근법(Constraint-based approaches)을 사용하고,

상기 추출된 서브 네트워크 간의 중요한 네트워크를 검정하는 과정은 상기 특허문서 내의 키워드 또는 IPC코드의 관찰빈도를 통한 조건부 독립성 검정(Conditional independence test)인 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 9

제 8항에 있어서,

상기 대상기술은 상기 베이지안 네트워크를 기반으로 한 시각화 과정에서 대상노드(Target nodes)이고, 상기 추출된 키워드 및 IPC코드는 베이지안 네트워크 그래프의 설명노드(Explanatory nodes)인 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 10

제 1항에 있어서, 상기 구조화된 특허 데이터에 대하여 빈도회귀분석을 실시하는 단계는,

회귀분석을 이용하여 상기 특허 데이터에 포함된 출원국가와 상기 구조화된 특허 데이터와의 영향력을 산정하는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 11

제 1항에 있어서, 상기 구조화된 특허 데이터에서 베이지안 네트워크를 기반으로 한 시각화를 수행하는 단계는,

베이지안 네트워크를 이용하여 상기 특허 데이터에 포함된 출원국가와 상기 구조화된 특허 데이터와의 네트워크 밀도를 산정하여 시각화하는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 12

제 11항에 있어서, 상기 빈도회귀분석을 실시한 결과와 상기 수행된 베이지안 네트워크의 시각화 정보를 결합하는 단계는,

상기 특허 데이터에 포함된 출원국가와 모두 관련된 공통 키워드 및 공통 IPC코드로부터 지속 가능한 기술을 도출하는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

청구항 13

제 12항에 있어서,

상기 베이지안 네트워크의 시각화 정보로부터 상호연결이 많은 순서대로 상기 지속 가능한 기술을 창출하는 출원국가의 순위를 결정하고,

상기 공통 키워드 및 공통 IPC코드로부터 도출한 지속 가능한 기술에 상기 상호연결이 가장 많은 출원국가의 기술을 추가하여 세계적으로 지속 가능한 기술을 도출하는 것을 특징으로 하는 베이지안 네트워크를 이용한 기술문서 분석방법.

발명의 설명

기술 분야

[0001] 본 발명은 기술문서 분석방법에 관한 것으로, 보다 상세하게는 특허분석을 이용한 대상기술의 핵심기술을 분석할 수 있는 베이지안 네트워크를 이용한 기술문서 분석방법에 관한 것이다.

배경 기술

[0003] 기술은 국가 및 기업 경영에서 가장 중요한 요소 중 하나로서, 많은 기업들이 기술 경쟁력을 향상시키기 위해 혁신적인 기술을 연구하고 개발하며, 경쟁자들과의 경쟁에서 기술의 지속가능성을 유지하기 위해 기술분석을 수행한다.

[0004] 혁신적인 기술을 개발하고 신제품을 출시한 회사가 시장을 장악하게 되며, 이러한 기술혁신 및 기술혁신에 의한 신제품 개발은 기업기술의 지속가능성을 유지하는 데 있어 매우 중요한 요소이다.

[0005] 지속 가능한 기술관리에 있어 기술혁신은 매우 중요한 요소로서, 혁신적인 혁신적인 기술관리를 성공적으로 수행한 대표적 기업으로 애플(Apple Inc.)을 들 수 있다. 애플은 혁신적인 기술과 제품을 만드는 대표적인 회사로서, 아이폰(iPhone) 및 아이패드(iPad)와 같은 스마트 장치를 통해 전세계에 많은 고객을 보유하고 있다. 많은 연구소와 비즈니스 스쿨은 기술에 관한 다양한 서적과 연구논문을 통해 애플의 기술혁신을 연구했는데, 이 연구들 중 일부는 애플의 기술혁신을 이해하기 위해 애플의 특허를 사용하여 기술분석을 수행하였다.

[0006] Jun & Park(2013)은 기술적 측면에서의 혁신들을 연구하기 위하여 애플의 출원특허 및 등록특허를 분석하였다. 그들은 애플의 특허분석을 위해 데이터 마이닝 방법(Data mining methods)과 소셜 네트워크 분석기법(Social network analysis techniques)을 고려하였다. Kim & Jun(2015)은 애플의 기술분석을 위한 통계 모델을 제안했다. 이 연구에서는 그래픽 기술을 이용한 인과관계추론(Causal inference), 코풀라회귀(Copula regression) 및 텍스트 마이닝 기법(Text mining techniques)을 사용하여 애플의 기술과 관련된 특허 데이터를 분석하였다.

[0007] 이외에도, 다른 회사의 기술분석에 대한 많은 연구가 수행되었는데, Jun & Park(2016)은 BMW와 현대자동차의 기술을 분석하고 서로 비교하여 이 두 회사 간의 기술경쟁을 조사하였다. 이 연구에서도 두 경쟁사 간의 기술혁신과 진화에 대한 특허분석을 통해 한국의 자동차시장에서 BMW와 현대자동차의 지속가능성을 비교하였다.

[0008] 이러한 이전의 많은 연구들은, 주어진 기술분야에서 지속 가능한 기술은 특허 데이터를 기반으로 한 기술분석 결과를 사용하여 식별할 수 있음을 의미한다. 그러나 이전 연구에서 기술분석 방법의 대부분은 기술전문가의 지식과 경험, 특히 델파이(Delphi), 기술로드맵(Technology road-mapping) 또는 시나리오 분석(Scenario analysis)과 같은 기술을 사용하여 이루어졌다. Hung et al.(2012)는 최종 사용자와 전문가를 기반으로 한 델파이 조사를 통해 대만의 PC 생태계에 대한 기술전망을 제공하였고, Keller & Gracht(2014)는 기술예측분석(Technology forecasting analysis)을 위해 델파이 설문조사를 사용하였다.

[0009] 최근에는 정성분석과 정량분석의 결과를 결합한 하이브리드 기술분석(Hybrid technology analysis)에 대한 많은 연구가 수행되었다. Huang et al.(2014)는 연구개발(R&D) 계획을 조사하기 위해 계량분석(Bibliometrics; 양적

기술분석(quantitative technology analysis))과 로드맵(Road-mapping; 질적 기술분석(qualitative technology analysis))을 결합하였다. Zhang et al.(2014)는 객관적 분석으로서 유의 분석 트리즈(Purposive analysis TRIZ)를 통한 정량적 접근방식 및 기술로드맵을 통한 정성적 접근방식을 사용한 다변량 분석(Multivariate analysis; 클러스터링(clustering), 요인분석(factor analysis), 토픽 모델링(topic model))을 결합한 기술분석을 수행하였다. 이들은 특허의 제목과 초록에서 추출한 키워드를 사용하여 기술정보(technological intelligence)를 위한 용어를 정리하였다.

[0010] 그러나 전술한 정성분석과 정량분석의 결과를 결합한 하이브리드 기술분석은 기술분석의 일부 또는 전부에서 Delphi 설문조사 또는 기술로드맵 작성과 같이 전문가의 지식과 경험에 의존하기 때문에 결과가 선택한 전문가 그룹에 따라 다를 수 있어, 기본적으로 결과가 주관적이고 안정된 결과를 얻을 수 없는 문제점이 있다.

[0011] 따라서 객관적이고 안정된 결과를 제공할 수 있도록 통계 모델링을 사용하여 기술문서를 정량적으로 분석하는 방안이 요구된다.

선행기술문헌

특허문헌

[0013] (특허문헌 0001) KR 10-2016-0051671 A (2016.05.11.)

(특허문헌 0002) KR 10-2015-0139130 A (2015.12.11.)

(특허문헌 0003) KR 10-1370735 B1 (2014.02.27.)

(특허문헌 0004) KR 10-1364851 B1 (2014.02.12.)

비특허문헌

[0014] (비특허문헌 0001) 1. Choi, J.; Jun, S.; Park, S. A patent analysis for sustainable technology management. Sustainability 2016, 8, 688. [CrossRef]

(비특허문헌 0002) 2. Kim, S.; Jang, D.; Jun, S.; Park, S. A novel forecasting methodology for sustainable management of defense technology. Sustainability 2015, 7, 16720-16736. [CrossRef]

(비특허문헌 0003) 3. Park, S.; Lee, S.; Jun, S. A network analysis model for selecting sustainable technology. Sustainability 2015, 7, 13126-13141. [CrossRef]

(비특허문헌 0004) 4. Roper, A.T.; Cunningham, S.W.; Porter, A.L.; Mason, T.W.; Rossini, F.A.; Banks, J. Forecasting and Management of Technology; JohnWiley & Sons: Hoboken, NJ, USA, 2011.

(비특허문헌 0005) 5. Jun, S.; Park, S. Examining technological innovation of Apple using patent analysis. Ind. Manag. Data Syst. 2013, 113, 890-907. [CrossRef]

(비특허문헌 0006) 6. Kim, J.; Jun, S. Graphical causal inference and copula regression model for Apple keywords by text mining. Adv. Eng. Inform. 2015, 29, 918-929. [CrossRef]

(비특허문헌 0007) 7. Lashinski, A. Inside Apple; John Murray: London, UK, 2012.

(비특허문헌 0008) 8. Gallo, C. The Apple Experience; McGraw-Hill: Columbus, OH, USA, 2012.

(비특허문헌 0009) 9. Jun, S.; Park, S. Examining technological competition between BMW and Hyundai in the Korean car market. Technol. Anal. Strateg. Manag. 2016, 28, 156-175. [CrossRef]

(비특허문헌 0010) 10. Hung, C.; Lee, W.; Wang, D. Strategic foresight using a modified Delphi with end-user participation: A case study of the iPad's impact on Taiwan's PC ecosystem. Technol. Forecast. Soc. Chang. 2012, 80, 485-497. [CrossRef]

(비특허문헌 0011) 11. Keller, J.; Gracht, H.A.V.D. The influence of information and communication

- technology (ICT) on future foresight processes—Results from a Delphi survey. *Technol. Forecast. Soc. Chang.* 2014, 85, 81-92. [CrossRef]
- (비특허문헌 0012) 12. Pincombe, B.; Blunden, S.; Pincombe, A.; Dexter, P. Ascertaining a hierarchy of dimensions from time-poor experts: Linking tactical vignettes to strategic scenarios. *Technol. Forecast. Soc. Chang.* 2013, 80, 584-598. [CrossRef]
- (비특허문헌 0013) 13. Huang, L.; Zhang, Y.; Guo, Y.; Zhu, D.; Porter, A.L. Four dimensional science and technology planning: A new approach based on bibliometrics and technology roadmapping. *Technol. Forecast. Soc. Chang.* 2014, 81, 39-48. [CrossRef]
- (비특허문헌 0014) 14. Zhang, Y.; Porter, A.L.; Hu, Z.; Guo, Y.; Newman, N.C. “Term clumping” for technical intelligence: A case study on dye-sensitized solar cells. *Technol. Forecast. Soc. Chang.* 2014, 85, 26-39. [CrossRef]
- (비특허문헌 0015) 15. Liour, T.; Chen, C.W.; Chen, C.B.; Chen, C.C. White-light LED lighting technology life cycle forecasting and its National and company-wide competitiveness. In *Proceedings of the International Conference on Asia Pacific Business Innovation & Technology Management*, Bali, Indonesia, 23-25 January 2011; pp. 1-14.
- (비특허문헌 0016) 16. Sung, W.T.; Lin, J.S. Design and implementation of a smart LED lighting system using a self adaptive weighted data fusion algorithm. *Sensors* 2013, 13, 16915-16939. [CrossRef]
- (비특허문헌 0017) 17. The IHS Technology Lighting and LEDs Team. Top Lighting and LED Trends for 2015. Available online: <https://technology.ihs.com/api/binary/520405> (accessed on 1 August 2017).
- (비특허문헌 0018) 18. Chen, D.-Z.; Lin, C.-P.; Huang, M.-H.; Chan, Y.-T. Technology forecasting via published patent applications and patent grants. *J. Mar. Sci. Technol.* 2012, 20, 345-356.
- (비특허문헌 0019) 19. Kim, G.; Park, S.; Jun, S.; Kim, Y.; Kang, D.; Jang, D. A study on forecasting system of patent registration based on Bayesian network. *Intell. Inf. Manag.* 2012, 4, 284-290. [CrossRef]
- (비특허문헌 0020) 20. AZoOptics. Global Smart Lighting Market Forecast to Grow at 36.4% CAGR to 2018; TVILIGHT Empowering Intelligence: Amsterdam, The Netherlands, 2015.
- (비특허문헌 0021) 21. Park, S.; Kim, J.; Jang, D.; Lee, H.; Jun, S. Methodology of technological evolution for three-dimensional printing. *Ind. Manag. Data Syst.* 2016, 116, 122-146. [CrossRef]
- (비특허문헌 0022) 22. Choi, S.; Jun, S. Vacant technology forecasting using new Bayesian patent clustering. *Technol. Anal. Strateg. Manag.* 2014, 26, 241-251. [CrossRef]
- (비특허문헌 0023) 23. Chen, G.; Xiao, L. Selecting publication keywords for domain analysis in bibliometrics: A comparison of three methods. *J. Informetr.* 2016, 10, 212-223. [CrossRef]
- (비특허문헌 0024) 24. Zhu, Y.; Yan, E.; Song, I.Y. The use of a graph-based system to improve bibliographic information retrieval: System design, implementation, and evaluation. *J. Assoc. Inf. Sci. Technol.* 2017, 68, 480-490. [CrossRef]
- (비특허문헌 0025) 25. Glanzel, W. National characteristics in international scientific co-authorship relations. *Scientometrics* 2001, 51, 69-115. [CrossRef]
- (비특허문헌 0026) 26. Yan, E.; Ding, Y.; Zhu, Q. Mapping library and information science in China: A coauthorship network analysis. *Scientometrics* 2010, 83, 115-131. [CrossRef]
- (비특허문헌 0027) 27. Leydesdorff, L.; Bornmann, L. Mapping (USPTO) patent data using overlays to Google Maps. *J. Assoc. Inf. Sci. Technol.* 2012, 63, 1442-1458. [CrossRef]
- (비특허문헌 0028) 28. Zhang, Y.; Zhang, G.; Zhu, D.; Lu, J. Scientific evolutionary pathways: Identifying and visualizing relationships for scientific topics. *J. Assoc. Inf. Sci. Technol.* 2017.

[CrossRef]

(비특허문헌 0029) 29. Feinerer, I.; Hornik, K.; Meyer, D. Text mining infrastructure in R. J. Stat. Softw. 2008, 25, 1-54. [CrossRef]

(비특허문헌 0030) 30. Feinerer, I.; Hornik, K. Package 'Tm' Ver. 0.6, Text Mining Package, CRAN of R Project. 2017. Available online: <https://cran.r-project.org/web/packages/tm/tm.pdf> (accessed on 1 August 2017).

(비특허문헌 0031) 31. Choi, J.; Hwang, Y.S. Patent keyword network analysis for improving technology development efficiency. Technol. Forecast. Soc. Chang. 2014, 83, 170-182. [CrossRef]

(비특허문헌 0032) 32. Feng, X.; Fuhai, L. Patent text mining and informetric-based patent technology morphological analysis: An empirical study. Technol. Anal. Strateg. Manag. 2012, 24, 467-479. [CrossRef]

(비특허문헌 0033) 33. Kim, J.; Jun, S. Zero-Inflated Poisson and negative binomial regressions for technology analysis. Int. J. Softw. Eng. Appl. 2016, 10, 431-448. [CrossRef]

(비특허문헌 0034) 34. Cameron, A.C.; Trivedi, P.K. Regression Analysis of Count Data, 2nd ed.; Cambridge University Press: New York, NY, USA, 2013.

(비특허문헌 0035) 35. Hilbe, J.M. Negative Binomial Regression, 2nd ed.; Cambridge University Press: Cambridge, UK, 2011.

(비특허문헌 0036) 36. R Development Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2017.

(비특허문헌 0037) 37. Mullahy, J. Specification and testing of some modified count data models. J. Econom. 1986, 33, 341-365. [CrossRef]

(비특허문헌 0038) 38. Ridout, M.; Demetrio, C.G.; Hinde, J. Models for count data with many zeros. In Proceedings of the XIXth International Biometric Conference, Cape Town, South Africa, 14-18 December 1998; Volume 19, pp. 179-192.

(비특허문헌 0039) 39. Scutari, M. Learning Bayesian networks with the bnlearn R package. J. Stat. Softw. 2009, 35, 1-22.

(비특허문헌 0040) 40. Korb, K.B.; Nicholson, A.E. Bayesian Artificial Intelligence, 2nd ed.; CRC Press: London, UK, 2011.

(비특허문헌 0041) 41. The United States Patent and Trademark Office (USPTO). Available online: <http://www.uspto.gov> (accessed on 1 December 2016).

(비특허문헌 0042) 42. WIPS Corporation. Available online: <http://www.wipson.com> (accessed on 1 August 2017).

(비특허문헌 0043) 43. Korea Intellectual Property Strategy Agency (KISTA). Available online: www.kista.or.kr (accessed on 15 December 2016).

(비특허문헌 0044) 44. World Intellectual Property Organization (WIPO). Available online: www.wipo.org (accessed on 1 March 2017).

(비특허문헌 0045) 45. International Patent Classification (IPC), World Intellectual Property Organization (WIPO). Available online: <http://www.wipo.int/classifications/ipc/en> (accessed on 1 March 2017).

발명의 내용

해결하려는 과제

- [0015] 이에 본 발명자는 상술한 제반 사항을 종합적으로 고려하여 기존의 기술분석방법이 지닌 주관적이고 불안정한 한계 및 문제점의 해결에 역점을 두어, 객관적이고 안정된 결과를 제공할 수 있는 정량적 접근법을 통하여 지속 가능한 기술관리 방법을 개발하고자 각고의 노력을 기울여 부단히 연구하던 중 그 결과로써 본 발명을 창안하게 되었다.
- [0016] 따라서 본 발명이 해결하고자 하는 기술적 과제 및 목적은 특허문서와 같은 기술문서로부터 빈도 데이터와 베이지안 네트워크 기반의 시각화 정보에 근거하여 목표 기술분야의 객관적인 핵심 기술정보를 제공할 수 있도록 하는 베이지안 네트워크를 이용한 기술문서 분석방법을 제공하는 데 있는 것이다.
- [0017] 또한, 본 발명이 해결하고자 하는 다른 기술적 과제 및 목적은 통계모델 외에도 양적 기술분석을 통하여 지속 가능한 기술관리가 가능하도록 하는 베이지안 네트워크를 이용한 기술문서 분석방법을 제공하는 데 있는 것이다.
- [0018] 또한, 본 발명이 해결하고자 하는 또 다른 기술적 과제 및 목적은 베이지안 통계(Bayesian statistics)를 기반으로 한 시각화 방법을 고려하고, 이 시각화를 특허 카운트 데이터 모델링(Patent count data modeling)에 추가함으로써 해당 기술분야에서 세계적인 지속가능 기술을 찾기 위한 기술경영(MOT; Management of technology)에서 효율적으로 연구개발(R&D)을 계획하는데 기여할 수 있는 베이지안 네트워크를 이용한 기술문서 분석방법을 제공하는 데 있는 것이다.
- [0019] 여기서 본 발명이 해결하고자 하는 기술적 과제 및 목적은 이상에서 언급한 기술적 과제 및 목적으로 국한하지 않으며, 언급하지 않은 또 다른 기술적 과제 및 목적들은 아래의 기재로부터 당업자가 명확하게 이해할 수 있을 것이다.

과제의 해결 수단

- [0021] 상술한 목적을 달성하기 위한 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법은, 분석하고자 하는 대상기술과 관련된 특허문서를 수집하는 단계; 상기 수집된 특허문서에서 데이터를 추출하고 이를 분석할 수 있도록 구조화된 특허 데이터로 가공하는 단계; 상기 구조화된 특허 데이터에 대하여 빈도회귀분석을 실시하는 단계; 상기 구조화된 특허 데이터에서 베이지안 네트워크를 기반으로 한 시각화를 수행하는 단계; 및 상기 빈도 회귀분석을 실시한 결과와 상기 수행된 베이지안 네트워크의 시각화 정보를 결합하는 단계;를 포함하는 것을 특징으로 한다.
- [0022] 이때, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 수집된 특허문서를 분석할 수 있도록 구조화된 특허 데이터로 가공하는 단계는, 텍스트 마이닝 기법을 통해 상기 특허문서로부터 키워드를 추출하는 과정과, 상기 특허문서에 포함된 IPC를 분리하여 저장하는 과정을 포함하는 것을 특징으로 한다.
- [0023] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 구조화된 특허 데이터는, 특허문서들을 나타내는 열(rows)과 키워드 또는 IPC코드에 대응하는 행(columns)의 행렬(matrix)을 포함하며, 행렬의 요소는 각 키워드 또는 IPC코드가 특정 특허문서에서 발생하는 빈도를 나타내는 것을 특징으로 한다.
- [0024] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 키워드 및 상기 IPC코드는, 카운트 회귀 모델 및 베이지안 네트워크를 기반으로 한 시각화를 통해 독립적으로 분석되는 것을 특징으로 한다.
- [0025] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 카운트 회귀 모델은, 포아송허들회귀분석(Poisson hurdle regression) 모델인 것이 바람직하다.
- [0026] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 포아송허들회귀분석 모델은 상기 대상기술을 응답변수로 하고, 추출된 키워드 및 IPC코드를 설명변수로 하며, 상기 응답변수 및 설명변수는 상기 베이지안 네트워크를 기반으로 한 시각화 과정에서 노드(nodes)로 표현되고, 상기 노드들 사이의 기술적 관계를 나타내는 에지(edges)로 표현되는 것을 특징으로 한다.
- [0027] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 가공된 특허 데이터에서 베이지안 네트워크를 기반으로 한 시각화를 수행하는 단계는, 전체 전체 네트워크에서 유의미한 서브 네트워크

추출하는 과정과, 추출된 서브 네트워크 간의 중요한 네트워크를 검정하는 과정으로 이루어지는 것을 특징으로 한다.

[0028] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 전체 전체 네트워크에서 유의미한 서브 네트워크 추출하는 과정은 제약기반 접근법(Constraint-based approaches)을 사용하고, 상기 추출된 서브 네트워크 간의 중요한 네트워크를 검정하는 과정은 상기 특허문서 내의 키워드 또는 IPC코드의 관찰 빈도를 통한 조건부 독립성 검정(Conditional independence test)인 것을 특징으로 한다.

[0029] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 대상기술은 상기 베이지안 네트워크를 기반으로 한 시각화 과정에서 대상노드(Target nodes)이고, 상기 추출된 키워드 및 IPC코드는 베이지안 네트워크 그래프의 설명노드(Explanatory nodes)인 것을 특징으로 한다.

[0030] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 구조화된 특허 데이터에 대하여 빈도회귀분석을 실시하는 단계는, 회귀분석을 이용하여 상기 특허 데이터에 포함된 출원국가와 상기 구조화된 특허 데이터와의 영향력을 산정하는 것을 특징으로 한다.

[0031] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 구조화된 특허 데이터에서 베이지안 네트워크를 기반으로 한 시각화를 수행하는 단계는, 베이지안 네트워크를 이용하여 상기 특허 데이터에 포함된 출원국가와 상기 구조화된 특허 데이터와의 네트워크 밀도를 산정하여 시각화하는 것을 특징으로 한다.

[0032] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 빈도회귀분석을 실시한 결과와 상기 수행된 베이지안 네트워크의 시각화 정보를 결합하는 단계는, 상기 특허 데이터에 포함된 출원국가와 모두 관련된 공통 키워드 및 공통 IPC코드로부터 지속 가능한 기술을 도출하는 것을 특징으로 한다.

[0033] 또한, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 있어서, 상기 베이지안 네트워크의 시각화 정보로부터 상호연결이 많은 순서대로 상기 지속 가능한 기술을 창출하는 출원국가의 순위를 결정하고, 상기 공통 키워드 및 공통 IPC코드로부터 도출한 지속 가능한 기술에 상기 상호연결이 가장 많은 출원국가의 기술을 추가하여 세계적으로 지속 가능한 기술을 도출하는 것을 특징으로 한다.

발명의 효과

[0035] 전술한 바와 같이 본 발명은 비전문가에게 수집된 특허 데이터로부터도 목표기술에 대한 객관적인 핵심 기술정보를 제공할 수 있는 장점이 있다.

[0036] 또한, 본 발명은 특허에서 키워드와 IPC코드 간의 기술적 관계를 시각화(Visualization) 함으로써, 특허 및 기술분석에서 기술의 연관분포(Associative distribution)를 용이하게 파악할 수 있는 효과가 있다.

[0037] 또한, 본 발명은 스마트 LED 시스템을 비롯한 대상기술 분야에서 세계적인 지속가능 기술을 찾아 효율적으로 연구개발을 계획할 수 있는 효과가 있다.

[0038] 또한, 본 발명은 대상기술 분야에서 가장 활발한 연구개발을 수행하고 있는 국가를 식별하는데 중요한 기여를 할 수 있다.

도면의 간단한 설명

[0040] 도 1은 구조화된 특허 데이터를 도식화한 도면,

도 2는 본 발명에 따라 베이지안 네트워크를 이용한 스마트 LED 기술을 시각화한 도면,

도 3은 본 발명에 따른 세계적인 지속가능 기술을 위한 통계적 모델링 프로세스를 나타낸 도면,

도 4는 매년 등록된 특허 건수를 나타낸 그래프,

도 5는 본 발명에 따라 키워드에 의한 “스마트 LED”의 기술구조를 나타낸 도면,

도 6은 본 발명에 따라 IPC코드에 의한 “스마트 LED”의 기술구조를 나타낸 도면,

도 7a 내지 도 7d는 본 발명에 따른 베이지안 네트워크를 이용하여 IPC코드를 시각화한 도면,
 도 8a 내지 도 8d는 본 발명에 따른 베이지안 네트워크를 이용하여 키워드를 시각화한 도면,
 도 9는 본 발명에 따라 도출된 스마트 LED 시스템을 위한 세계적으로 지속 가능한 기술을 도식화한 도면.

발명을 실시하기 위한 구체적인 내용

- [0041] 이하에서는 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 대한 실시 예를 도면을 참조하여 상세히 설명한다. 이하에서 설명되는 실시 예는 통상의 지식을 가진 자에게 발명의 범주를 완전하게 알려주기 위하여 제공되는 것으로, 본 발명은 이하에서 개시되는 실시 예에 한정되지 않고 다양한 형태로 구현될 수 있다.
- [0042] 도면들 중 동일한 구성들은 가능한 한 어느 곳에서든지 동일한 부호들을 나타낸다. 하기의 설명에서 구체적인 특정 사항들이 나타나고 있는데, 이는 본 발명의 보다 전반적인 이해를 돕기 위해 제공된 것일 뿐, 본 발명을 특정한 실시 형태에 대해 한정하려는 것이 아니며, 본 발명의 사상 및 기술 범위에 포함되는 모든 변경, 균등물 내지 대체물을 포함하는 것으로 이해되어야 한다. 그리고 본 발명을 설명함에 있어, 관련된 공지 기능 혹은 구성에 대한 구체적인 설명이 본 발명의 요지를 불필요하게 흐릴 수 있다고 판단되는 경우 그 상세한 설명을 생략한다.
- [0043] 다르게 정의되지 않는 한, 기술적이거나 과학적인 용어를 포함해서 여기서 사용되는 모든 용어들은 본 발명이 속하는 기술 분야에서 통상의 지식을 가진 자에 의해 일반적으로 이해되는 것과 동일한 의미를 가지고 있다. 일반적으로 사용되는 사전에 정의되어 있는 것과 같은 용어들은 본 발명에서의 기능을 고려하여 정의된 것으로서, 이는 본 발명의 기술적 사상에 부합되는 개념과 당해 기술분야에서 통용 또는 통상적으로 인식되는 의미로 해석되어야 하며, 본 출원에서 명백하게 정의하지 않는 한, 이상적이거나 과도하게 형식적인 의미로 해석되지 않는다.
- [0044] 아울러 어떤 부분이 어떤 구성요소를 포함한다고 할 때, 이는 특별히 반대되는 기재가 없는 한 다른 구성 요소를 제외하는 것이 아니라 다른 구성요소를 더 포함할 수 있는 것을 의미한다.
- [0045] 첨부된 블록도의 각 블록과 흐름도의 각 단계의 조합들은 컴퓨터 프로그램 인스트럭션들에 의해 수행될 수도 있다. 이들 컴퓨터 프로그램 인스트럭션들은 범용 컴퓨터, 특수용 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서에 탑재될 수 있으므로, 컴퓨터 또는 기타 프로그램 가능한 데이터 프로세싱 장비의 프로세서를 통해 수행되는 그 인스트럭션들이 블록도의 각 블록 또는 흐름도의 각 단계에서 설명된 기능들을 수행하는 수단을 생성하게 된다.
- [0047] 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법은 주로 양적 특허분석에 기반을 두고 전문가의 주관적인 판단을 최대한 제한하기 때문에 객관적으로 기술을 분석할 수 있다. 본 발명은 통계 모델 외에도 양적 기술분석을 통한 지속 가능한 기술관리방법론을 개발하는 것을 목적으로 한다. 본 발명의 실제 적용 가능성을 검증하기 위해 본 발명이 적용되는 유망기술분야 중 하나를 일례로 들어 설명한다.
- [0048] 본 발명에서는 스마트 LED(Smart Light Emitting Diode) 시스템을 대상기술(Target technology)의 일례로 선택하여 실용영역에 적용하였다.
- [0049] 먼저, 선행 연구들을 살펴보면, Liour et al.(2011)는 로지스틱 회귀모델을 사용하여 전통적인 LED 기술이 2016년까지 포화단계에 도달할 것으로 예측하였다. 전통적인 LED 기술의 보급으로 인해 스마트 LED 시스템이 개발되었고, 스마트 LED 기술의 개발 및 관리에 대한 요구가 계속 증가하고 있다. Sung & Lin(2013)은 자체 적응 가중 데이터 융합(self-adaptive weighted data fusion)을 이용한 스마트 LED 시스템의 설계 및 구현을 제안하였다. 그들은 스마트 장치로 제어되는 스마트 LED 시스템을 개발하였다. IHS Technology Lighting & LEDs Team(2015), Chen et al.(2012), Kim et al.(2012)는 스마트 LED 기술의 개발 필요성에 대해 설명하였고, AZoOptics(2015)는 스마트 조명 시스템의 연평균 성장률이 2014년에서 2018년 사이에 36.4 %가 될 것이라고 예측하였다.
- [0050] 따라서 스마트 LED 시스템은 현재와 미래의 주요 신흥 기술로서, 본 발명에서는 스마트 LED 시스템의 기술분석을 통해 기술을 이해하고 분석결과를 활용하여 스마트 LED 분야에서 지속 가능한 기술을 확보하기 위한 효율적인 R&D 전략을 수립할 수 있음을 설명한다. 스마트 LED 시스템의 기술을 분석하기 위해서는 관련 특허를 분석할

필요가 있다.

- [0051] 특허분석과 관련된 많은 연구에서 특허문헌에서 키워드와 국제특허분류(IPC) 코드를 추출하고 이를 분석하였다. 키워드는 통계적 분석(Bibliometrics)과 같은 많은 분야에서 사용된다.
- [0052] Chen & Xiao(2016)는 통계적 분석을 위해 저널논문 데이터베이스에서 공개 키워드를 선택하여 조사하였다. 서지 정보(Bibliographic) 검색을 개선하기 위해 Zhu et al.(2017)은 자료(Papers)에서 추출한 용어구조(term structure)에 기반한 그래프 시스템을 사용하였다.
- [0053] 인용분석(Citation analysis)은 통계적 분석의 또 다른 유용한 기술로서, Glanzel(2001)은 국제과학분야에서 공동저자(co-authorship) 관계의 국가적 특징을 연구하였다. 또한, Yan et al.(2010)은 도서관과 정보과학을 매핑한 공동저자 네트워크 분석을 소개하였다. 그것들은 모두 저자들 간의 인용분석과 관련이 있다.
- [0054] 이에 따라, 본 발명에서는 특허발명자들을 분석하기 위해 공동저자 인용분석을 고려할 수 있다. 일반적으로, 정량적인 특허분석에 사용되는 데이터는 키워드의 출현횟수 또는 특허문서의 대표적 카운트 데이터(count data)인 IPC코드로 구성된다. 따라서 특허분석을 위해서는 카운트 데이터 모델(count data models)이 필요하다. 그러나 특허분석에 대한 대부분의 이전 연구에서는 분석결과의 정확성을 도출하는데 있어 카운트 데이터 모델을 고려하지 않았다.
- [0055] 따라서 이러한 문제를 해결하기 위해 본 발명에서는 카운트 데이터 모델을 이용하여 키워드와 IPC코드를 분석하는 방법을 제안한다. 또한, 본 발명에서는 특허에서 키워드와 IPC코드 간의 기술적 관계를 시각화한다. 시각화(Visualization)는 특허 및 기술 분석에서 기술의 연관분포(associative distribution)를 파악하는데 용이하다.
- [0056] Leydesdorff & Bornmann(2012)은 특허 데이터 매핑을 시각화하기 위해 구글지도(Google maps)를 사용하였다. 또한, Zhang et al.(2017)는 과학적 활동의 진화적 관계를 발견하기 위하여 과학지도(Science map)에 기반한 시각화 방법을 제안하였다.
- [0057] 본 발명에서는 베이지안 통계(Bayesian statistics)를 기반으로 한 시각화 방법을 고려하고, 이 시각화를 특허 카운트 데이터 모델링(Patent count data modeling)에 추가한다.
- [0058] 본 발명의 실시 예에서는, 스마트 LED 기술에 대한 사례연구를 통해 본 발명이 실제로 어떻게 적용될 수 있는지 설명한다.
- [0060] 먼저, 본 발명에 따른 베이지안 네트워크를 이용한 기술문서 분석방법에 적용되는 일반적인 카운트 데이터 모델링 기법을 살펴보면, 텍스트 마이닝 기술(Text mining techniques)을 사용하여 분석을 위해 수집된 특허문헌에서 키워드 및 IPC코드와 같은 기타 정보를 추출한다. 이 특허 키워드들과 IPC코드는 발생빈도로 표시된다. 빈도 값(Frequency values)은 일반적인 카운트 데이터이다.
- [0061] 기존의 많은 연구에서 빈도 데이터(Frequency data)는 연속 데이터 분석(Continuous data analysis)방법을 사용하여 분석되었다. 그러나 보다 효과적인 특허분석을 수행하려면 빈도 데이터 분석(Frequency data analysis)을 고려해야 한다. Kim & Jun(2016)은 영과잉-포아송분포(Zero-inflated Poisson distribution)와 음이항 모형(Negative binomial regression)을 사용하여 특허 키워드를 분석하였다. 이 방법은 카운트 데이터 분석에 사용할 수 있는 몇 가지 모델 중 하나이다.
- [0062] 많은 특허문서에서 응답변수 및 설명변수(키워드 또는 IPC코드)는 음수가 아닌 정수 데이터이다. 따라서 카운트 데이터 모델을 사용하여 설명변수(X)로 응답변수(Y)를 설명할 필요가 있다. 첫째, 응답변수가 카운트 데이터이기 때문에 이산확률분포(Discrete probability distribution)가 고려될 수 있다. 카운트 데이터 분석에 종종 사용되는 포아송분포는 하기의 수학적 1과 같은 확률질량함수(Probability mass function)를 가진다.

수학식 1

$$P(Y = y) = \frac{e^{-m} m^y}{y!}, m > 0 \text{ and } y = 0, 1, 2, \dots$$

[0064]

[0065]

여기서, m 은 포아송분포의 기본파라미터(Basic parameter)이며 이벤트 발생의 평균수를 나타낸다. Y 의 기대치는 m 이고, Y 의 분산도 m 이다. 본 발명에서는 포아송회귀분석(Poisson regression)을 카운트 데이터 모델로 간주한다. 포아송회귀분석은 포아송분포공식(Poisson distribution formula)과 그 파라미터(m)를 기반으로 한다. 포아송회귀분석은 카운트 데이터 분석을 위한 대표적인 모델이다. 포아송분포를 사용하여 다음과 같이 주어진 X 에 대한 Y 의 포아송밀도(Poisson density)를 구한다($f(Y = y_i | X = x_i)$). 포아송회귀분석모델(Poisson regression model)의 기본버전은 하기의 수학식 2와 같이 파라미터화 된다.

수학식 2

$$m = E(Y|X) = \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_p x_p)$$

[0067]

[0068]

여기서, m 은 주어진 X 에 대한 Y 의 기대치와 조건부 분산이 있는 포아송분포의 파라미터이다. 본 발명에서는 빈도분석을 위한 기법으로 포아송모델을 고려한다. 또한, 키워드와 IPC코드 사이의 연관성을 이해하기 위해 베이저안 네트워크(Bayesian networks)를 기반으로 하는 시각화를 고려한다.

[0070]

이하, 도 1 내지 도 3을 참조하여, 카운트 데이터 모델 및 베이저안 네트워크 기반 시각화를 이용한 기술분석을 상세히 설명한다. 도 1은 구조화된 특허 데이터를 도식화한 도면이고, 도 2는 본 발명에 따라 베이저안 네트워크를 이용한 스마트 LED 기술을 시각화한 도면이며, 도 3은 본 발명에 따른 세계적인 지속가능 기술을 위한 통계적 모델링 프로세스를 나타낸 도면이다.

[0071]

특허 데이터는 연구개발(R&D) 관리를 위한 기술을 분석할 수 있는 유용한 자원이다. 따라서 많은 연구에서 기술 분석을 위해 특허 데이터를 사용하였고, 채택된 특허분석방법도 매우 다양하다. 본 발명에서는 포아송허들회귀분석(Poisson hurdle regression)과 특허분석을 위한 베이저안 네트워크에 근거한 시각화를 결합한다. 하기에서 실시 예로서 설명할 스마트 LED 시스템 기술분야는 본 발명에 따른 기술문서 분석방법이 실제문제에 어떻게 적용될 수 있는지 보여준다.

[0072]

본 발명에 따른 베이저안 네트워크를 이용한 기술문서 분석방법에서는 먼저 키워드와 IPC코드를 특허문서에서 추출한다. 이 단계는 특허문서를 전처리하기 위해 텍스트 마이닝 기법이 사용된다.

[0073]

본 발명에서는 R 데이터 언어와 텍스트 마이닝 패키지 'tm'을 사용한다. 이 소프트웨어는 통계분석이 구조화된 데이터에만 적용 가능하므로, 통계 모델링을 사용하여 특허분석을 위해 문서를 구조화된 데이터로 변환하는 다양한 기능을 제공한다. 도 1은 본 발명에 따라 제안된 특허 데이터의 데이터 구조를 나타낸다.

[0074]

구조화된 데이터는 특허문서들을 나타내는 열(rows)과, 키워드 또는 IPC코드에 대응하는 행(columns)의 행렬(matrix)을 포함한다. 행렬의 요소는 각 키워드 또는 IPC코드가 특정 특허문서에서 발생하는 빈도를 나타낸다. 도 1에 도시된 바와 같이, 키워드 및 IPC코드는 편의상 하나의 구조화된 특허 데이터에 포함된다. 그러나 실제 분석과정에서 키워드 및 IPC코드 데이터는 카운트 회귀 모델 및 베이저안 네트워크를 기반으로 한 시각화를 통해 독립적으로 분석된다. 구조화된 데이터에 포함된 발생빈도는 전형적인 카운트 데이터이다. 따라서 본 발명에서는 카운트 분석 모델이 필요하고 포아송허들회귀분석을 고려한다. 이 모델은 Mullahy(1986)에 의해 소개되었으며 빈도 데이터를 0과 0이 아닌 부분으로 나누는 것에 의해 만들어졌다. 하기의 수학식 3은 Mullahy 및 Ridout 등이 제안한 포아송허들모델을 나타낸다.

수학식 3

$$P(Y = y) = \begin{cases} \pi, & y = 0 \\ (1 - \pi) \frac{(1-\pi)e^{-m}m^y}{(1-e^{-m})y!}, & y = 1, 2, 3, \dots \end{cases}$$

[0076]

[0077]

여기서 π 는 $Y = 0$, $p = P(Y = 0)$ 의 확률이다. 회귀분석은 모델을 0과 0이 아닌 부분으로 나누어 수행된다. 빈도가 0일 때 Y 는 π 에 의해 계산되고, 빈도가 0보다 큰 경우, Y 는 가중치가 $(1-\pi)$ 인 포아송분포로 모델링된다. 따라서 포아송허들모델은 특허 키워드의 빈도분석을 위한 효율적인 접근방법이다.

[0079]

또한, 본 발명에서는 응답변수 Y 에 대하여 두 가지 변수를 고려한다. 본 발명의 실시 예에서 대상기술이 스마트 LED 시스템과 관련되어 있기 때문에, 응답변수 Y 에 대한 두 가지 변수는 "Smart" 및 "LED"이다. 한편, Smart 및 LED 이외의 다른 모든 키워드는 설명변수로 사용된다. 캡처된 모든 IPC코드는 Smart 및 LED의 설명변수로 사용된다. 스마트 LED 시스템의 기술분석을 수행하기 위해 다음과 같은 포아송허들회귀모델 쌍을 고려한다.

[0081]

Model 1: (Smart + LED) = f (other keywords)

[0082]

Model 2: (Smart + LED) = f (all IPC codes)

[0084]

회귀파라미터(regression parameters)는 도 1의 구조화된 데이터를 두 개의 제안된 모델과 맞추므로써 추정된다. 본 발명은 키워드 및 IPC코드 분석결과를 활용하여 스마트 LED 시스템과 관련된 하위 기술 간의 기술적 관계를 파악할 수 있다.

[0085]

본 발명에서 기술분석을 위하여 고려된 다른 방법은 베이지안 네트워크를 기반으로 한 시각화이다. 본 발명에서는 베이지안 네트워크를 기반으로 한 이 시각화를 호출한다. 베이지안 네트워크는 그래프 이론에 기초한 통계모델이다. 이 모델에서 응답 및 설명 변수는 네트워크의 노드이다. 베이지안 네트워크에서 연결과 화살표는 노드 사이의 확률적 의존관계(Probabilistic dependencies)를 나타낸다. 베이지안 네트워크의 구조는 하기의 수학식 4와 같이 방향성 비순환 그래프(Directed Acyclic Graph, DAG)이다.

수학식 4

$$G_{\text{Bayesian network}} = (N, E)$$

[0087]

[0088]

여기서, N 과 E 는 각각 노드(nodes)와 에지(edges)를 나타낸다. 에지는 베이지안 네트워크의 노드 사이의 연결이다. 본 발명에서 N 은 키워드와 IPC코드로 구성되며 $N = \{\text{Smart}, \text{LED}, \text{기타 키워드}, \text{모든 IPC코드}\}$ 로 표현되며, E 는 키워드들 또는 IPC코드들 사이의 기술적 관계를 나타낸다. 노드들의 결합확률함수(joint probability function)는 다음과 같이 나타내진다.

수학식 5

$$p(n_1, n_2, \dots, n_p) = p(n_1)p(n_2|n_1) \cdots p(n_p|n_1, \dots, n_{p-1})$$

[0090]

[0091]

노드들이 키워드 또는 IPC코드를 나타내므로, 이 조건부 확률(Conditional probability)은 해당 키워드 또는

IPC코드의 그래프를 구성하는데 사용된다. 기술적 구조를 이해하기 위한 베이지안 네트워크 기반 시각화 과정은 두 단계로 구성된다.

- [0093] 1단계: 전체 네트워크에서 유의미한 서브 네트워크 추출.
- [0094] 2단계: 추출된 서브 네트워크 간의 중요한 네트워크 검정.
- [0096] 1단계에서 베이지안 네트워크를 기반으로 한 시각화를 위해 구조학습알고리즘(structure learning algorithm)이 사용된다. 시각화를 위해 학습하는 베이지안 구조는 주어진 기술 데이터에 대한 초기믿음(Initial belief)으로 시작된다. 초기믿음은 사전분포(Prior distribution)이며 개별 데이터의 우도함수(Likelihood function)로 곱해지며 갱신된 사후분포(Updated posterior distribution)를 야기한다. 이 알고리즘은 두 가지 가능한 방법, 즉 제약기반 접근법과(Constraint-based approaches) 및 점수기반 접근법(Score-based approaches)에 의존하는데, 제약기반 접근법이 통계적 검정을 위한 더 강력한 이론적 프레임워크(theoretical framework)를 제공하기 때문에 이를 사용한다.
- [0097] 2단계에서 조건부 독립성 검정(Conditional independence test)가 수행된다. 이 검정에서는 특허문서 내의 키워드 또는 IPC코드의 관찰빈도를 통해 조건부확률함수를 사용한다. 본 발명에서 검정 데이터는 다음과 같이 구성된다.
- [0099] 검정 데이터: Smart~LED | 키워드 그룹 + IPC코드 그룹.
- [0101] "물결표(~)"는 "~"의 양쪽에 있는 두 개의 키워드가 동시에 응답변수로 간주됨을 의미한다. 키워드 Smart 및 LED는 대상노드(Target nodes)이고 다른 키워드 및 IPC코드는 베이지안 네트워크 그래프의 설명노드(Explanatory nodes)이다. 도 2에 베이지안 네트워크를 기반으로 한 제안된 시각화의 기본 구조를 나타내었다. 도 2는 본 발명에 따라 베이지안 네트워크를 이용한 스마트 LED 기술을 시각화한 도면이다.
- [0102] 베이지안 네트워크는 Smart 및 LED 주변의 두 노드를 설명하는 다양한 노드로 구성된다. 물론 Smart 및 LED를 제외한 다른 노드들도 그들의 연결이 중요한 경우 서로 연결할 수 있다. 더하여, 본 발명은 시각화 결과의 통계적 유의성을 확인하기 위해 조건부 독립성 검정을 수행한다. 이때, 본 발명에서는 확률값(p-value)에 대한 임계값으로 0.05를 사용한다. 즉, p값이 0.05보다 작으면 통계적 검사결과가 유의하다고 결론지을 수 있다.
- [0103] 본 발명에서는 실시 예로서 카운트 데이터 분석과 베이지안 네트워크를 기반으로 한 시각화를 동시에 고려한 스마트 LED 시스템에 대한 기술분석을 수행한다. 도 3은 스마트 LED 시스템에서 지속 가능한 기술을 찾기 위해 본 발명에 따른 전반적인 절차를 도시한 도면으로, 본 발명에 따른 세계적인 지속가능 기술을 위한 통계적 모델링 프로세스를 나타낸 것이다.
- [0104] 키워드 및 IPC코드는 텍스트 마이닝 기술을 사용하여 생성된다. 본 발명은 이러한 구조화된 데이터를 사용하여 카운트 회귀 분석을 통해 지속 가능한 기술을 추출하고 베이지안 네트워크를 기반으로 한 시각화를 통해 스마트 LED 기술을 개발하는데 가장 활발한 국가를 인지한다. 결론적으로, 본 발명에 따른 기술문서 분석방법은 스마트 LED 시스템의 지속 가능한 기술환경에 대한 자세한 설명을 제공하고 스마트 LED 시스템에서 가장 활발하게 기술 개발이 이루어지는 국가를 식별한다.
- [0106] 다음으로, 일 실시 예로서, 스마트 LED 조명 시스템과 관련된 특허 데이터를 이용하여 기술분석을 수행함으로써, 본 발명에 따른 기술문서 분석방법이 실제영역에서 실제 문제에 어떻게 적용될 수 있는지 상세히 설명한다.
- [0107] 먼저, 관련 특허문서는 미국 특허청 및 WIPS의 특허 데이터베이스에서 검색하였고, 한국지적재산권전략연구원(KISTA)의 전문가들의 도움을 받아 스마트 LED 조명 시스템과 관련된 특허문서 수집을 위한 검색공식을 하기와 같이 구성하였다.

- [0109] 검색식 = (smart or intelligent or intelligence) and (light emitting or light-emitting or light diode or led chip) and (package or module or circuit or pcb or board or panel or performance or maintain or heat or fever or thermal or radiant or radiation or system or drawing design; layout or lay out or lay-out; sense or illumination or body or signal or movement or motion or gesture or communication or internet or correspondence or wire or wireless or radio or network or control or failure or breakdown or break-down or trouble or remote or warning or material or ingredient or resource or physical or part or component)
- [0111] 또한, 아시아와 유럽, 북미의 3개 지역에 대한 특허를 검색하였고, 아시아와 북미의 대표 국가로 중국과 미국을 선정하였다. 유효한 특허를 필터링하여 4226개의 샘플을 얻었으며, 이 중에서 중국, 유럽, 미국 특허 건수는 각각 3043건, 184건, 999건이다. 도 4는 매년 등록된 특허 건수를 나타낸 그래프이다.
- [0112] 도 4를 참조하면, 유럽과 미국에 비해 중국은 스마트 LED 기술에 대한 특허를 더 많이 등록하였다. 전반적으로, 2000년대 후반부터 스마트 LED 조명 시스템과 관련된 특허는 극적으로 증가하였다.
- [0113] 검색된 특허에 대하여, 텍스트 마이닝 기법을 사용하여 특허문서에서 키워드와 IPC코드를 추출하였다. 본 발명에 따른 실시 예에서는 텍스트 마이닝을 위해 R 데이터언어와 'tm' 패키지를 사용하였다.
- [0114] 특허문서에서 확인된 상위 10개의 주요 IPC코드를 하기의 표 1에 정리하였다. 표 1은 스마트 LED 조명 시스템과 관련된 특허의 상위 10개의 주요 IPC코드와 WIPO(World Intellectual Property Organization)의 기술정의를 포함한다.

표 1

IPC Code	Technological Definition
H05B	Electric heating and light
H01L	Semiconductor devices, electric solid state devices
F21S	Non-portable lighting devices and systems
F21V	Functional features or details of lighting devices or systems, structural combinations of lighting devices with other articles
G09G	Arrangements or circuits for control of indicating devices using static means to present variable information
G09F	Displaying, advertising, signs, labels or name-plates, seals
F21K	Light sources
G11B	Information storage based on relative movement between record carrier and transducer
C09K	Materials for applications, applications of materials
G08B	Signaling and calling systems, order telegraphs, alarm systems

[0116]

- [0118] 표 1에 나타난 상위 10개의 IPC코드는 스마트 LED 기술의 개발에 큰 영향을 미치는 것으로 알려져 있다. 우리는 각 IPC코드에 해당하는 기술이 어떻게 스마트 LED 기술에 영향을 줄 수 있는지 설명하기 위해 IPC코드를 사용하였다. 본 발명에 따른 통계모델에서 IPC코드는 독립변수이며 키워드 Smart 및 LED는 하기의 수학적식 6과 같이 종속변수가 된다.

수학식 6

$$(smart + LED) = f(H05B, H01L, F21S, F21V, G09G, G09F, F21K, G11B, C09K, G08B) + e$$

여기서 오류항(error term)은 $e \sim N(0, \sigma^2)$ 이다. 본 발명에 따라 세 지역에 대하여 통계적 회귀분석을 수행하여 하기 표 2의 결과를 얻었다.

표 2

Keyword	China		Europe		US		Total	
	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value
Control	0.0395	<0.0001	-0.0983	0.0155	0.0214	0.0511	0.0478	<0.0001
Lamp	0.0522	<0.0001	0.1150	<0.0001	0.0694	<0.0001	0.0633	<0.0001
Circuit	0.0187	<0.0001	0.0797	<0.0001	0.0459	<0.0001	0.0258	<0.0001
Power	0.0232	<0.0001	0.0430	0.2086	0.0540	<0.0001	0.0292	<0.0001
Device	0.0041	0.2990	-0.2465	<0.0001	0.0015	0.9087	0.0049	0.2030
Layer	-0.0832	<0.0001	-0.2011	0.0007	-0.1550	<0.0001	-0.0844	<0.0001
Signal	-0.0177	0.0006	-0.0049	0.8681	-0.0434	0.0009	-0.0296	<0.0001
Wireless	0.0400	<0.0001	0.3101	<0.0001	0.1358	<0.0001	0.0525	<0.0001
Material	-0.0812	<0.0001	-0.2032	0.0017	-0.1780	<0.0001	-0.0967	<0.0001
Heat	0.0277	<0.0001	-0.3520	0.0004	0.1082	<0.0001	0.0380	<0.0001

이 결과로부터 p값이 0.05 미만인 중요한 키워드를 찾을 수 있다. 지역에 따라 키워드 목록에 약간의 차이가 있지만 전반적인 결과는 서로 비슷하다는 것을 확인할 수 있다. 각 키워드의 추정치(Estimate value)는 스마트 LED에 대한 영향의 정도를 나타낸다. 예를 들어, 총계(Total)에서 키워드 control과 관련된 기술이 1단위 증가할 때, 추정치 0.0478은 스마트 LED와 관련된 전체 기술 개발이 0.0478 단위만큼 증가한다는 것을 의미한다. 총계의 추정치는 중국, 유럽 및 미국의 모든 데이터를 사용한 회귀분석의 결과이다. 표 2의 결과를 토대로 키워드 별로 스마트 LED 시스템의 기술구조를 도 5와 같이 생성하였다.

도 5는 본 발명에 따라 키워드에 의한 “스마트 LED”의 기술구조를 나타낸 도면으로서, 도 5에서 중국은 Control, Lamp, Circuit, Power, Layer, Signal, Wireless, Material 및 Heat를 기반으로 한 기술을 통해 스마트 LED 기술의 개발에 영향을 준다. 이들 키워드는 표 2에서 0.05미만의 p값을 갖는다. 유럽, 미국 및 총계의 결과는 중국의 경우와 마찬가지로 설명된다. 굵게 표시된 키워드는 모든 지역에서 공통적으로 사용되며, 스마트 LED 시스템을 위한 기반기술(Underlying technologies)을 대표한다.

한편, 본 발명에서는 IPC코드에 의한 기술분석에 대한 또 다른 접근법을 고려하였다. 하기의 표 3은 상위 10개의 IPC코드를 기반으로 한 통계분석 결과를 정리한 것이다.

표 3

IPC Code	China		Europe		US		Total	
	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value	Estimate	p-Value
H05B	0.3315	<0.0001	2.2759	<0.0001	0.2541	<0.0001	0.3119	<0.0001
H01L	-0.1796	<0.0001	0.0910	0.7691	-0.0355	0.0835	-0.1478	<0.0001
F21S	0.1471	<0.0001	-0.1040	0.9085	0.2351	<0.0001	0.1983	<0.0001
F21V	0.0372	<0.0001	0.9100	0.0142	0.1067	<0.0001	0.0670	<0.0001
G09G	0.0266	0.3145	-0.3980	0.6318	-0.7344	<0.0001	-0.0012	0.9610
G09F	0.1465	<0.0001	-0.7724	0.6061	0.3659	0.0022	0.2450	<0.0001
F21K	0.0740	0.1584	-0.9757	0.2625	0.0688	0.5120	-0.1021	0.0250
G11B	-6.4294	0.9406	-0.4168	0.2768	-3.2243	<0.0001	-3.2011	<0.0001
C09K	-0.9770	<0.0001	-0.6168	0.3789	-1.3882	0.0010	-0.9275	<0.0001
G08B	0.3501	0.0029	NA	NA	0.3269	<0.0001	-0.0323	0.6040

[0129]

[0131] 표 3에서 NA는 "not available"을 의미한다. 본 발명에서는 스마트 LED 시스템과 관련된 모든 특허에서 발생한 상위 10개의 IPC코드를 사용하였지만, G08B를 인용하는 유럽특허 중 샘플이 없기 때문에 모든 지역에 걸쳐 균일한 결과를 산출하지는 않는다. 표 2의 키워드 분석의 경우와 마찬가지로 p값이 0.05 미만인 IPC코드를 선택하여 기술분석에 사용하였다. 표 2의 키워드의 경우와 마찬가지로 각 IPC코드의 추정치는 전체 스마트 LED 기술에 대한 영향을 보여준다.

[0132] 표 3의 결과를 사용하여 도 6의 기술구조를 도출하였다. 도 6은 본 발명에 따라 IPC코드에 의한 “스마트 LED”의 기술구조를 나타낸 도면이다.

[0133] 키워드를 기반으로 한 기술구조와 대비할 때 IPC코드의 기술구조는 단순한 형식을 취한다. 이것은 IPC코드가 키워드보다 넓은 범위의 기술을 나타내기 때문이다. 도 5에 도시한 키워드 기술구조와 마찬가지로 도 6에 도시한 IPC코드의 경우에도 기술의 공통코드는 굵은 글꼴로 표시된다. 도 6을 참조하면, IPC코드 H05B 및 F21V가 스마트 LED 시스템의 개발을 위한 공통 코드로서, 하기와 같이 정의된다.

[0135] H05B: electric heating and lighting.

[0136] F21V: functional features or details of lighting devices or systems, structural combinations of lighting devices with other articles.

[0138] 따라서 도 4와 5의 결과를 사용하여 스마트 LED 시스템의 기술적 지속 가능성에 대한 대표적인 기술을 다음과 같이 도출하였다.

[0140] 지속 가능한 기술 1: electric devices for heating

[0141] 지속 가능한 기술 2: lamp and electric lighting

[0142] 지속 가능한 기술 3: functional structures combining lighting devices with wireless (circuit)

[0143] 지속 가능한 기술 4: functional features related to layers and materials

[0145] 이러한 결과는 스마트 LED 회사들이 위의 네 가지 기술에 대한 지속적인 연구개발(R&D)을 통해 기술 지속 가능성을 유지할 수 있음을 시사한다. 또한, 이 결과는 스마트 LED 시스템을 제조하는 회사들이 연구개발을 계획하는데 기여할 수 있다.

- [0147] 다음 단계는 스마트 LED 시스템의 기술분석에 대한 또 다른 접근방법인 시각화이다. 도 7a 내지 도 7d 및 도 8a 내지 도 8d은 지역에 따른 베이저안 네트워크를 기반으로 한 시각화 결과로서, 도 7a 내지 도 7d는 본 발명에 따른 베이저안 네트워크를 이용하여 IPC코드를 시각화한 도면이고, 도 8a 내지 도 8d은 본 발명에 따른 베이저안 네트워크를 이용하여 키워드를 시각화한 도면이다.
- [0148] 네트워크 밀도가 높을수록 기술에 대한 연구개발이 더욱 활성화된다. 따라서 중국의 노드(IPC코드)가 가장 많은 상호연결을 보여주기 때문에 스마트 LED 기술이 유럽이나 미국에 비해 중국에서 활발하게 개발되고 있음을 알 수 있다. 중국은 분명히 가장 적극적이며, 미국과 유럽이 뒤따른다.
- [0149] 도 7a 내지 도 7d의 시각화 결과를 사용하여 스마트 LED 기술에 직접 또는 간접적으로 영향을 주는 IPC코드의 관련성을 확인할 수도 있다. 예를 들어, 중국에서는 C09K기술이 LED 및 H05B기술에 영향을 끼치며, LED기술은 스마트기술에 영향을 준다. 또한, G11B기술은 스마트 LED 기술뿐만 아니라 다른 IPC코드 기술과도 관련이 없다. 이는 G11B노드가 다른 노드에 연결되어 있지 않고 단독으로 존재하기 때문이다.
- [0150] 도 7a 내지 도 7d의 IPC코드 시각화 결과와 마찬가지로 키워드 기반 시각화 결과인 도 8a 내지 도 8d을 참조하면, 중국은 베이저안 기술 키워드 네트워크를 기반으로 한 시각화에서도 가장 활동적인 국가로 나타났다. 또한, 시각화 결과의 화살표 연결을 통해 각 키워드 간의 기술 연관성을 파악할 수 있다. 예를 들어, 유럽에서는 무선 및 제어 기술이 스마트기술에 직접적인 영향을 미치는 것으로 나타났다. 또한, 장치기술은 무선기술을 통해 간접적으로 스마트기술에 영향을 준다. 따라서 지속 가능한 스마트 LED 기술을 창출하는 데 있어 각국의 순위는 다음과 같이 결정될 수 있다.
- [0152] 제1지역: 중국
- [0153] 제2지역: 미국
- [0154] 제3지역: 유럽
- [0156] 마지막으로, 본 발명에 따른 기술문서 분석방법의 시각화 구조의 결과에 대한 통계적 유의성을 입증하기 위해 조건부 독립성 검정을 수행하였다.
- [0157] 본 발명에서 상기 검정은 하기의 키워드 및 IPC코드를 사용하여 수행되었다.
- [0159] 키워드 검정식: Smart ~ LED | Control + Lamp + Circuit + Power + Device + Layer + Signal + Wireless + Material + Heat
- [0160] IPC코드 검정식: Smart ~ LED | H05B + H01L + F21S + F21V + G09G + G09F + F21K + G11B + C09K + G08B
- [0162] 하기의 표 4는 지역별 키워드 및 IPC코드의 조건부 독립성에 대한 검정 결과를 나타낸다.

표 4

Node	Region	p-Value
Keywords	China	1.896×10^{-4}
	Europe	6.778×10^{-2}
	US	5.981×10^{-5}
	Total	2.148×10^{-7}
IPC codes	China	1.581×10^{-9}
	Europe	1.121×10^{-5}
	US	5.700×10^{-8}
	Total	2.200×10^{-16}

[0164]

[0166] 유럽의 키워드 검정의 경우를 제외하고는 모든 결과가 통계적으로 유의미하였다. 따라서 본 발명에 따른 기술문서 분석방법에 있어서 베이지안 네트워크를 기반으로 한 시각화 결과가 통계적으로 유의하다는 것을 확인하였다. 그러므로 본 발명에 따른 기술문서 분석방법은 도 9에 나타난 바와 같이 카운트 회귀분석 모델의 결과와 베이지안 네트워크를 기반으로 한 시각화를 통해 스마트 LED 시스템에서 세계적인 지속 기술의 계층적 기술구조를 구축하였다.

[0167] 도 9는 본 발명에 따라 도출된 스마트 LED 시스템을 위한 세계적으로 지속 가능한 기술을 도식화한 도면으로서, 도 9를 참조하면, 본 발명은 중국, 미국, 유럽의 공통기술에서 추출한 4가지 지속 가능한 기술을 기반으로, 스마트 LED 시스템을 위한 지속 가능한 기술의 글로벌 개발구조에 중국의 고유기술들을 추가하였다. 이에 따라, 본 발명은 스마트 LED 시스템을 위한 세계적으로 지속 가능한 기술을 찾기 위해 4가지 지속 가능한 기술과 중국의 고유기술들을 통합하였다.

[0168] 그러므로 스마트 LED 시스템 분야에서, 기업 및 연구기관은 기술 지속 가능성을 위하여 “electric devices of heating”, “lamp and electric lighting”, “functional structure combining lighting devices with wireless”, 및 “functional features related to layers and materials”의 기술들을 습득해야 한다. 또한, 기업 및 연구기관은 “control, power, and signal”와 관련된 중국의 기술 및 하기의 다섯 가지 IPC코드와 관련된 기술들을 습득해야 한다.

[0170] H01L- basic electric elements

[0171] F21S- non-portable lighting devices; systems

[0172] G09F- displaying, advertising, signs, labels or name-plates; seals

[0173] C09K- material for applications not otherwise provided for; applications of materials not otherwise provided for

[0174] G08B- signaling or calling systems; order telegraphs; alarm systems

[0176] 본 발명에 따른 시험결과를 이용하여, KISTA의 스마트 LED 기술 전문가들의 도움을 받아 하기와 같은 지속 가능한 기술그룹을 도출하였다.

[0178] 1. basically electric lighting devices

[0179] 2. wireless control system

[0180] 3. layers and materials

[0181] 4. power signal

[0183] 따라서 상기 네 가지 기술을 사용하여 스마트 LED 시스템을 연구 및 생산하는 회사는 시장에서 세계적인 지속 가능성을 유지할 수 있다. 또한, KISTA 전문가 그룹은 본 발명에 따른 기술문서 분석방법에 대한 실현 가능성을 확인하였다.

[0184] 전술한 바와 같이, 본 발명은 스마트 LED 시스템의 기술분석을 실시 예로 들어 통계적 모델링을 기반으로 한 베이저안 네트워크를 이용한 기술문서 분석방법을 제안하였다.

[0185] 본 발명은 베이저안 네트워크를 기반으로 카운트 회귀분석 모델과 시각화를 고려하여, 스마트 LED 시스템과 같은 분석하고자 하는 기술과 관련된 특허문서를 수집하고 특허 키워드와 IPC코드로 구성된 구조화된 데이터 매트릭스로 변환한다. 이때, 행렬의 각 요소는 특허문서의 키워드 또는 IPC코드의 출현빈도이다.

[0186] 빈도값(카운트 데이터)을 분석하기 위해 카운트 데이터 분석에 적절한 포아송확률분포를 이용한 포아송 카운트 데이터 분석을 사용하고, 베이저안 네트워크를 사용하여 키워드 및 IPC코드의 시각화 네트워크를 형성한다. 실시 예에서 카운트 데이터 모델링 결과로 스마트 LED 시스템의 지속 가능한 기술을 확인하였다.

[0187] 가장 활동적인 국가는 베이저안 네트워크를 기반으로 한 시각화 결과를 사용하여 확인되며, 베이저안 네트워크를 기반으로 한 카운트 데이터 모델링과 시각화 결과를 결합하여 스마트 LED 네트워크를 위한 세계적인 지속 가능한 기술을 도출하였다. 이에 따라, 스마트 LED 기술은 네 가지 지속 가능한 기술을 기반으로 중국 고유의 스마트 LED 기술을 추가함으로써 기술 지속성을 유지할 수 있다.

[0188] 본 발명에 따른 베이저안 네트워크를 이용한 기술문서 분석방법은 스마트 LED뿐만 아니라 다양한 다른 기술분야에도 적용될 수 있으며, 세계적으로 지속 가능한 기술을 찾기 위한 기술경영(MOT; Management of technology)에서 효율적으로 연구개발(R&D)을 계획하는데 기여할 수 있으며, 대상기술 분야에서 가장 활발한 연구개발을 수행하고 있는 국가를 식별하는데 중요한 기여를 할 수 있다.

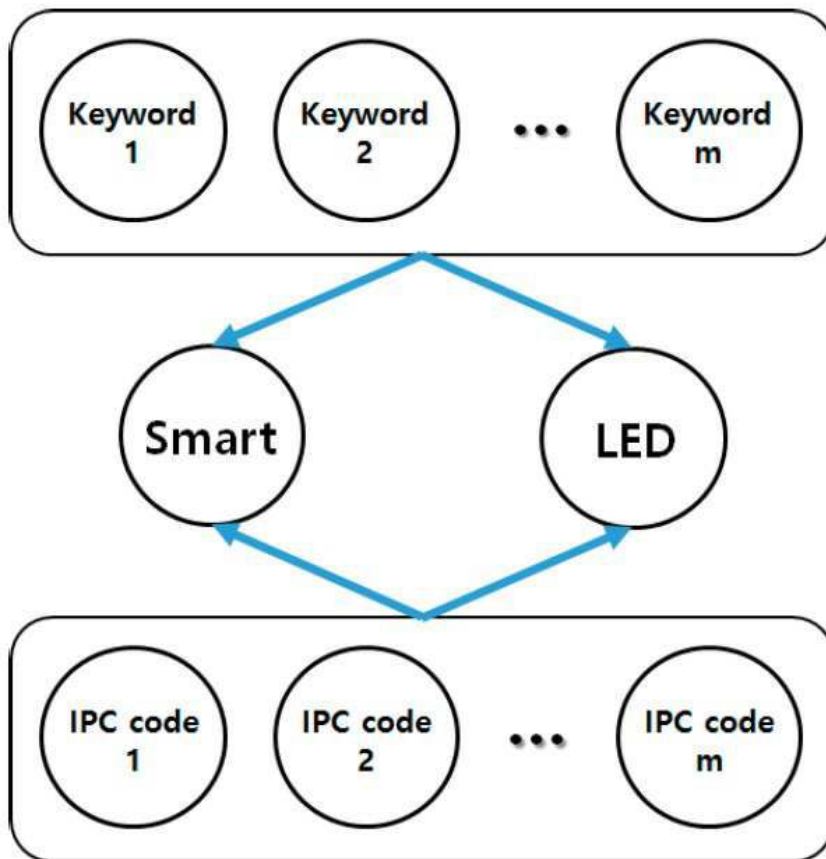
[0189] 한편, 본 발명의 상세한 설명에서는 첨부된 도면에 의해 참조되는 바람직한 실시 예를 중심으로 구체적으로 기술되었으나, 본 발명의 범위에서 벗어나지 않는 한도 내에서 여러 가지 변형이 가능함은 물론이다. 그러므로 본 발명의 범위는 설명된 실시 예에 국한되어 정해져서는 안되며 후술하는 특허청구의 범위뿐 아니라 이 특허청구의 범위와 균등한 것들에 의해서 정해져야 한다.

도면

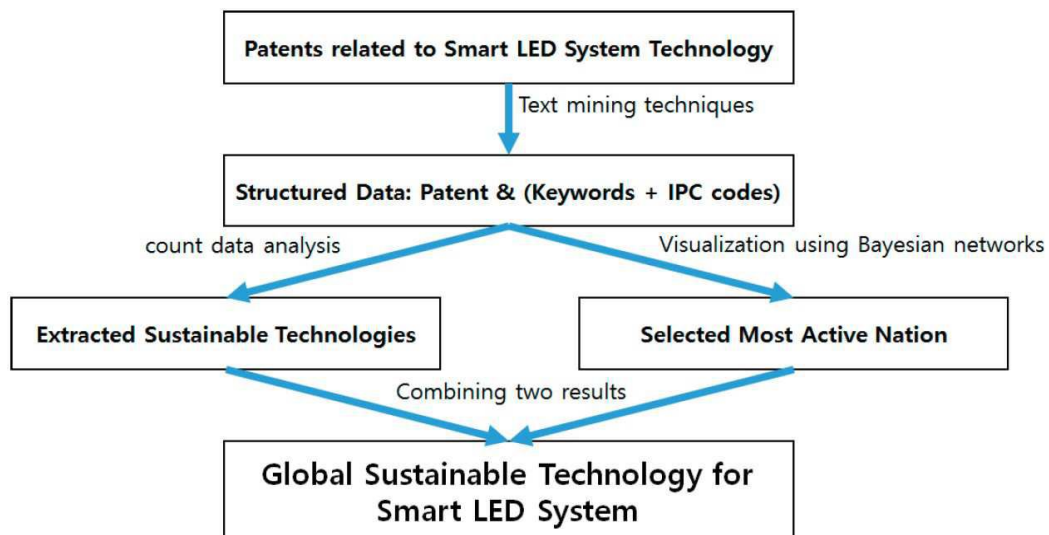
도면1

Patent	Keyword ₁	Keyword ₂	...	Keyword _{p1}	Patent	IPC ₁	IPC ₂	...	IPC _{p2}
1	Freq _{ij} : Occurred frequency of Keyword _j in patent _i				1	Freq _{ij} : Occurred frequency of IPC _j in patent _i			
2					2				
i					i				
n					n				

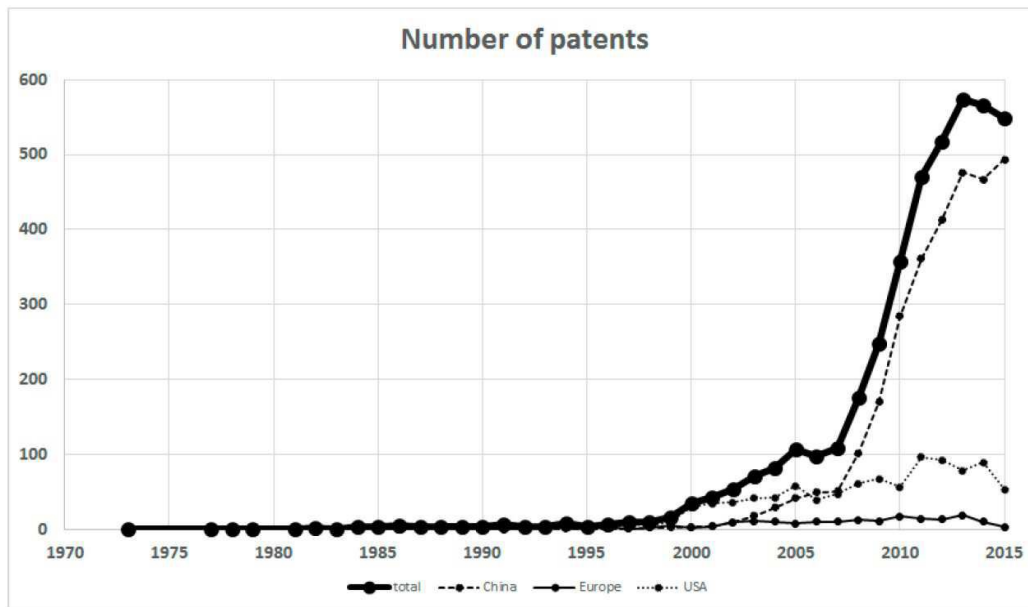
도면2



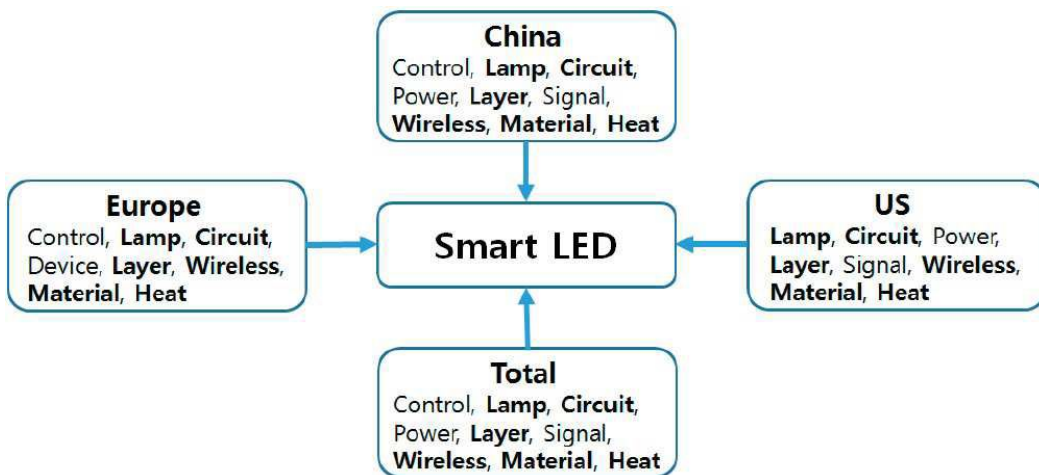
도면3



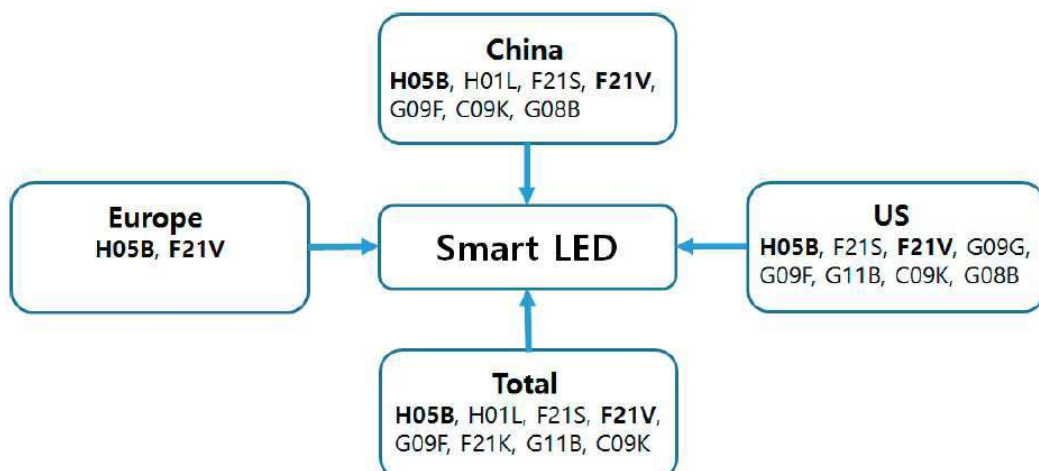
도면4



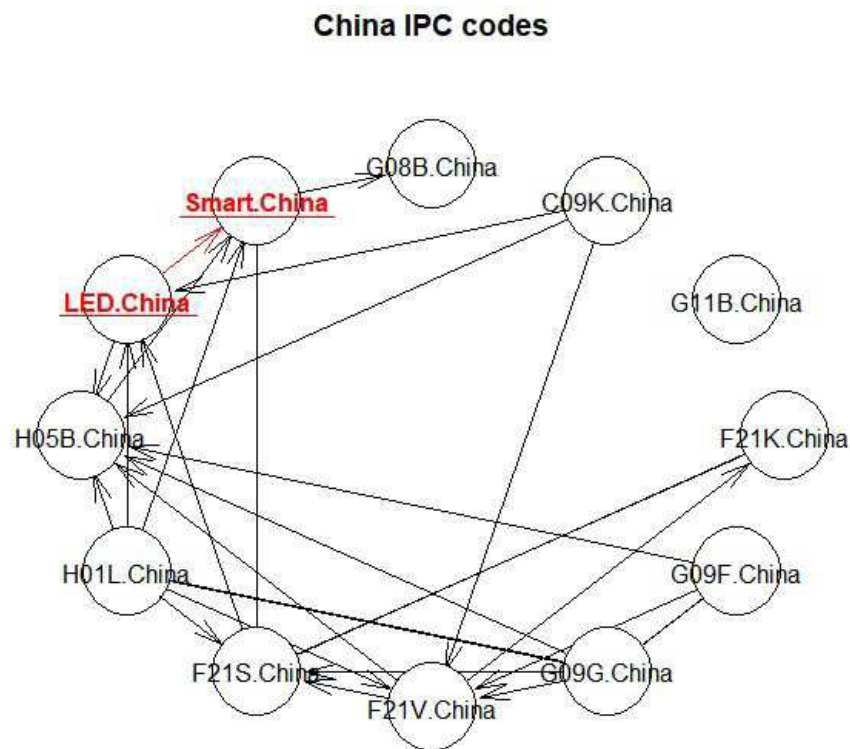
도면5



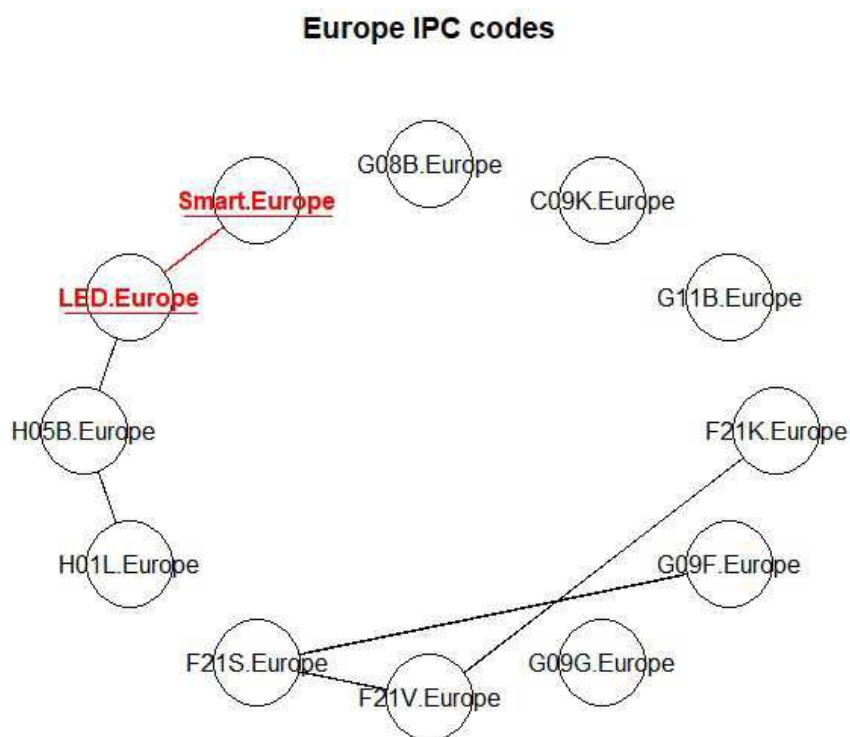
도면6



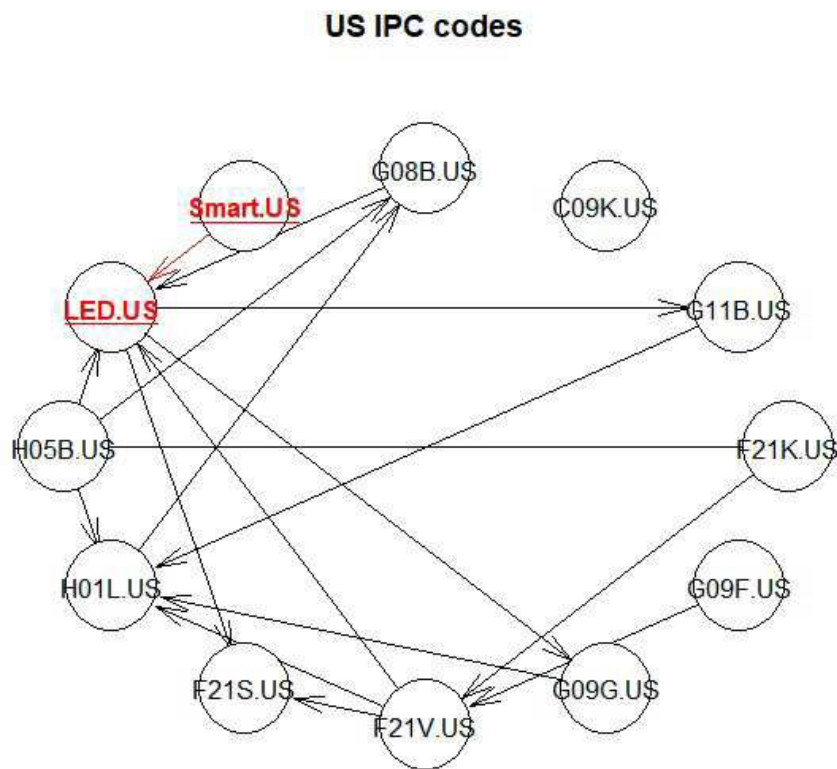
도면7a



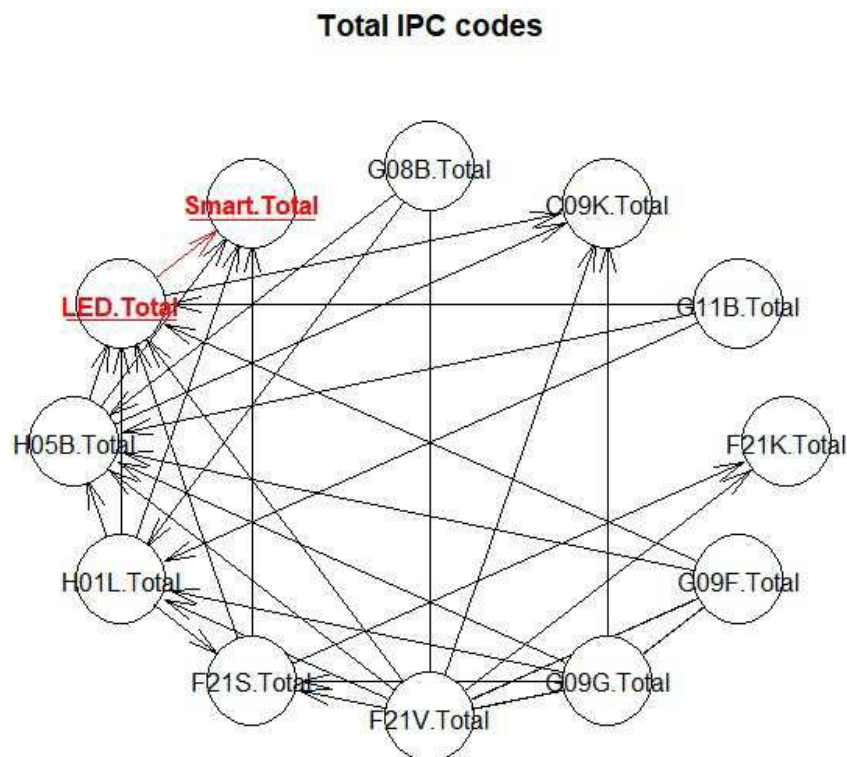
도면7b



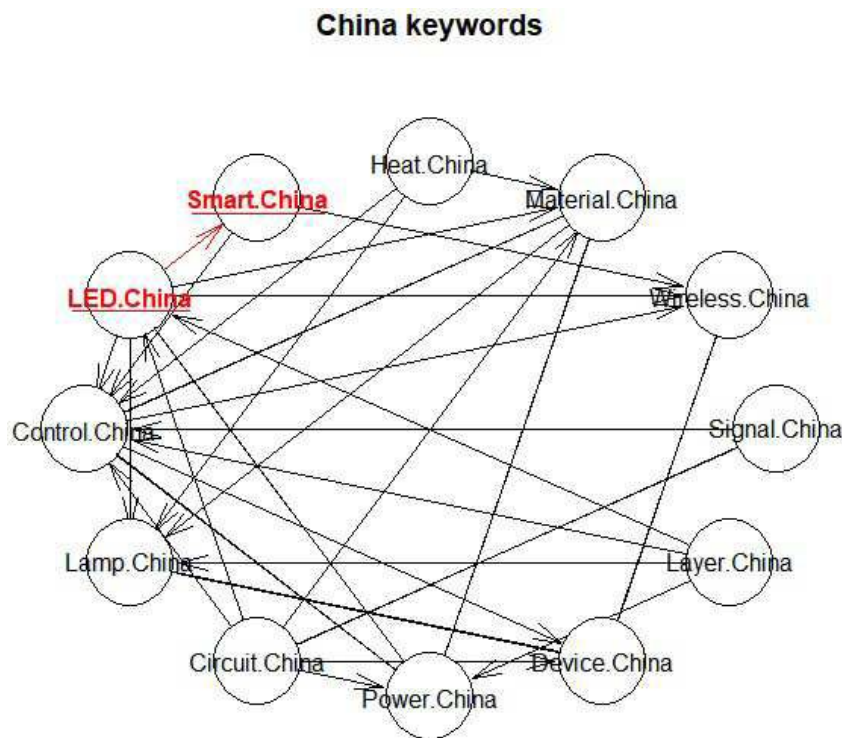
도면7c



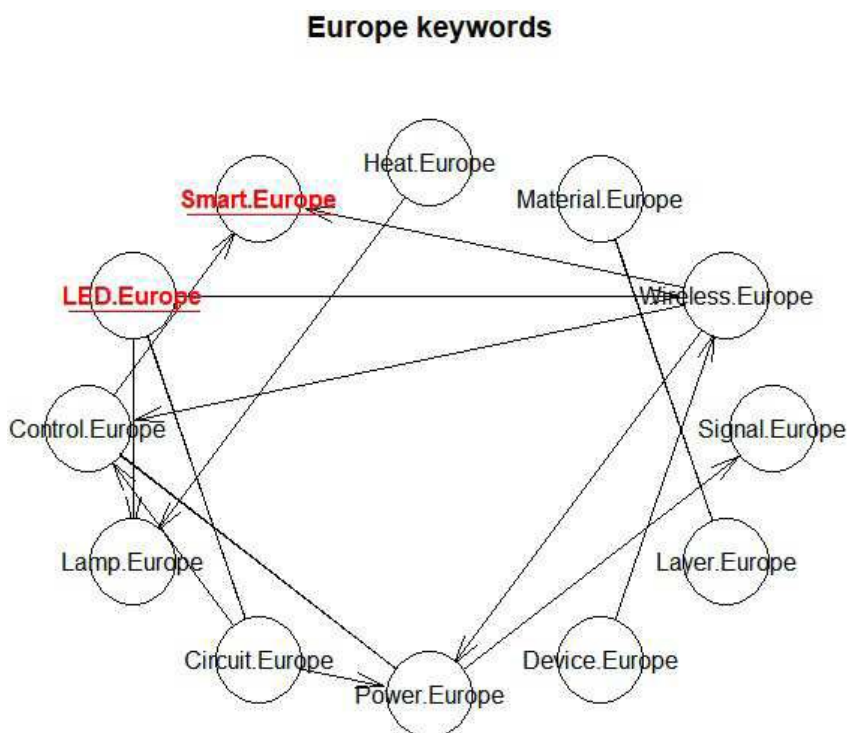
도면7d



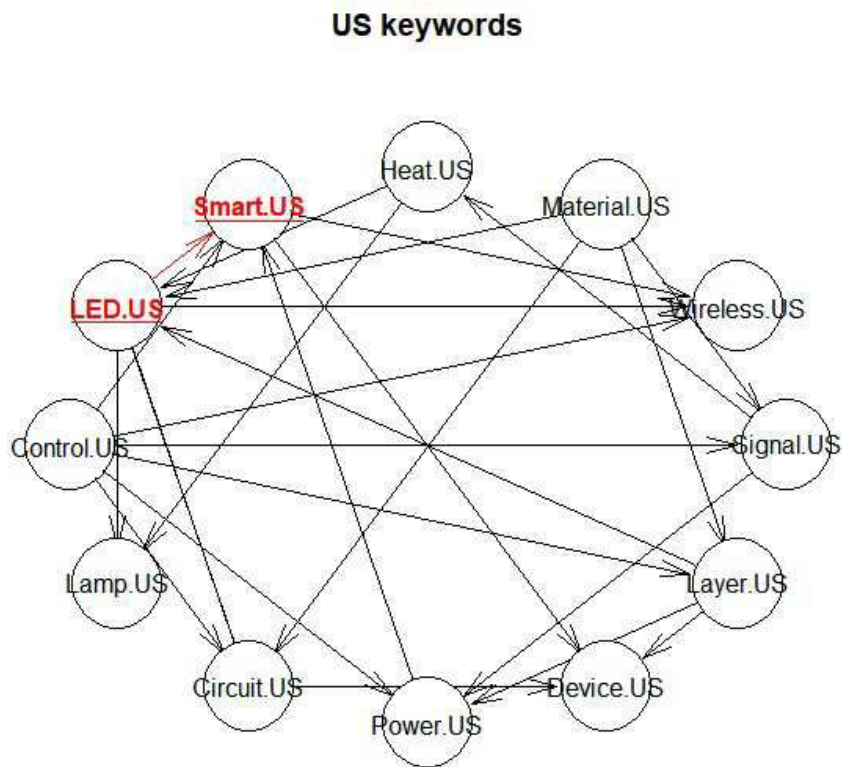
도면8a



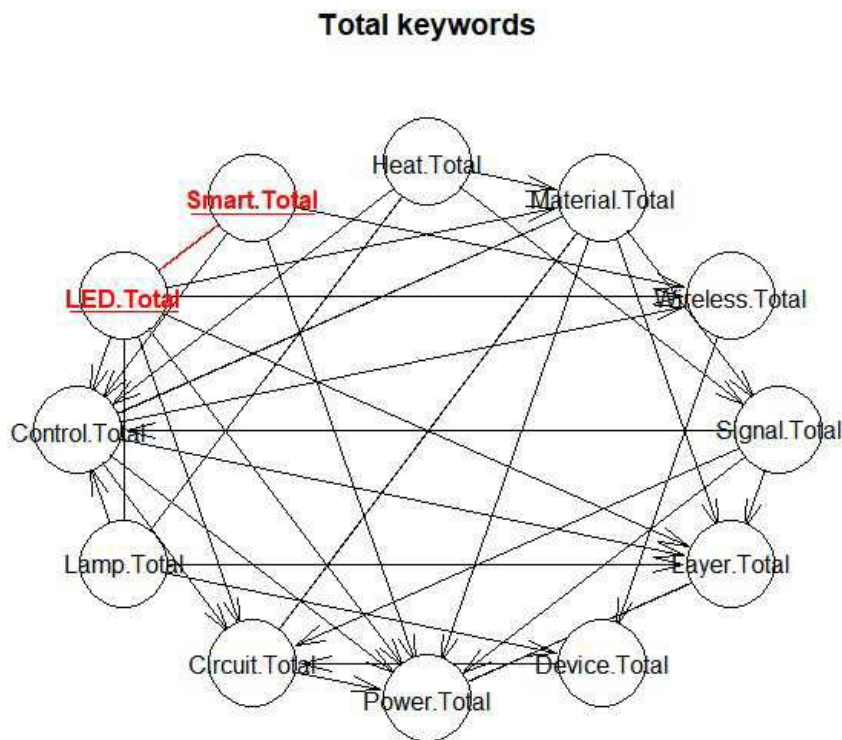
도면8b



도면8c



도면8d



도면9

