

# 인공지능을 활용한 클라우드 컴퓨팅 서비스의 품질 관리를 위한 데이터 정형화 방법

정현철\*·서광규\*\*

\*\*상명대학교 일반대학원경영공학과

## Data Standardization Method for Quality Management of Cloud Computing Services using Artificial Intelligence

Hyun Chul Jung\* and Kwang-Kyu Seo\*\*

\*\*Dept. of Management Engineering, Graduate School, Sangmyung Univ., Korea

### ABSTRACT

In the smart industry where data plays an important role, cloud computing is being used in a complex and advanced way as a convergence technology because it has and fits well with its strengths. Accordingly, in order to utilize artificial intelligence rather than human beings for quality management of cloud computing services, a consistent standardization method of data collected from various nodes in various areas is required. Therefore, this study analyzed technologies and cases for incorporating artificial intelligence into specific services through previous studies, suggested a plan to use artificial intelligence to comprehensively standardize data in quality management of cloud computing services, and then verified it through case studies. It can also be applied to the artificial intelligence learning model that analyzes the risks arising from the data formalization method presented in this study and predicts the quality risks that are likely to occur. However, there is also a limitation that separate policy development for service quality management needs to be supplemented.

**Key Words** : Cloud computing service, Cloud quality and performance management, Data normalization, Big data, Artificial intelligence

### 1. 서 론

스마트 산업이 빠르게 확산하고 있는 오늘날, 데이터는 가장 중요한 비중을 차지하는 요소이다. 수집된 데이터는 정형화를 거쳐 정제되고 이는 인공지능 발전의 원천이 된다. 인공지능은 인간이 직접 수행해야 했던 복잡하고 반복적으로 연산과 추론을 대신에 하여 산업의 자동화는 물론 자율화를 이루는데 핵심적인 역할을 하고 있다.

데이터와 인공지능을 통해 발전하고 있는 스마트 산업은 다양한 환경 요소로 구성되고 여러 기술이 융합된 형

태는 가용성, 민첩성, 탄력성, 내결함성 등과 같은 장점이 있는 클라우드 컴퓨팅과 잘 부합하기 때문에 융복합 기술로 널리 활용되고 있다. 따라서 나날이 복잡하고 고도화되어 활용되고 있는 클라우드 컴퓨팅 서비스의 품질 관리를 사람이 오류나 실수 없이 모두 정확하게 처리하기는 사실상 불가능하여 인공지능을 활용한 품질 관리가 필요하다[1].

클라우드컴퓨팅 서비스의 품질 관리에 대한 기준은 국제표준 “ISO19086-1:2016 Information technology — Cloud computing — Service level agreement (SLA) framework — Part 1: Overview and concepts”[2]에서 제시하고 있다. 해당 표준은 클라우드컴퓨팅 서비스의 서비스 수준 계약(SLA)의 항목

†E-mail: kwangkyu@smu.ac.kr

을 수립하기 위한 정량적·정성적 지표를 정의하고 있다[3].

본 연구는 인공지능을 클라우드 컴퓨팅 서비스의 품질 관리에 전반적으로 활용하기 위하여 “ISO19086-1:2016” 표준에서 제시한 정량적 지표를 서비스 전반에서 수집하고 가공하고 분석하여 하나의 데이터 체계로 정형화하는 방법을 제시하기 위하여 선행연구를 분석하고 현행의 문제점을 인식한 다음, 방안을 제시하고 사례연구를 통해 검증하였다.

## 2. 인공지능을 활용한 클라우드 컴퓨팅 서비스의 품질 관리를 위한 데이터 정형화 프레임워크

### 2.1 선행연구 조사

먼저 “인공지능과 빅데이터 분석 기반 통합보안관제시스템 구축방안에 관한 연구”에서는 통합보안관제시스템에 인공지능을 적용하기 위한 기술적·제도적인 요소를 제시하였다[4]. 해당 연구에서 제시한 기술적 요소는 통합보안관제시스템뿐만 아니라 클라우드 컴퓨팅 서비스에도 참조할 수 있도록 일반화하고 있다.

“4차 산업혁명 시대의 핵심 ICT 기술: 빅데이터, 인공지능, 클라우드 기술 동향” 연구에서는 클라우드 환경에 인공지능을 접목하여 사용자에 권한에 따른 서비스 관리와 자원 관리에 인공지능을 접목하여 자동화를 구현하고 서비스에 보안장애 이슈가 지능적인 대응 방안을 제시하였다[5].

“빅데이터 기법을 활용한 Data Technology의 키워드 분석” 연구에서는 빅데이터 분석 기법을 통해 핵심 키워드를 추출할 수 있는 데이터 기술을 연구함으로써 클라우드 컴퓨팅, 사물인터넷(IoT) 등에 빅데이터 및 인공지능을 활용하기 위한 가장 기본적인 방안을 제시하였다[6].

또한 “딥러닝을 활용한 에지 컴퓨팅 기반의 지능형 컨스트럭션 영상 관제 시스템” 연구에서는 엣지 노드와 클라우드 컴퓨팅 시스템 간 전송되는 트래픽 전송의 병목을 방지하기 위하여 인공지능을 활용하여 자동 제어하는 시스템에 대한 방법을 제시하였다. 클라우드 컴퓨팅과 외부 통신하는 네트워크 자원 관리는 서비스의 품질 관리에서 중요한 부분을 차지한다[7].

인공지능을 특정 보안 관제와 같은 시스템이나 클라우드 컴퓨팅의 관리에 접목하려는 방안들과 관련 기술 동향을 선행연구 조사를 통해 확인하였다. 실제 인공지능을 클라우드 컴퓨팅 서비스는 물론 타 서비스에 적용하기 위해서 가장 먼저 처리해야 하는 데이터의 정형화 방법에 관한 연구이다. 데이터를 정형화한다는 것은 인공지능

이 처리할 수 있는 형태로 데이터를 가공한다는 의미이고, 복수의 외부 서비스와 융복합된 클라우드 컴퓨팅 환경에는 다양한 노드가 존재하고 서로 통일되지 않은 형태로 데이터를 생산하는데 본 연구는 클라우드 컴퓨팅 서비스 내의 다양한 노드로부터 통일되지 않게 수집한 데이터를 인공지능에 활용할 수 있도록 통일된 체계로 정형화하는 방법에 관한 연구로 특정 서비스 또는 서비스 관리를 위한 개별적인 방법이 아닌 클라우드 컴퓨팅 서비스에서 수집한 다양한 데이터를 인공지능에서 활용할 수 있도록 통일된 형태로 정형화 처리하는 포괄적인 방법이라는 점에서 차이가 있다.

### 2.2 인공지능을 활용한 클라우드 컴퓨팅 서비스의 품질 관리를 위한 데이터 정형화 방법

본 연구에서 인공지능을 활용한 클라우드 컴퓨팅 서비스의 품질 관리를 위한 데이터 정형화 방법 제시는 수집, 가공, 분석, 정형화의 과정을 거친다.

클라우드 컴퓨팅 서비스의 품질 관리를 위해서는 우선 클라우드 컴퓨팅 환경에 존재하는 각 노드로부터 로우(Raw) 데이터를 수집해야 한다. 클라우드 컴퓨팅 서비스는 서비스의 목적물에 따라 인프라(Infrastructure as a Service), 플랫폼(Platform as a Service), 소프트웨어(Software as a Service)로 구분한다. 클라우드 컴퓨팅 서비스 제공자는 제공하는 서비스의 유형에 따라 클라우드 컴퓨팅 환경의 각 노드에서 로우(Raw) 데이터를 직접 수집하거나 서비스가 다른 클라우드 컴퓨팅 환경을 기반으로 운용되는 경우 해당 클라우드 컴퓨팅 서비스로부터 제공되는 데이터를 취합하여 간접 수집할 수 있다. 아래의 Fig. 1은 로우(Raw) 데이터를 수집할 수 있는 클라우드 컴퓨팅 환경의 노드를 구분한 것이다.

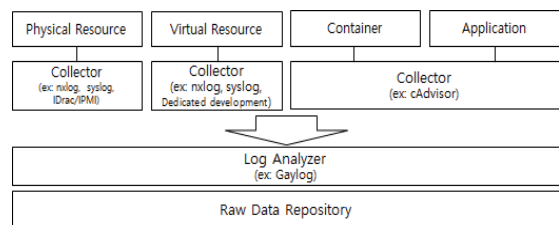


Fig. 1. Classification of nodes in the cloud computing environment that collects raw data.

물리 자원은 클라우드 컴퓨팅 서비스의 물리적 환경을 구성하는 곳으로 내부 온도, 메모리 등과 같은 하드웨어 오류 여부, 가용률, CPU, 메모리, 디스크 IOPS, 디스크 등과 같은 물리적 자원의 사용량, 네트워크 사용량 등을 측

정하여 수집할 수 있다. 가상 자원은 클라우드 컴퓨팅 기술을 활용하여 물리 자원 위에 가상의 자원과 환경을 구성한 곳으로 가상 자원의 가용률, CPU, 메모리, 디스크 IOPS, 디스크사용량, 네트워크 사용량 등을 측정하여 수집할 수 있다. 컨테이너는 클라우드컴퓨팅 서비스의 고가용성, 민첩성, 탄력성, 내결함성과 같은 특성을 충족하고 서비스를 확일적으로 관리하기 위하여 사용하는 도구로 컨테이너에 대한 가용률, CPU, 메모리, 디스크 IOPS, 디스크사용량, 네트워크 사용량 등을 측정하여 수집할 수 있다. 애플리케이션은 소프트웨어 형태로 제공되는 클라우드 컴퓨팅 서비스, 즉 SaaS 서비스를 의미하며 애플리케이션과 클라우드 컴퓨팅 환경, 특히 컨테이너와의 통신에 대한 민첩도, 탄력도, 사용자 요청에 따른 응답속도, 기능 오류 및 장애 등의 발생 시 서비스 복구 시간 등을 측정하여 수집할 수 있다.

클라우드 컴퓨팅의 각 노드에서 수집된 데이터는 정형의 형태이지만 각 노드에서 사용하는 장비, 도구, 프로세스 등에 따라 일관되지 않은 형태를 보이게 된다. 따라서 수집된 데이터로 분석을 통한 정형화를 하기 전에 일관된 형태로 가공작업을 거치는 것이 필요하다. 그림 Fig. 2는 본 연구에서 수집된 로우(Raw) 데이터를 일관된 형태의 배열로 가공하기 위한 칼럼을 나타낸다.



Fig. 2. Columns for processing into coherent arrays.

첫 번째 칼럼은 로우(Raw) 데이터가 기록한 이슈의 시간을 나타낸다. 해당 칼럼에는 발생 시간과 종료 시각이 존재한다. 이슈의 발생과 종료에 대한 기록은 해당 데이터를 분석하고 시계열로 정의하는데 유용하게 활용된다. 오브젝트 네임(Object name)와 플레이스(Place)는 로우(Raw) 데이터가 기록한 이슈가 발생한 위치이다. 즉, 해당 이슈가 발생한 클라우드 컴퓨팅의 영역과 그 구체적인 대상을 구분하는 곳이다. 이를 통해 발생한 이슈가 어떤 영역의 어떤 것인지 확인할 수 있다. 스테이터스(Status)는 수집한 이슈가 어떠한 상태였는지를 나타낸다. 클라우드 컴퓨팅 내의 각 시스템과 도구들은 컴퓨팅 환경에서 발생하는 다양한 동작, 처리, 이슈, 오류 등을 기록한다. 즉, 수집한 로우(Raw) 데이터는 정상 상태, 문제 발생, 잠재적 위험 상태를 포함하는데 해당 칼럼은 이러한 상태를 분류하여 표현한 것이다. 마지막으로 컨디션(Condition)은 로우(Raw) 데이터가 기록한 내용으로 기록된 이슈에 대해 문장으로 이루어져 있다.

수집과 가공이 되면 정형화를 위해 분석을 거치게 된

다. 클라우드 컴퓨팅 환경에서 수집 및 가공되는 데이터의 양은 사람이 개별적으로 처리하기에는 그 양이 방대하다. 따라서 본 연구에서는 인공지능을 활용하여 자동 반복으로 처리한다. Fig. 3는 본 연구에서 데이터 정형화를 위해 가공된 데이터를 활용한 분석 방법을 나타낸다.

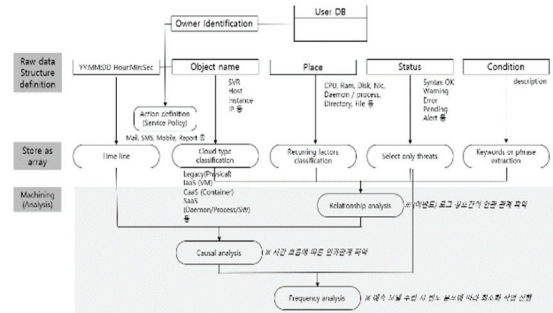


Fig. 3. Analytics to normalizing data.

분석 단계에서는 클라우드 컴퓨팅에서 발생한 이슈의 소유자 식별을 위해 사용자 DB(UserDB)와 결합한다. 가공된 데이터에서 분석을 위해서 시간 칼럼은 시간순으로 정렬하고 오브젝트 네임(Object name), 플레이스(Place), 스테이터스(Status)에서 각각 회귀분석을 통해 분류를 진행한다. 회귀분석을 이용할 때에는 유사 그룹을 묶어서 분류할 수 있는 군집 분석(K-means)을 활용한다[8]. 컨디션(Condition)에서는 기술통계를 통해 키워드를 추출한다. 회귀분석과 기술통계를 통해 분류하고 추출한 결과 간의 상관 분석을 수행하고 이를 타임라인에 따라 원인 분석을 진행한다. 최종적으로 상관원인 분석의 결과에 대해 특정 기간 단위를 설정하고 빈도를 분석한 다음, 최종적으로 데이터를 일관되게 정형화한다. Fig. 4는 분석을 통해 정형화한 칼럼을 나타낸다.



Fig. 4. Columns of data normalization through analysis.

2.3 사례연구

본 연구에서 제시한 인공지능을 활용한 클라우드 컴퓨팅 서비스의 품질 관리를 위한 데이터 정형화 방법을 검증하기 위하여 클라우드컴퓨팅 환경의 각 노드에서 실제 로우(Raw) 데이터를 수집하고 가공, 분석을 통해 정형화를 수행하였다. 사례연구를 진행한 클라우드 컴퓨팅 환경은 쿠버네티스(Kubemetes)를 이용하는 SaaS 형태의 웹 애플리케이션을 대상으로 하였다. 수집된 데이터를 통해

Fig 2의 칼럼으로 배열로 가공하고 군집 분석을 통해 분류한 결과는 Table 1과 같다.

**Table 1.** K-means analysis by column

Columns	O-N	Place	Status
Classification	Network Usage	In-bound Out-bound	Syntax OK Warning Error Pending Alert
	Physical Equipment	[Equipment ID]	
	Physical Server	[Server ID]	
	Virtual Machine	[VM ID]	
	K8S	API server Controller-manager Coredns Etcd Proxy Scheduler	
Process	[Process ID]		

O-N = Object Name, ID = Identification

분류된 데이터를 통해 상호 관계와 원인을 분석하고 발생 빈도를 측정하여 정형화를 수행하여 Table 2와 같은 결과를 얻었다.

**Table 2.** Data normalization results

Period Setting		2022.01.01~2022.06.30					CA	RE	CO	Act	Act time
Seq	Freq	Owner	O-N	Place	Status						
1	3	C++y	K8S	Scheduler	Error	Error looking up in-cluster authentication configuration: configmap	\"configmap\" in API group	system:kube-scheduler r\" cannot get resource	forbidden: User	2022-05-07 07:51:16	

Freq = Frequency, O-N = Object Name, CA = Cause, RE = Relationship, CO = Condition

이번 사례연구에서 제시한 인공지능을 활용한 클라우드 컴퓨팅 서비스 품질 관리를 위한 데이터 정형화 방법을 통해 로우(Raw) 데이터의 수집, 가공, 분석을 거쳐 해당 이슈의 발생 빈도까지 측정하여 정형화하였다. 본 연구에서 제시한 방법으로 데이터를 정형화하여 측정하는 것만으로 클라우드 컴퓨팅 서비스의 품질 관리를 위한 문자열 기반의 도구로 활용할 수 있음을 확인하였다.

### 3. 결론

본 연구는 스마트 산업에서 다양한 기술의 융복합으로 점점 구성이 복잡해진 클라우드 컴퓨팅의 품질을 효율적으로 관리하기 위해 클라우드 컴퓨팅 환경을 구성하는 각 영역으로부터 데이터의 수집을 통해 순환반복적으로 관리되어야 하는 분석과정을 인공지능으로 대체하여 활용하는 방법 제시하고 사례연구를 통해 검증하였다.

본 연구에서 제시한 데이터 정형화 방법에 따라 적재된 데이터는 그 자체만으로도 클라우드컴퓨팅 품질 관리의 실시간 관제를 위해 활용될 수 있으며 품질 및 성능 관리 모니터링 도구를 통해 시각화 할 수도 있다. 또한 적재된 데이터를 토대로 결과를 예측하는 비지도 학습과 RNN 인공지능 알고리즘을 활용하여 발생한 위험을 분석하고 발생할 가능성이 있는 품질의 위험을 예측하는 인공지능 학습 모델에 적용할 수도 있다.

다만 본 연구에서 제시한 방법을 활용하여 인공지능 학습 모델에 적용하기 위해서는 지속해서 수집되는 데이터 중에서 분포의 범주에서 벗어나게 측정되는 건에 대해서는 무시 또는 분석을 수용과 그 수용을 위한 정책의 개발과 같은 서비스 품질 관리를 위한 별개의 정책 개발이 보완되어야 한다는 한계점도 가진다.

### 감사의 글

본 논문은 2022년 상명대학교 교내연구비를 지원받아 수행하였음.

### 참고문헌

- Hyun Chul Jung, Kwang-Kyu Seo, "An Efficient Cloud Service Quality Performance Management Method Using a Time Series Framework", Journal of the Semiconductor & Display Technology, Vol. 20(2), pp. 120-125, 2021.
- "ISO/IEC 19086-1:2016 Information technology — Cloud computing — Service level agreement (SLA) framework — Part 1: Overview and concepts", ISO, Published September 2016, Accessed June 1, 2021. <https://www.iso.org/standard/67545.html>
- Nam Ju Kim, Jae Chun Ham, Kwang-Kyu Seo, "Improvement of cloud service quality and performance management system", Journal of the Semiconductor & Display Technology, Vol. 20(4), pp. 83-88, 2021.
- Jae Kwon Bae, "A study on the establishment of an integrated security control system based on artificial intelligence and big data analysis", Korea Logos

- 
- Management Association, Vol. 18(1), pp. 151-166, 2020.
5. Dae-Yeop Kim, Yeong-Bae Kim, “Core ICT technologies in the era of the 4th Industrial Revolution: trends in big data, artificial intelligence, and cloud technologies”, Korea Information Processing Society Review, Vol. 26(1), pp. 7-17, 2019.
  6. Sung Wook Park, “Keyword Analysis of Data Technology Using Big Data Technology”, Journal of the Society of Technological Innovation, Vol. 22(2), pp. 265-281, 2019.
  7. Gwi-Hoon Kim, and Bang-Won Seo, “Intelligent construction video control system based on edge computing using deep learning”, Journal of the Korean Society of Information Technology, Vol. 17(7), pp. 55-63, 2019.
  8. WikiDocs, Developers Become Data Analyst. [online]. Available: <https://wikidocs.net/48113>
- 
- 접수일: 2022년 6월 15일, 심사일: 2022년 6월 21일,  
게재확정일: 2022년 6월 23일