

CPC 코드 기반 사물인터넷(IoT) 특허의 기술 연관성 규칙 분석

심재륜*

Analysis of Technology Association Rules Between CPC Codes of the 'Internet of Things(IoT)' Patent

Jaeruen Shim*

요약 본 연구는 4차 산업혁명 ICT 기반기술의 핵심인 사물인터넷 특허의 CPC 코드 기반 기술 연관성 규칙 분석에 관한 것이다. 데이터 마이닝을 위한 오픈 소스인 R을 이용하여 CPC 코드간 기술 연관성 규칙을 도출하였다. 이를 위해 2019년 7월까지 특허청에 출원된 사물인터넷(Internet of Things) 관련 특허 605건 중 복합 CPC 코드를 가지는 369건을 대상으로 서브클래스(Subclass) 수준까지 분석하였다. 기술 연관성 규칙 분석 결과 지지도가 높은 CPC 코드는 [H04W → H04L](18.2%), [H04L → H04W](18.2%), [G06Q → H04L](17.3%), [H04L → G06Q](17.3%), [H04W → G06Q](9.8%), [G06Q → H04W](9.8%), [G06F → H04L](7.9%), [H04L → G06F](7.9%), [G06F → G06Q](6.2%), [G06Q → G06F](6.2%), [G06F → G06Q](6.2%) 순이고, CPC 코드간 상호 연결망을 분석한 결과 기술 연관성 관련 핵심 CPC 코드는 G06Q와 H04L이다. 본 연구 결과를 활용하면 앞으로의 특허 경향을 예상해 볼 수 있다.

Abstract This study deals with the analysis of the technology association rules between CPC codes of the Internet of Things(IoT) patent, the core of the Fourth Industrial Revolution ICT-based technology. The association rules between CPC codes were extracted using R, an open source for data mining. To this end, we analyzed 369 of the 605 patents related to the Internet of Things filed with the Patent Office until July 2019, with a complex CPC code, up to the subclass-level. As a result of the technology association rules, CPC codes with high support were [H04W → H04L](18.2%), [H04L → H04W](18.2%), [G06Q → H04L](17.3%), [H04L → G06Q](17.3%), [H04W → G06Q](9.8%), [G06Q → H04W](9.8%), [G06F → H04L](7.9%), [H04L → G06F](7.9%), [G06F → G06Q](6.2%), [G06Q → G06F](6.2%). After analyzing the technology interconnection network, the core CPC codes related to technology association rules are G06Q and H04L. The results of this study can be used to predict future patent trends.

Key Words : Technology Association Rules, CPC(Cooperative Patent Classification), Internet of Things, Support, Confidence

1. 서론

4차 산업혁명의 핵심 기술인 사물인터넷은 우리 주변의 사물과 사물, 사람과 사물을 연결시키고 지능화시키는 최고의 통신 인프라이다. 특히 사물인터넷 기술은 서로 다른 분야의 기술들이 모여서 새로운 부가가치를 만들어내고 있는 기술 융복합이

활발히 이루어지고 있는 분야이다.

일반적으로 기술 융복합은 서로 다른 기술 분야의 기술들이 상호 결합하여 형성된 새로운 기술 분야라고 정의할 수 있다[1]. 특허 관점에서 기술 융복합과 관련한 개념은 명확하게 정립되어 있지 않지만, 다양한 방법론으로 기술 융복합 분석이 시도되고 있다

This work was supported by the research grant of the Busan University of Foreign Studies in 2019.

* Division of Embedded IT, Busan University of Foreign Studies(BUFS)

Received August 21, 2019

Revised September 04, 2019

Accepted September 09, 2019

[2-3]. 최근 특허문헌의 기술 분류 코드를 이용하여 특허 기술의 기술 융복합을 분석한 사례로는 IPC(International Patent Classification) 코드의 주분류-부분류 관계를 활용하여 사물인터넷 특허의 기술 융복합을 분석한 결과[4]와 CPC(Cooperative Patent Classification)의 상세 분류체계를 이용하여 특허에 포함된 기술 분류 분석 모델을 제안한 연구[5] 및 CPC 코드 기반 전자상거래(G06Q) 분야에서 사물인터넷 기술의 기술 융복합을 분석한 결과[6] 등이 있고, 데이터 마이닝에서 활용되고 있는 연관 규칙(Association Rules) 마이닝을 이용하여 IPC 코드 분석을 시도하기도 하였다[7].

현재 특허청에서는 4차 산업혁명과 관련된 종전 7대 기술분야(인공지능, 빅데이터, 클라우드, 사물인터넷, 지능형 로봇, 자율 주행차, 3D 프린팅)를 16대 기술분야(인공지능, 빅데이터, 클라우드 컴퓨팅, 차세대 통신, 사물인터넷, 지능형 로봇, 자율 주행차, 드론, 가상증강현실, 스마트시티, 맞춤형 헬스케어, 혁신신약, 지능형 반도체, 첨단소재, 신재생 에너지, 3D 프린팅)로 확대 및 개선하여 CPC 코드 체계에 추가하여 수립하였다[8].

특허문헌에서 특허 분석을 통한 기술 트렌드를 예측하고 기술 융복합의 방향성을 가능할 수 있는 특허 정보학(Patent Informatics & Informetrics)이 중요해지고 있다[9]. 이에 본 연구에서는 사물인터넷의 기술 융복합과 특허 경향을 확인하기 위해 2019년 7월까지 특허청에 출원된 사물인터넷 관련 특허 605건 중 복합 CPC 코드를 가지는 369건을 대상으로 기술 연관성 규칙을 분석하였다.

2. 기술 연관성 규칙 분석 및 CPC 코드

2.1 기술 연관성 규칙 분석

특허문헌의 CPC 코드간 의미있는 기술 연관성 규칙 분석을 위해 연관성 분석의 3가지 척도(measure)인 지지도(support), 신뢰도(confidence), 향상도(lift)의 기본 정의는 다음 식과 같다[10].

$$\text{Support}(A \rightarrow B) = P(A \cap B) \quad (1)$$

$$\begin{aligned} \text{Confidence}(A \rightarrow B) &= P(B|A) \quad (2) \\ &= \frac{P(A \cap B)}{P(A)} \end{aligned}$$

$$\text{Lift}(A \rightarrow B) = \frac{\text{Confidence}(A \rightarrow B)}{\text{Support}(B)} \quad (3)$$

지지도는 A와 B가 동시에 일어날 확률이고, 신뢰도는 A가 발생한 조건에서 B가 발생할 확률이다. 향상도는 A와 B의 신뢰도를 B의 지지도로 나눈 값으로 1보다 크면 A와 B가 상호 보완 관계이고, 1보다 작으면 상호 대체이며, 1인 경우 상호 독립을 뜻한다. 지지도의 경우 $\text{Support}(A \rightarrow B)$ 와 $\text{Support}(B \rightarrow A)$ 의 확률이 같기 때문에 신뢰도를 통해 CPC 코드간 기술의 선후(전후) 관계를 파악한다. 아이템과 아이템간의 트랜잭션(transaction) 집합은 각 특허문헌의 복합 CPC 코드를 할당한다[11].

2.2 CPC 코드

CPC(Cooperative Patent Classification)는 미국특허청과 유럽특허청이 공동으로 개발·관리·유지하는 특허분류체계로 IPC와 마찬가지로 섹션, 클래스, 서브클래스 및 메인그룹 또는 서브그룹의 계층적 구조로서 하나의 서브클래스 기호에 1~3자리 숫자, 사선(/) 및 2~6자리 숫자가 IPC 표준에 따라 구성됨으로써 IPC보다 세분화되고 더 많은 글자가 추가된 형태이다[12].

CPC(또는 IPC)는 발명과 관련된 기술내용(혹은 기술주제, 기능 및 용도 포함)이 단독일 경우 '하나의 주분류'를 부여하고, 기술내용이 여러 개일 경우에는 출원된 발명을 대표할 수 있는 기술내용을 '주분류(Main Category)'로 제시하고, 그 외의 다른 기술내용을 '부분류(Sub Category)'로 제시한다.

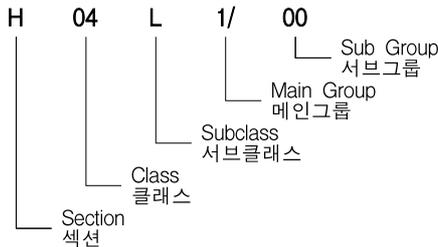


그림 1. H04L 1/00 CPC 코드 체계
Fig. 1. H04L 1/00 CPC Code Hierarchy

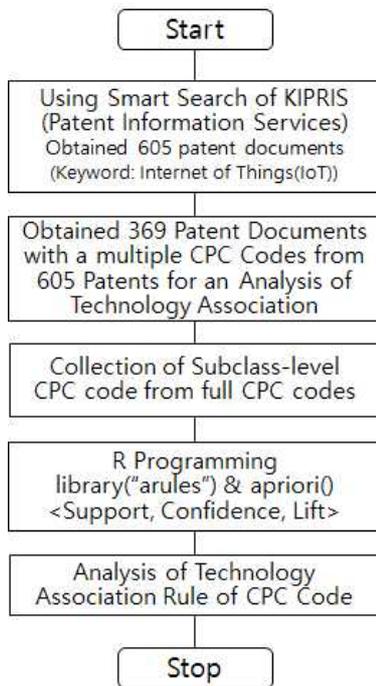


그림 2. 특허문헌을 통한 사물인터넷 기술 연관성 규칙 분석 절차
Fig. 2. Analysis Procedure of Technology Association Rules through Patents

3. 자료 수집 및 연구방법

그림 2에 사물인터넷 기술의 기술 연관성 분석 절차를 제시하였다.

사물인터넷 관련 특허 자료는 특허청의 특허정보넷(KIPRIS)의 스마트 검색을 통해 확보하였다[13]. 먼저 특허정보넷(KIPRIS, www.kipris.or.kr)의 「스마트 검색」 창에서 ‘발명(고안)의 명칭(TL)’에 ‘사물인터넷’, ‘IoT’, ‘Internet of Things’를 입력

(TL=[IoT+사물인터넷+Internet*of*Things])하여 2019년 7월까지 특허청에 출원된 사물인터넷 특허 605건을 1차적으로 확보하였다. 이후 검색된 각 특허의 CPC 코드를 확인하면서 서브클래스 수준의 CPC 코드간 중복 여부를 확인하였다. 이런 과정을 통해 서브클래스 수준에서 중복되지 않은 2개 이상의 CPC 코드를 가지는 특허문헌 369건을 최종 확보하였다.

마지막으로 데이터 마이닝을 위한 오픈 소스인 R의 library(“arules”) 패키지와 apriori() 함수를 이용하여 데이터를 트랜잭션 클래스 형태로 변환하고, 지지도와 신뢰도 및 향상도 등의 연관성을 분석하였다[14].

4. 기술 연관성 분석 결과

본 연구에서 확보한 605건의 사물인터넷 관련 특허 중 단일 CPC 코드를 가지는 특허는 236건이고, 복합 CPC 코드를 가지는 특허는 369건이다.

표 3과 표 4에 사물인터넷 관련 특허의 기본적인 특허 정보 현황을 정리하였다. 표 3은 사물인터넷 관련 특허문헌 605건의 출원년도별 법적 상태를 정리한 것으로 등록(Registration) 297건(49%), 공개(Exposure) 225건(37%), 거절(Rejection) 68건(11%), 취하(Withdrawal) 9건, 포기(Abandonment) 5건, 소멸(Extinction) 1건 등으로 이루어져 있다. 표 3에서 보듯이 2015년 이후부터 사물인터넷 관련 특허 출원이 크게 늘어나고 있다.

표 4에 단일 CPC 코드를 가지는 사물인터넷 관련 특허 236건의 CPC 코드 현황을 조사하였다. 조사결과 H04L【‘디지털 정보의 전송’】(157건)과 H04W【‘무선통신 네트워크’】(26건)이 전체의 약 78%를 차지한다. H04L과 H04W이 사물인터넷 관련 특허의 가장 대표적인 CPC 코드에 해당한다.

표 3. 사물인터넷 관련 특허의 출원년도별 법적 상태
Table 3. Legal Status of IoT Patent by Year of Filing

Year of Filing	Legal Status						No. of Patent App.
	Exposure	Registration	Rejection	Abandonment	Withdrawal	Extinction	
2019	5	1					6
2018	14	21	1				36
2017	81	77	12	1	1		172
2016	61	94	31	2	1		189
2015	48	73	17	2		1	141
2014	15	24	6		7		52
2013	1	2					3
2012		4	1				5
2011		1					1
Total	225	297	68	5	9	1	605

표 4. 단일 CPC 코드를 가지는 사물인터넷 관련 특허의 CPC 코드 현황
Table 4. The status of CPC Code of IoT Patents with Single CPC Code

No.	CPC Code Subclass	No. of Patent (Freq.)	No.	CPC Code Subclass	No. of Patent (Freq.)
1	A01G	1	12	G08B	5
2	A47J	1	13	G09B	2
3	A61B	4	14	G16H	2
4	A61F	1	15	H01R	2
5	A61L	2	16	H04B	3
6	B25J	1	17	H04H	2
7	B66B	1	18	H04L	157
8	E06B	1	19	H04N	4
9	G05B	1	20	H04W	26
10	G06F	5	21	H05B	1
11	G06Q	13	22	H05K	1

표 5에 복합 CPC 코드를 가지는 사물인터넷 관련 특허 369건에서 지지도(support)가 상위 20위에 해당하는 기술 연관성 규칙을 정리하였다. CPC 코드간 규칙은 총 3,111개로 계산되었다. CPC 코드간 연관성 규칙은 데이터 마이닝을 위한 오픈 소스인 R을 이용하였다.

표 5에 보듯이 지지도가 가장 높은 규칙인 [H04W→H04L]과 [H04L→H04W]의 지지도는 18.2%로 같지만, 신뢰도(confidence)는 66.3%와 35.4%로 다르다. 신뢰도(A→B)는 A가 발생한 조건에서 B가 발생할 확률 P(B|A)로 [H04W→H04L]의 신뢰도가 상대적으로 높다는

것은 H04W이 H04L 보다 선행 기술라는 것이다. [H04W→H04L]의 향상도는 1보다 크므로 H04W와 H04L은 상호 보완 관계에 해당한다.

표 5. 지지도 상위 20위 사물인터넷 관련 특허의 기술 연관성 규칙 현황
Table 5. The Status of Technology Association Rules of Top 20 IoT Patents with High Support

No.	Rules (A→B)	Support	Confidence	Lift
1	H04W → H04L	18.2%	66.3%	1.3
2	H04L → H04W	18.2%	35.4%	1.3
3	G06Q → H04L	17.3%	40.5%	0.8
4	H04L → G06Q	17.3%	33.9%	0.8
5	H04W → G06Q	9.8%	35.6%	0.8
6	G06Q → H04W	9.8%	22.8%	0.8
7	G06F → H04L	7.9%	49.2%	1.0
8	H04L → G06F	7.9%	15.3%	1.0
9	G06F → G06Q	6.2%	39.0%	0.9
10	G06Q → G06F	6.2%	14.6%	0.9
11	G05B → G06Q	5.4%	83.3%	1.9
12	G06Q → G05B	5.4%	12.7%	1.9
13	A01G → G06Q	4.3%	94.1%	2.2
14	G06Q → A01G	4.3%	10.1%	2.2
15	G06K → H04L	4.1%	51.7%	1.0
16	{G06Q, H04W} → {H04L}	4.1%	41.7%	0.8
17	{H04L, H04W} → {G06Q}	4.1%	22.4%	0.5
18	{G06Q, H04L} → {H04W}	4.1%	23.4%	0.9
19	G06K → G06Q	3.5%	44.8%	1.0
20	H04N → G06Q	3.0%	45.8%	1.1

두 번째로 지지도가 높은 규칙은 [G06Q→H04L]과 [H04L→G06Q]로 지지도는 모두 17.3%이고, 신뢰도는 40.5%와 33.9%로 다르다. [G06Q→H04L]의 향상도는 1보다 작으므로 전자상거래를 해당하는 G06Q[‘관리용, 상업용, 금융용, 경영용, 감독용 또는 예측용으로 특히 적합한 데이터 처리 시스템 또는 방법’]와 H04L[‘디지털 정보의 전송’]은 상호 대체 관계에 해당한다.

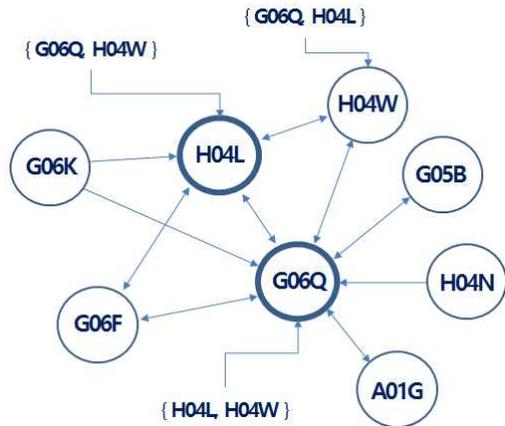


그림 3. 사물인터넷 관련 특허의 CPC 코드간 상호 연결망 분석 (표 5의 결과를 재구성함)
 Fig. 3. Analysis of CPC Code Interconnections (Reconstructed the results in Table 5)

표 6. 단계(연도)별 지지도 상위 5위의 기술 연관성 규칙
 Table 6. Top 5 Support Technology Association Rules by Stage(Year)

No.	Rules (A↔B)	Support(%)		
		[1] 2012~2015	[2] 2012~2016	[3] 2012~2019(July)
1	H04W↔H04L	20.5	21.9	18.2
2	G06F↔H04L	11.1	9.3	7.9
3	G06Q↔H04L	9.4	12.2	17.3
4	H04W↔G06Q	7.7	10.1	9.8
5	G06F↔G06Q	7.7	8.0	6.2

그림 3은 표 5에서 제시된 CPC 코드간 기술 연관성 규칙을 상호 연결망으로 해석하여 시각화하였다. 그림 3에서 확인할 수 있듯이 G06Q는 총 8개(H04L, H04W, H04N, A01G, G05B, G06F, G06K, {H04L, H04W})의 CPC 코드와 연결되어 있고, H04L은 총 5개(H04W, G06Q, G06F, G06K, {G06Q, H04W})의 CPC 코드와 연결되어 있다. 사물인터넷 관련 특허의 핵심 CPC 코드는 G06Q와 H04L임을 알 수 있다.

표 6에 단계(연도)별 지지도 상위 5위의 기술 연관성 규칙을 정리하였다. 표 6에서 보듯이 [G06F↔H04L]는 매 단계별 지속 감소(1단계_11.1%→2단계_9.3%→3단계_7.9%)하고 있는 반면, [G06Q↔H04L]는 급속히 증가(1단계_9.4%→2단계_12.2%→3단계_17.3%)하고 있다.

5. 결론

본 연구는 4차 산업혁명 ICT 기반기술의 핵심인 사물인터넷 특허의 CPC 코드에 의한 기술 연관성 규칙 분석에 관한 것이다. 2019년 7월까지 특허청에 출원된 사물인터넷 관련 특허 605건 중 복합 CPC 코드를 가지는 369건을 대상으로 서브클래스(Subclass) 수준까지 분석하였다.

기술 연관성 규칙 분석 결과 지지도가 높은 규칙은 [H04W→H04L](18.2%), [H04L→H04W](18.2%), [G06Q→H04L](17.3%), [H04L→G06Q](17.3%), [H04W→G06Q](9.8%), [G06Q→H04W](9.8%), [G06F→H04L](7.9%), [H04L→G06F](7.9%), [G06F→G06Q](6.2%), [G06Q→G06F](6.2%) 순이다. 지지도가 가장 높은 규칙은 [H04W→H04L]과 [H04L→H04W]으로 지지도는 모두 18.2%이지만, 신뢰도는 66.3%와 35.4%로 다르다. [H04W→H04L]의 신뢰도가 [H04L→H04W] 보다 높다는 것은 H04W이 H04L 보다 선행 기술이라는 것이다.

단일 CPC 코드를 가지는 236건의 CPC 코드 현황을 정리한 결과, H04L(157건)과 H04W(26건)이 전체의 약 78%에 해당하지만, CPC 코드간 상호 연결망을 분석한 결과 기술 연관성 관련 핵심 CPC 코드는 G06Q와 H04L이다.

본 연구 결과를 활용하면 앞으로의 특허 경향을 예상해 볼 수 있다.

REFERENCES

[1] Jae-Shin Jo, The Convergency Analysis of International Patent Classification(IPC) and Research for Utilization of an Examination and a Trial, Korea Intellectual Property Society, Vol.38, pp.91-130, 2012
 [2] Pang-Ryong Kim and Sung-Hyun Hwang, A Study on the Projection of the IT-based Promising Technologies Utilizing Patent Database, The Journal

- of Korea Information and Communications Society, Vol.34, No.10C, pp.1021-1030, 2009
- [3] Jinho Choi, Heesu Kim, and Namgyu Im, Keyword Network Analysis for Technology Forecasting, The Journal of the Intelligence Information Systems, Vol.17, No.4, pp.227-240, 2011
- [4] Jaeruen Shim, Analysis of Technology Convergence of 'Internet of Things' Patents by IPC Code Analysis, The Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol.9, No.3, pp.266-272, 2016
- [5] Soo-Hyeon Chae and Jangwon Gim, A Study of CPC-based Technology Classification Analysis Model of Patents, Journal of The Korea Contents Association, Vol.18, No.10, pp.443-452, 2018
- [6] Jaeruen Shim, Analysis of Technology Convergence of 'Internet of Things' Patents in the Electronic Commerce by the CPC Code Technology, The Journal of Korea Institute of Information, Electronics, and Communication Technology, Vol.11, No.6, pp.678-683, 2018
- [7] Sunghae Jun, IPC Code Analysis of Patent Documents Using Association Rules and Maps-Patent Analysis of Database Technology, Database Theory and Application, Bio-Science and Bio-Technology, pp.21-30, 2011
- [8] KIPO(Korean Intellectual Property Office), https://www.kipo.go.kr/kpo/HtmlApp?c=33001&catmenu=m06_07_06/
- [9] Patent Informatics & Informetrics, CSIR-NISCAIR Biennial Report, 2014 <http://www.niscair.res.in/>
- [10] J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, Acorn Pub., 2015
- [11] Sunghae Jun, S.S. Park, and D.S. Jang, Patent Analysis and Technology Forecasting, Kyowoo Pub., 2014
- [12] KIPO, CPC Manual(Korean), 2014
- [13] KIPRIS (Korea Intellectual Property Rights Information Service), <http://kpat.kipris.or.kr/kpat/searchLogina.do?next=MainSearch/>
- [14] RStudio <https://www.rstudio.com/products/rstudio/download/>

저자약력

심재륜(Jaeruen Shim)

[중신회원]



- 1990년 2월 : 한국과학기술원 전기및전자공학과 공학사
- 1992년 2월 : 포항공과대학교 전자전기공학과 공학석사
- 1999년 8월 : 포항공과대학교 전자전기공학과 공학박사
- 2000년 3월 ~ 현재 : 부산외국어대학교 임베디드IT학부 교수

〈관심분야〉 RF, 무선통신, 사물인터넷, 기술창업