공항 수하물 처리 시스템 이벤트 로그의 프로세스 관점 분석 방안 연구

A Process Perspective Event-log Analysis Method for Airport BHS (Baggage Handling System)

박신념 · 송민석

포항공과대학교 산업경영공학과

요 약

급증하는 항공 여객의 성장세에 맞춰 여객 터미널의 규모가 대형화됨에 따라 출발, 도착, 환승 여객들이 소지한 수하물을 최단 시간 내에 신속, 정확하게 처리할 수 있는 다양한 데이터 기술들이 접목된 첨단수하물 처리시스템(Baggage Handling System; 이하 BHS)이 필수 요소가 되었다. 따라서 본 연구에서는 공항 수하물 처리시스템 운영의 고도화를 위해, 프로세스 관점의 데이터 분석 방법론을 통한 국내 공항의수하물 처리능력 분석 방법을 소개하고, 이벤트 로그 기반 주요 지점의 정확한 부하예측 방법을 제시하여, 자원의 선제적 배치 및 flight-carrocel 스케줄링 최적화 문제 등 향후 첨단화된 BHS 운영 전략으로 이어질수 있도록 한다. 분석에 사용된 데이터는 공공데이터 포털에서 얻을 수 있는 '전국 공항 수송실적', '항공기운항 정보' API를 활용하였다. 국내 공항 BHS 시뮬레이션 모델에 해당 방법을 적용한 결과, 높은 수준의 예측성능을 확인 할 수 있었다.

■ 중심어 : 수하물 처리 시스템, 프로세스 마이닝, 물류 시뮬레이션, 기계 학습

Abstract

As the size of the airport terminal grows in line with the rapid growth of aviation passengers, the advanced baggage handling system that combines various data technologies has become an essential element in order to handle the baggage carried by passengers swiftly and accurately. Therefore, this study introduces the method of analyzing the baggage handling capacity of domestic airports through the latest data analysis methodology from the process point of view to advance the operation of the airport BHS and the main points based on event log data. By presenting an accurate load prediction method, it can lead to advanced BHS operation strategies in the future, such as the preemptive arrangement of resources and optimization of flight-carrousel scheduling. The data used in the analysis utilized the APIs that can be obtained by searching for "Korea Airports Corporation" in the public data portal. As a result of applying the method to the domestic airport BHS simulation model, it was possible to confirm a high level of predictive performance.

■ Keyword: Baggage Handling System(BHS), Process Mining, Logistics Simulation, Machine Learning

Ⅰ. 서 론

국제공항운영협의회(ACI)에 따르면 최근 10 년간 아시아・태평양 지역의 연평균 항공이용 여객 증가율은 7.3% 이상이며 항공기 운항 횟수는 연평균 5.4%의 성장을 하고 있다. 또 앞으로 전 세계 항공 여객은 장기적으로 '15년 72억 명에서 2040년 240억명으로 3배 이상 증가 할 것으로 전망하고 있으며 중동과 아시아・태평양지역에서 이를 주도할 것으로 나타나고 있다[1]. 따라서 이러한 항공 여객의 성장세에 맞춰여객 터미널의 규모가 대형화됨에 따라 출발,도착,환승 여객들이 소지한 수하물을 최단 시간 내에 신속,정확하게 처리할 수 있는 다양한데이터 기술들이 접목된 첨단 수하물 처리시스템(Baggage Handling System; 이하 BHS)이 필수요소가 되었다.

수하물 처리시스템(BHS)은 공항 운영과 고객서비스의 성패를 좌우하는 핵심 시스템이다. 국내 K 항공사의 경우 2010년 지연 운항 원인의 대부분이 수하물 관련 문제였으며, 인천공항은 2016년 1월3일 수하물 처리시스템 문제로 159대의 항공편이 지연되고 5,200개의 수하물을 제때 싣지 못하는 수하물 대란을 겪은 바 있다 [2]. 해외의 경우, 덴버, 스키폴, 히드로, 하네다. 시드니, 샬드골 공항에서도 BHS 문제로 공항 운영에 큰 혼란이 발생한 적이 있으며, 세계적으로 항공사, 공항당국 등 항공업계에서는 잘못처리되는 수하물로 인하여 여객 불편은 물론 막대한 경제적 손실로 골치를 앓고 있다 [3].

따라서 본 연구는 공항 수하물 처리시스템 운영의 고도화를 위해, 프로세스 관점의 데이터 분석 방법론을 통한 국내 공항의 수하물 처리능력 분석 방법을 소개하고, 이벤트 로그(eventlog) 데이터 기반 주요 지점(i.e., departure sorter carrocel)의 정확한 부하예측 방법을 제시하여, 자원의 선제적 배치 및 flight-carrocel 스케줄링

최적화 문제 [4] 등 향후 첨단화된 BHS 운영 전략으로 이어질 수 있도록 하는 것에 목적을 둔다.

논문의 구성은 다음과 같다. 제 2장에서는 공항 BHS와 데이터 기술의 접목과 관련한 선행연구에 대해 소개한다. 제 3장에서는 본 연구에사용된 구체적인 방법론을 절차에 따라 설명한다. 제 4장에서는 프로세스 마이닝 분석과 BHS부하 예측 분석 결과를 다루며 제 5장은 연구 결론과 시사점으로 정리되었다.

Ⅱ. 관련 연구

공항 BHS는 전 세계 대형공항을 중심으로 다양한 분과에서 매우 활발하게 진행 중인 연구주제이지만, 초대형 BHS인 인천공항을 제외한국내공항 BHS 연구는 상대적으로 그 수가 적으며, 용량분석, 안정적 시설 운영을 위한 연구가

〈표 1〉관련 연구 요약

저 자	내 용		
V Denisov	프로세스 마이닝 활용한 출발 BHS 내		
(2019)	X-ray 부하 Monitoring, 예측 분석		
M Frey	출발 BHS 의 Flight-Carrocel Scheduling		
(2017)	최적화		
Ziyi Li	통계적 고객 위탁 수하물 분석을 통한 공		
(2017)	항별 최적 Check-in counter 개수 결정		
이예원	무선 주파수 RFID 를 이용한 개인화 서비		
(2016)	스 모형 제시		
도기영	시뮬레이션을 활용한 출발 BHS 능력 분석		
(2015)			
J Hansen	Bag-tag log data 기반 프로세스 마이닝을		
(2014)	활용한 출발 BHS 물류 정체 원인 분석		
이재진	키오스크 도입에 따른 개선비용 시뮬레이		
(2013)	션 분석		
정종욱	모바일 기술과 RFID 기술을 적용한 RHS		
(2012)	모델 제시		
박원진	이디버르 토취 IT 거요비사 타게 여기		
(2012)	인터뷰를 통한 IT 적용분야 탐색 연구		

대부분이었다. 따라서 이벤트 로그 분석 기반 BHS 운영 효율화를 주제로 하는 본 연구와 밀접하게 관련된 국내외 선행 연구 요약은 표 1 과 같다 [5].

문헌 연구를 바탕으로 수행된 본 연구의 목표는, 최근 네덜란드 암스테르담 공항의 BHS 부하 분석 사례로 제시된 Vadim densiov(2019)의' 프로세스 마이닝 기반 performance spectrum 분석'방법론을 도기영(2015)의 연구내용에서 확인 할 수 있는 정보를 바탕으로 국내공항 BHS에 알맞게 적용하는 방법을 제시하는 것이다.

Ⅲ. 연구 내용

3.1 개요

본 연구의 범위는 국내 19개 공항, 국내선 대 상 1997~2018년으로 분석 대상 범위를 한정하 였다. 또한, 지연 발생에 영향을 미치는'출발 BHS '시스템을 대상 범위로 한다. 본 연구의 방 법 및 절차 개요는 그림 1로 나타내었다. 첫째, 공공데이터 포털 API, 한국공항공사에서 제공 되는 데이터를 바탕으로 연구 대상이 되는 공항 을 탐색적 자료 분석(Exploratory Data Analysis) 방법으로 선정한다. 선정 기준은 여객수, 운항 수, 화물수, 국내선 노선 수 등이 있다. 둘째, 제 공되는 데이터에는 BHS 관련 정보가 부재하므 로, 선정된 대상 공항의 이용 여객 수 정보와 BHS 시스템 세부 정보를 바탕으로 BHS 시뮬레 이션 모델을 생성한다. 시뮬레이션 로그를 이용 해 이후 BHS 분석을 진행한다. 셋째, 시뮬레이 션 로그를 기반으로 프로세스 마이닝을 통한 프 로세스 모델 도출, 병목 지점 탐색, 성과 스펙트 럼 분석(Performance Spectrum Analysis) [6], 이 후 예측 부하 분석을 위한 feature selection 등을 수행한다. 넷째, 프로세스 마이닝을 통해 얻어진 feature selection 정보 를 바탕으로 BHS의 주요 지점(i.e., departure sorter carrocel)의 부하예측을 다양한 기계학습(Machine Learning) 알고리즘으 로 수행하고 성능을 확인한다.



〈그림 1〉연구 방법 및 절차

3.2 분석 대상 공항 선정

탐색적 자료 분석(Exploratory Data Analysis) 방법론은 수집된 데이터에 대한 기초 이해 과정 으로, 데이터 애널리틱스 프로젝트의 최종 목적 달성을 위해 반드시 수행되는 단계이다. 대표적 인 방법으로 기술 통계량(평균, 중앙값, 최대값, 최소값, 공분산 등) 산출, 시각화, 이상점 발견 등이 있으며, 본 연구에서는 '분석 대상공항 선 정'과 '공항 BHS 시뮬레이션 모델 생성'의 입력 값을 산출하기 위한 과정으로 수행된다. 수집된 항공 데이터를 분석한 결과, 국내선 기준으로 항공 이용객 수 TOP3 공항은 차례로 제주공항, 김포공항, 김해공항 순으로 나타났다. 제주공항 은 2009년(1,238만명/년)을 기점으로 김포공항 이용객 수를 추월하였다. 김포⇔제주 노선은 100 회/월 이상, 김포⇔김해, 김해⇔제주 노선은 30~ 100회/월, 나머지 노선은 30회/월 미만인 것으로 나타났다. 또한 2017년 8월 2,884,951명/월로 최고치를 기록하였으며 2011년 2월 1,393,811 명/월로 최저치를 기록한 것을 확인할 수 있 다. 따라서, 꾸준한 항공 수요 증가에 따라 BHS 부하가 예상되는 공항은 제주국제공항 으로, 해당 공항을 분석 대상 공항으로 선정 하였다.

3.3 BHS 시뮬레이션 모델 생성

분석 대상 공항을 선정한 후 BHS 이벤트 로그 정보를 얻기 위한 시뮬레이션 모델을 생성하였다. 실제 제주공항 BHS 시설 정보는 대외비자료이기 때문에, 도기영(2015) [7] 에서 연구한 김포공항 시뮬레이션 정보를 바탕으로 BHS를설계하였으며, 입력값은 제 2장 2절에서 분석한제주공항 수요 및 공항 정보값을 반영하였다.국내선 출발 1호기 체크인(check-in) 카운터 4개운영 시 상황을 가정하였으며, 물류 시뮬레이션전용 패키지인 AnyLogic(8.5 university ver.) [8]을 사용하여 10일간 시뮬레이션 결과, 총 2,249개 수하물의 29,445개 이벤트 로그를 획득하였다.



〈그림 2〉BHS 시뮬레이션 모델

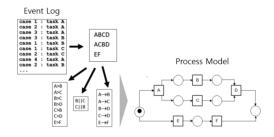
3.4 프로세스 마이닝

프로세스 마이닝은 기업내 다양한 정보 시스템에 기록된 이벤트 로그에 기반하여 업무 및 공정 프로세스 모델을 자동으로 도출하고 개선하는 일련의 방법론을 말한다 [9]. 유럽 및 네덜란드 아인트 호벤 공대를 중심으로 발전해온 프로세스 마이닝 기술은 현재 오픈소스 툴의 (ProM) 배포, 프로세스 마이닝 메니페스토(manifesto) 제정 [10-11] 을 통한 연구 표준 가이드라인 제시 등으로 헬스케어나 서비스 산업 등 여러 도메인에서 빅데이터 및 IoT(사물인터넷) 기술의 발전과 동시에, 국내외 관련 연구자들이 증가하는 추세이다. 최근에는 항공 정보 시스템

의 고도화에 따라, 공항 이용 프로세스, BHS 등 공항 적용 사례가 점차 많아지고 있다. 본 연구에서는 BHS 처리능력 분석, 병목 지점 탐색, feature selection을 위해 오픈소스 ProM(6.9), Disco(2.2)를 사용하였다.

3.4.1 프로세스 모델 도출

이벤트 로그롤 통해 프로세스 모델을 자동으로 도출하는 방법은 프로세스 마이닝에서 제공하는 alpha mining, genetic mining, heuristic mining 등 다양한 miner에 의해 수행된다. 이 중에서도 가장 기본 원리가 되는 ' α algorithm' 방법 [12] 을 간단한 도식으로 나타내면 그림 3과 같다.

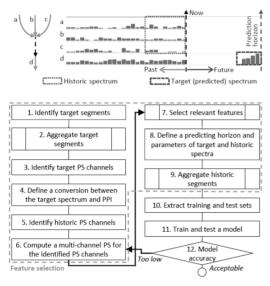


〈그림 3〉 프로세스 모델 도출 절차 예시 [12]

3.4.2 성과 스펙트럼 분석

일반적으로 프로세스 마이닝 기술은 하나의 객체 내(inner-case) 능력을 주로 분석하는 방법으로 사용되어왔다. 하지만 최근 프로세스 성과스펙트럼 분석(Performance Spectrum Analysis, 이하 PSA) 방법론이 제안되면서 [13], 다중 객체가 상호작용할 경우, 객체간(inter-case) 능력 분석이 가능해졌다. 해당 방법론은 유럽의 핵심허브인 암스테르담 공항의 BHS 부하예측 문제에서 well-working 하는 것으로 보고되었다 [6]. 그림 4에서 나타난 바와 같이 본 연구에서 다루는 문제도 각기 다른 a,b,c (i.e., check-in counter)수하물이 d(i.e., X-ray)로 물량이 합쳐져서 부하능력이 계산되는 경우이므로, 미래 시점에 대한

수하물 체류 부하 예측은 PSA 절차에 따라 진행할 수 있다.



〈그림 4〉Inter-case 문제의 PSA 방법론 [6]

3.5 기계 학습

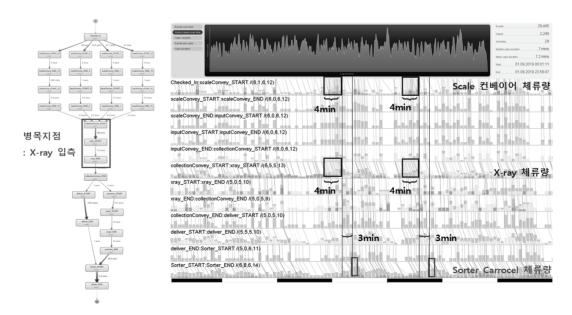
기계학습(Machine Learning)이란 컴퓨터가 주 어진 데이터를 학습하여 새로운 문제의 결과를 예측하는 방법론을 칭한다 [14]. 대표적으로 지도 학습, 비지도 학습, 강화학습이 있다. 본 연구에서는 BHS 부하예측에 지도학습 방법론으로 회귀 분석(LR, MLP, Random Forest) 또는 앙상블 학습(XGBoost) [15-17]을 적용하여 예측성능을 비교 하였다.

Ⅳ. 결 과

4.1 프로세스 마이닝 결과

이벤트 로그 정보를 바탕으로 공항 BHS 프로 세스 마이닝 분석 결과를 요약하면 그림 5와 같 다. 수하물은 check-in counter > scale 컨베이어 > input 컨베이어 > collection 컨베이어 > X-ray (AT) > 이송 컨베이어 > sorter carrocel 순으로 carrocel 1회전까지 평균 7.2분이 소요되었다.

이중에서 X-ray 입측은 평균 98초간 대기하는 주요 병목 지점으로 나타났다. PSA 결과에서 대체적으로 sorter carrocel 부하에 약 3분전 Scale 컨베이어, X-ray 부하가 밀접한 영향을 미치는 것으로 확인되었다. 일반적으로 BHS에서 수하



〈그림 5〉BHS 프로세스 모델 도출 및 PSA 결과

물 검색을 담당하는 X-ray는 주요 병목 지점이 다 [7]. 따라서 대부분의 대형 BHS에서 X-ray (EDS급)가 위치한 주요 sorter carrocel의 부하예 측 관리는 전체 BHS 부하 능력을 컨트롤 하는 매우 중요한 문제이다 [6]. 하지만 본 연구에서 다루는 시스템은 X-ray(AT급) 입측까지 이동 소 요시간이 상대적으로 짧고 물류 traffic이 적기 때문에 X-ray 부하 대신, 현재 수동 sorter 방식 으로 유영중인 carrocel의 자원배치 효율화 및 flight-carrocel 스케줄링 최적화 문제 [4] 등 향후 첨단화된 BHS 운영에 양질의 입력 정보 제공을 목표로 sorter carrocel 부하를 예측하였다. PSA feature selection 결과를 바탕으로 선택된 feature(X)는 4분전~현재 시점의 scale 컨베이어, X-ray의 수하물 체류량, feature(Y)는 3분뒤 시 점의 1분간 sorter carrocel 수하물 예측 체류량 이다.

	Activity	Date	Time
1	Checked_In	01.09.2019	09:28:16
2	scaleConvey_START_10	01.09.2019	09:28:50
3	scaleConvey_END_10	01.09.2019	09:28:55
4	inputConvey_START_10	01.09.2019	09:28:56
5	inputConvey_END_10	01.09.2019	09:28:59
6	collectionConvey_START	01.09.2019	09:29:00
7	xray_START	01.09.2019	09:30:34
8	xray_END	01.09.2019	09:30:41
9	collectionConvey_END	01.09.2019	09:30:54
10	deliver_START	01.09.2019	09:30:55
11	deliver_END	01.09.2019	09:32:41
12	Sorter_START	01.09.2019	09:32:42
13	Sorter_END	01.09.2019	09:35:05

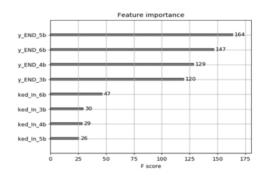
〈그림 6〉 수하물 이벤트 로그

4.2 BHS 부하 예측 결과

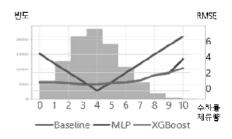
본 연구에서는 BHS 부하예측 분석에 회귀분석(Regression Analysis) 기법을 사용하였다. 회귀분석에서 key step으로 수행하는 feature selection 과정은 3분뒤 sorter carrocel 수하물 체류량을 예측하기 위한 모델에 4장 1절 결과를 사용하였다. baseline approach는 feature selection을 수행하지 않은, 현시점 scale 컨베이어 체류량을 X인자로 설정하고 3분뒤 sorter carrocel 체

류량을 Y인자로 설정한 선형 회귀분석 모델이다. 그 외 approach는 프로세스마이닝 기반 feature 정보 + 선형 회귀/랜덤 포레스트(random forest)/다중 레이어 퍼셉트론(multi-layered perceptron)/XGBoost의 학습 알고리즘 비교를 위한모델이다.

그림 7의 feature 중요도(f-score)는 20이상 준수한 수준으로 나타났으며, 그림 8에 나타난 결과에 따르면 baseline 대비 RMSE(Root Mean Squared Error)가 47% 향상된 결과를 확인할 수있다. 수하물 범위(0~10개)의 오차범위는 (1~2개) 수준으로 비교적 정확한 예측성능을 보인다. 특히 대표적인 앙상블 배깅(ensemble bag



〈그림 7〉BHS 부하 예측 모델 feature 중요도



모델	Adj.R ²	RMSE	MAPE
Baseline	0.02	1.70	48.1
Linear Regression	0.71	0.90	22.3
Random Forest	0.68	0.93	22.5
MLP	0.73	0.87	21.3
XGBoost	0.73	0.87	21.2

〈그림 8〉BHS 부하 예측 모델 성능

ging) 기법으로 알려진 XGBoost 경우 [16], 수하물 체류량이 많은 상황에서도 가장 좋은 성능을 나타내는 것으로 확인되었다.

Ⅴ. 결 론

본 연구에서는 항공 데이터를 활용하여 프로 세스 마이닝 방법론 기반 국내공항 BHS 처리능력 및 부하를 예측하는 방법을 소개하였다. 해당 접근법은 물류 복잡성이 가중된 대형 BHS에서도 충분히 확장 적용이 가능한 방법으로, BHS문제지점 및 부하 지점을 탐지하고 최적의 자원분배 및 관련 조치를 사전에 가능하게 한다. 현분석에서는 시뮬레이션 데이터를 활용하였지만향후 RFID(Radio-frequency identification) 기반수하물 추적 시스템 [17] 등 BHS 고도화를 통한양질의 이벤트 로그 정보 수집이 가능해진다면더욱 활용도가 높은 분석 방법이 될 것으로 기대된다.

참고문헌

- [1] ACI World Airport Traffic Forecasts 2016-2040.
- [2] 이재진, *국내선 공항 승객 및 수하물 처리 프로세 스 개선 방안 연구*, 한국항공대학교 학위논문(박 사), 2012.
- [3] SITA THE BAGGAGE REPORT, 2018.
- [4] M Frey et al., "Column Generation for Outbound Baggage Handling at Airports", Transportation Science 51(4) 1031-1386, 2017.
- [5] 김종서, 수하물 처리 시스템(BHS)의 서비스 품질 이 공공기관 조직성과에 미치는 영향, 한국항공 대학교 학위논문(박사), 2019.
- [6] Vadim Denisov et al., "Predictive Performance Monitoring of Material Handling Systems Using

- the Performance Spectrum", ICPM Proceedings, 2019.
- [7] 도기영, 國內空港 수하물처리시스템의 問題點 및 改善方案: K, G, J공항 출발 수하물처리시스템 중심으로, 중앙대학교 학위논문(석사), 2015.
- [8] Zhang, Yongan, Ying Wang, and Long Wu. "Research on demand-driven leagile supply chain operation model: a simulation based on anylogic in system engineering." Systems Engineering Procedia 3: 249-258, 2012.
- [9] Van Der Aalst, Wil. "Process mining: Overview and opportunities." ACM Transactions on Management Information Systems (TMIS) 3.2: 1-17, 2012.
- [10] Van Der Aalst, Wil, et al. "Process mining manifesto." International Conference on Business Process Management. Springer, Berlin, Heidelberg, 2011.
- [11] Van Dongen, Boudewijn F., et al. "The ProM framework: A new era in process mining tool support." International conference on application and theory of petri nets. Springer, Berlin, Heidelberg, 2005.
- [12] W. van der Aalst, T. Weijters and L. Maruster, "Workflow mining: discovering process models from event logs," in IEEE Transactions on Knowledge and Data Engineering, vol. 16, no. 9, pp. 1128-1142, Sept. 2004.
- [13] Vadim Denisov et al., "Unbiased, Fine-Grained Description of Processes Performance from Event Data", BPM Proceeding, 2018.
- [14] Michie, Donald, David J. Spiegelhalter, and C.C. Taylor. "Machine learning." Neural and Statistical Classification 13.1994: 1-298, 1994.
- [15] Liaw, Andy, and Matthew Wiener. "Classification and regression by randomForest." R news 2.3: 18-22., 2002.

- 1 2 2
- [16] Chen, Tianqi, and Carlos Guestrin. "Xgboost: A scalable tree boosting system." Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. 2016.
- [17] Gardner, Matt W., and S. R. Dorling. "Artificial neural networks (the multilayer perceptron)—a review of applications in the atmospheric sciences." Atmospheric environment 32.14-15: 2627-2636, 1998.
- [18] Tsaur et al. "The evaluation of airline service quality by fuzzy MCDM", Tourism Management 23(2):107-115, 2002.
- [19] W. van der Aalst, "Process Mining: Discovery, Conformance and Enhancement of Business Processes", Springer Verlag, Berlin (ISBN 978-3-642-19344-6).
- [20] Zhang, Ting, Yuanxin Ouyang, and Yang He. "Traceable air baggage handling system based on RFID tags in the airport." Journal of Theoretical and Applied Electronic Commerce Research 3.1: 106-115, 2008.

저 자 소 개



박 신 념(Shin-nyum Park)

- ·2017년 : 한양대학교 산업공학 과(공학사)
- ·2019년~현재 : 포항공과대학 교 산업경영공학과(공학석사)
- ·관심분야 : 프로세스 마이닝, 데이터 마이닝, 디지털 트윈



송 민 석(Minseok Song)

- ·1998년 : 포항공과대학교 산 업경영공학과(공학사)
- ·2000년 : 포항공과대학교 산 업경영공학과(공학석사)
- ·2005년 : 포항공과대학교 산 업경영공학과(공학박사)
- ·2010년~2016년 : 울산과학기술원(UNIST) 경영 공학부 조교수, 부교수
- ·2016년~현재 : 포항공과대학교 산업경영공학 과 부교수
- ·관심분야: 비즈니스 애널리틱스, 비즈니스 정보 시스템, 프로세스 마이닝, 블록체인