

## 클러스터링 기법을 활용한 출발 여객 체류 시간 분석

# Analysis of Departing Passengers' Dwell Time using Clustering Techniques

안덕배·김휘양·백호종\*

한국항공대학교 항공교통물류학부

Deok-bae An · Hui-yang Kim · Ho-jong Baik\*

School of Air Transport and Logistics, Korea Aerospace University, Gyeonggi-do 10540, Korea

### [요 약]

본 연구는 실제 공항에서 수집되는 여객 데이터를 활용하여 공항 내 여객의 체류 시간을 분석한 연구이다. 여객의 체류 시간은 공항 터미널 설계, 공항의 수익성에 영향을 주어 중요한 여객 특성으로 간주되어 왔지만 실제 여객 데이터 수집의 어려움으로 그에 대한 분석이나 실시간 공항 운영에 활용하기가 어려웠다. 하지만 스마트 공항의 일환으로 세계 유수의 공항에서 방대한 양의 여객 데이터를 수집하고 있고, 축적된 데이터를 활용하여 공항 내 여객 체류 시간 분석이 가능해졌다. 본 연구에서는 인천 국제 공항에서 수집된 여객 데이터를 활용하여 여객 체류 시간 분석을 수행하였으며, 방대한 양의 자료를 효율적으로 처리하기 위해 데이터 마이닝 기법인 클러스터링을 활용하여 여객을 체류 시간에 따라 구분하였다. 분석 결과 인천 국제 공항 출발 여객은 체류 시간에 따라 1) 체류 시간이 짧고 대부분의 시간을 에어사이드에서 보내는 여객, 2) 평균 3 시간 정도의 체류 시간을 갖는 여객, 3) 총 체류 시간이 압도적으로 긴 여객 등 크게 3 개의 클러스터로 구분할 수 있는 것으로 나타났다.

### [Abstract]

This paper is concerned with departure passengers' dwell time analysis using real system data. Previous researches emphasize the importance of dwell time analysis from perspective of airport terminal planning and non-aeronautical revenue. However, short-term airport operation using passengers' dwell time is considered impossible due to absence of passengers' behavior data. Recently, in accordance with the wave of smart airport, world leading airports are systematically collecting passenger data. So there is high possibility of analyzing passengers' dwell time with the data stacked in the airport database. We conducted dwell time analysis using data from Incheon Int'l airport. In order to handle passenger data, we adapted clustering algorithm which is one of data mining techniques. As a clustering result, passengers are divided into 3 clusters. One is the cluster for passengers whose dwell time is relatively short and who tend to spend longer time in the airside. Another is the cluster for passengers who have near 3 hours dwell time. The other is the cluster for passengers whose total dwell time is extremely long.

**Key word** : Smart airport, Passenger flow management, Dwell time, Passengers' pattern, Clustering, K-means.

<https://doi.org/10.12673/jant.2019.23.5.380>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 7 October 2019; Revised 8 October 2019

Accepted (Publication) 25 October 2019 (30 October 2019)

\*Corresponding Author

Tel: +82-2-300-0150

E-mail: hojongbaik@gmail.com

## I. 서론

### 1-1 연구의 배경 및 목적

세계적인 공항들은 스마트 공항을 실현하기 위하여 많은 노력을 기울이고 있다. 스마트 공항의 개념 및 정의는 계속해서 진화하고 있는 상황이다. 하지만 승객의 흐름을 예측하고 관리하는 여객흐름관리 (PFM; passenger flow management)는 스마트 공항을 실현하는 하나의 좋은 전략이 될 수 있다.

PFM은 공항 시설물의 접점에서 여객 탐지 및 추적을 통해 공항 내에서 여객의 흐름을 예측하고 관리하는 것을 말한다[1]. PFM을 구현하기 위해서는 여객 출현 시점, 여객 체류 시간 등과 같은 여객 행동 특성에 대한 분석이 요구된다. 설문 조사를 통해 수집되는 장기적인 공항 이용객 특성과는 다르게 여객 데이터 수집의 어려움으로 PFM을 위해 필요한 단기적인 여객의 행동 특성을 파악하는 것은 어렵다고 여겨져 왔다[2].

하지만 최근 영국 히드로 공항, 싱가포르 창이 공항, 독일 뮌헨 공항 등과 같은 세계적인 공항에서 실시간으로 여객 데이터를 시스템적으로 수집하고 있으며, 이러한 여객 데이터에는 여객의 성별, 나이, 국적 등과 같은 여객 기본 정보뿐 아니라 체크인 시각, 보안검색대 통과 시각 등과 같은 여객의 행동을 유추할 수 있는 정보가 포함되어 있다. 이렇게 실시간으로 수집되는 여객 관련 정보를 활용하여 PFM 실현을 위한 단기적인 여객 행동 특성 분석이 가능할 것으로 예상된다.

본 연구는 PFM을 실현하기 위한 기초 연구로서, 인천국제공항에서 수집된 여객 데이터를 바탕으로 공항 내 여객 흐름 예측에 필수적인 출발 여객의 체류 시간에 대한 분석을 수행하였다. 항공기 운항 정보에서 추출할 수 있는 항공기 출발 시각 (STD; schedule time of departure)과 여객 데이터에서 추출할 수 있는 체크인카운터 통과 시각, 보안 검색대 통과 시각의 차를 활용하여 출발 여객의 터미널 총 체류 시간과 에어사이드 체류 시간을 계산하고 데이터 마이닝 기법 중 하나인 클러스터링을 활용하여 여객을 체류 시간에 따라 몇 개의 클러스터로 구분하였다.

### 1-2 연구의 범위

본 연구의 수행을 위해 수집된 자료의 공간적, 시간적 범위와 본 연구의 목적을 달성하기 위해 설정된 연구의 내용적 범위는 다음과 같다.

- 1) 공간적 범위: 인천국제공항 제1터미널 출발층 (3층)
- 2) 시간적 범위: 2018년 7월 23일~7월 27일 (총 5일)
- 3) 내용적 범위:
  - 공항 내 여객 체류 시간에 대한 기존문헌 조사
  - 체류 시간 분석에 사용될 클러스터링 기법 선정
  - 여객 데이터를 활용한 공항 내 출발 여객 체류 시간 분석
  - 클러스터링 결과 해석 및 결론 도출

## II. 기존연구고찰

본 장에서는 공항 터미널 계획, 공항의 수익성과 관련된 여객 체류 시간에 대한 기존 연구를 고찰하고 공항 내 여객 체류 시간 분석의 중요성에 대해 확인한다.

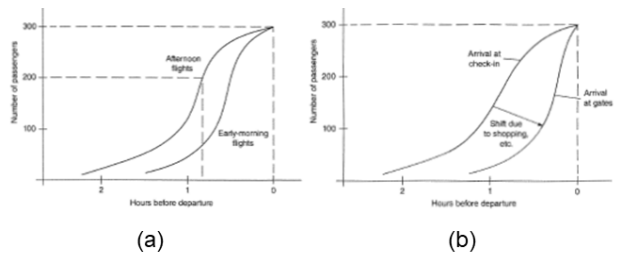
### 2-1 공항 터미널 계획

Neuville et al.(2013)은 공항 터미널 공간 계획에 있어 공항 내 여객 체류 시간 분석의 중요성에 대해 언급하였다. 공항 내에서의 여객 체류 시간은 여객이 일정 시간동안 특정 공간을 점유하고 있는 것을 의미하며, 여객의 체류 시간에 따라 터미널 설계 시 고려해야할 공간의 크기가 달라진다. 여객 체류 시간이 공항 터미널 크기를 결정하는 중요한 요소지만 시간대별, 공항별, 계절별 다양한 패턴을 나타내는 여객 체류 시간 분석에 대한 어려움에 대해서도 서술하고 있다[2].

그림 1은 유럽 주요 공항의 여객 도착 분포를 보여주고 있다. 그림 1의 (a)는 항공기 출발 시각에 따른 시간대별 여객의 도착 분포를 나타내는데, 오후 운항편 여객은 출발 시각으로부터 대략 3 시간 정도 전에 공항에 도착하는 반면, 이른 아침 운항편 여객은 출발 예정 시각에 거의 근접하게 오는 것을 알 수 있다. 같은 운항 시간대에서도 국제선 승객은 국내선 승객에 비해 일찍 공항에 도착하며, 비행 거리가 짧은 비즈니스 여객의 경우 거의 출발 예정 시각에 가깝게 공항에 도착하는 경향이 있다. 또한 여객의 도착 분포는 측정 지점에 따라서도 상이하다. 그림 1의 (b)에서 체크인 도착 분포에 비해 보딩게이트 도착 분포는 압축된 형태로 나타난다. 체크인 후 공항의 시설물 이용에 따라 여객은 출발 예정 시각에 맞추어 보딩게이트에 도착하게 되어 보딩게이트에서 보다 압축된 형태의 도착 분포를 갖는다.

### 2-2 공항 수익성

Graham(2014)은 여객의 공항 내 체류 시간이 여객 쇼핑 특성에 직접적으로 영향을 미치며, 셀프체크인의 증가, 강화된 보안 검색, 저비용항공사의 이른 보딩 시간 등이 여객의 체류 시간에 영향을 미친다고 설명하고 있다. 영국공항공단 (BAA; british



(출처: Airport Systems: Planning, Design and Management, 2013, pp.588-589)

그림 1. 유럽 주요 공항의 누적 여객 도착 분포  
Fig. 1. The cumulative arrival diagrams for a major European airports.

airport authority) 조사에 따르면 여객은 공항에서 주어진 추가 10 분당 국내선 여객은 0.8 파운드(약 1,200 원), 국제선 여객은 1.6 파운드(약 2,400 원) 가량을 소비하는 것으로 나타났다. 또한 국내선 여객의 경우 전체 공항 체류 시간의 68%를 에어사이드에서 보내는 반면, 국제선 여객은 출국 심사 등과 같은 추가적인 수속 시설물 통과에 따라 상대적으로 보다 짧은 에어사이드 체류 시간을 갖는 것으로 나타났다[3].

Torres et al.(2005)은 스페인 아스투리아스 공항의 여객을 대상으로 여객의 공항 내 체류 시간과 소비 사이의 관계에 대한 분석을 수행하였다. 휴가 여객의 공항 내 체류 시간(93.62 분)은 비즈니스 여객의 체류 시간(82.30 분) 보다 큰 것으로 조사되었으며, 두 종류의 여객 모두 공항 내 평균 소비(4.26 유로, 약 5,500 원)의 1/3을 주차비, 1/4를 식음료에 소비하는 것으로 나타났다. 또한 휴가, 비즈니스 여객 모두 공항 체류 시간이 길어질수록 공항 내 소비가 증가하고 체류 시간이 상대적으로 긴 휴가 여객의 소비가 더 큰 것으로 나타났다[4].

**2-3 시사점**

기존 문헌 및 연구 결과에서 여객의 체류 시간이 공항의 설계, 수익성, 운영에 영향을 준다는 것을 알 수 있었다. 출발 여객의 체류 시간은 공항 시설 혼잡에 직접적으로 영향을 미치는 주요 변수이며, 체류 시간에 따라 공항 설계 및 공항 운영 효율성이 달라진다. 또한 여객의 체류 시간이 공항의 수익성에 영향을 미친다는 사실은 여러 연구를 통해 증명되고 있다. 따라서 공항 운영자의 입장에서 각 공항의 여객 체류 시간 분석은 필수적으로 수행되어야 하며, 분석 결과 활용에 대한 고민을 해야 한다.

**III. 연구방법**

본 장에서는 여객 체류 시간과 관련하여 인천 국제 공항에서 수집하는 자료에 대해 알아보고, 수집된 자료를 활용하여 여객 체류 시간 분석을 위한 방법론에 대해 설명한다.

**3-1 자료 수집**

본 연구에서는 출발 예정 시각과 탑승권 발권시각, 보안검색대 통과 시각과의 차이를 각각 총 체류 시간과 에어사이드 체류 시간으로 정의하고 이를 산출한 뒤, 여객 체류 시간 분석을 수행하였다.

1) 공용체크인시스템 (AirCUS; airport common user system)

인천 국제 공항 내의 체크인카운터, 수하물카운터, 보딩게이트에서 수집되는 여객 정보로 사용 체크인카운터, 탑승권 발권 시각, 수하물 탁송 시각, 이용 항공사, 도착지, 승객 예약 기록 (PNR; passenger name record) 등의 정보가 수집된다.

2) 여객흐름관리시스템 (PFMS; passenger flow management system)

인천 국제 공항 내의 출국장 및 보안검색대에서 수집되는 여객 정보로 여객의 국적, 성별, 나이, 출국장 통과시각, 보안검색대 통과시각, 출발 예정 시각 등이 포함되어 있다.

**3-2 클러스터링 알고리즘**

1) K-means 알고리즘

K-means 알고리즘은 가장 유명하고 널리 사용되는 클러스터링 방법으로 주어진 데이터 간의 유사성(혹은 차이성) 정도에 따라 데이터를 구분하는 방법이다. 유사성은 클러스터 중심 (centroid)으로부터 각 데이터까지의 거리로 측정되며 일반적으로 해당 거리는 유클리드 거리 제곱 값으로 계산되어 진다.(수식 1 참고)

$$D(X_i, Q_l) = \sum_{i=1}^n (X_i - Q_l)^2 \tag{1}$$

수식 1에서  $X_i$ 는  $m$ 개의 특성을 갖는  $i$ 번째 데이터 벡터를 의미하며,  $[X_{i,1}, X_{i,2}, \dots, X_{i,m}]$ 으로 표현된다.  $Q_l$ 은  $l$  클러스터의 중심 벡터를 나타내고,  $D(X_i, Q_l)$ 은  $i$ 번째 데이터와  $l$  클러스터 중심까지의 거리를 의미한다. 수식 1에서 보면 데이터 간 유사성을 판단하는 식이 간단해 K-means 알고리즘은 대용량 자료 처리에 효율적이라는 장점을 갖고 있다[5], [6]. 수식 2는 K-means 알고리즘의 비용 함수식이며, 해당 식을 최소화하는 방향으로 클러스터가 구분된다.

$$P(W, Q) = \sum_{l=1}^k \sum_{i=1}^n w_{i,l} \times D(X_i, Q_l) \tag{2}$$

where  $w_{i,l} = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ object is in } l^{\text{th}} \text{ cluster} \\ 0 & \text{otherwise} \end{cases}$

$$\sum_{l=1}^k w_{i,l} = 1$$

**표 1. 공용체크인시스템 샘플 데이터**  
**Table 1. AirCUS data (sample).**

ISSUE_DATE	ISSUE_TIME	WKS_ID	PNR_CODE	ARR_APO	AIRLINE_CODE	FLIGHT	DEP_DATE	SEATNO
20180723	094704	ICN1CC32	G7H1XJ	HKG	UO	0619	20180723	007E
20180723	094756	ICN1CC32	VYF73A	HKG	UO	0619	20180723	007A
20180723	094759	ICN1CC32	VYF73A	HKG	UO	0619	20180723	007B
20180723	094801	ICN1CC32	VYF73A	HKG	UO	0619	20180723	007C
20180723	094803	ICN1CC32	VYF73A	HKG	UO	0619	20180723	007D

**표 2. 여객흐름관리시스템 샘플 데이터**  
**Table 2. PFMS data (sample).**

DT_VB	NM_V4	ENT_CD	PSG_MBT_ID	REG_DATE
20180727	0653	0800	20180727PR04850031K_MUHAU	2018-07-27 06:53:21.0
20180727	0653	0800	20180727LJ02310053D_NOUKNW	2018-07-27 06:53:49.0
20180727	0653	0800	20180727LJ02310053E_OOJE	2018-07-27 06:53:56.0
20180727	0704	0800	20180727LJ02310005E_BEOI	2018-07-27 07:04:42.0
20180727	0704	0800	20180727LJ02310006E_NBGI	2018-07-27 07:04:50.0

$w_{i,l}$ 은  $i$ 번째 데이터가  $l$  클러스터에 포함될 때 1의 값을 갖는 정규화된 가중치를 의미한다. 모든 데이터는 오직 하나의 클러스터에 포함되어야 하며, 본 연구에서는 하나의 데이터가 여러 군집에 포함되는 명확하지 않은 경우(fuzzy case)에 대해서는 다루지 않는다.

2) K-means++ 알고리즘

일반적인 K-means 알고리즘은 초기 클러스터 중심을 임의 선택한 후 일련의 계산을 진행한다. 따라서 초기 클러스터 중심 선택이 클러스터 분석 결과에 민감하게 작용하며, 알고리즘 수행 결과 지역해(local minima)가 산출될 가능성이 크다. 이런 약점을 극복하고자 제안된 것이 K-means++ 알고리즘이다. 이 알고리즘은 극단적인 클러스터 중심의 선택을 피하고 고루 분산된 클러스터 중심을 선택하기 위해 각 데이터와 클러스터 중심 사이 거리와 데이터 밀집 정도에 대한 가중치를 둔 확률 식을 기반으로 클러스터 중심을 결정한다[7]. K-means++ 알고리즘을 활용한 초기 클러스터 중심 선택 방법은 다음과 같다.

Step 1-a. 전체 데이터 중 임의로 한 개의 클러스터 중심 선택

$$b. \frac{D(x)^2}{\sum_{x \in X} D(x)^2}$$

( $D(x)$ 는 데이터와 가장 가까운 클러스터 중심 사이의 거리)

c. 설정된 클러스터 개수만큼 1-b 반복

Step 2. K-means 알고리즘 수행

3) 클러스터 개수(K) 결정

K-means 알고리즘을 수행할 때는 앞서 언급한 클러스터의 초기 중심 선택 문제뿐만 아니라 총 클러스터 개수의 결정 또한 중요하게 논의되어야 한다. 클러스터링 기법에서 몇 개의 클러스터로 데이터를 구분하느냐는 연구자에 따라 달라질 수 있다. 이러한 관점에서 K-means 알고리즘은 분석 대상 데이터와 관련된 전문가와의 협업이 요구된다[8]. 본 연구에서는 수식 2에서 제시된 비용 함수 값의 변화 추이를 고려하여 유의미하다고 판단되는 클러스터의 개수를 선택하였다.

그림 2는 클러스터 개수에 따른 비용 함수 값의 변화를 나타내는 그래프이다. 그래프에서 클러스터 개수가 증가함에 따라 비용 함수 값이 감소하는 것을 알 수 있다. 클러스터 개수가 3개로 증가할 때까지는 비용 함수의 감소폭이 크지만, 그 이후에는 큰 변화가 없는 것으로 나타나 본 연구에서는 클러스터 개수를 3개로 선정하고 여객 체류 시간에 대한 분석을 수행하였다.

IV. 결과분석

4-1 총 체류 시간 클러스터링 결과

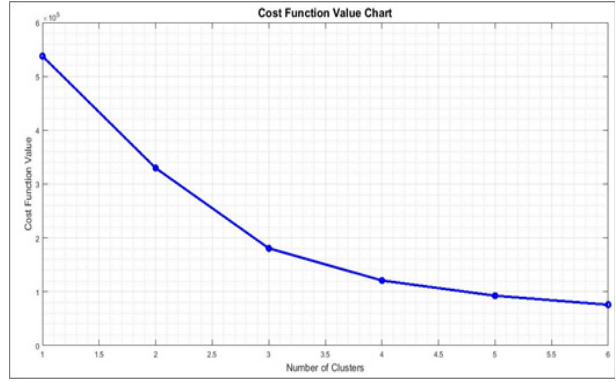


그림 2. 클러스터 개수에 따른 비용 함수 값  
Fig. 2. Cost function VS. number of clusters.

그림 3과 표 3은 공항 내 출발 여객의 총 체류 시간에 대한 클러스터링 결과를 보여준다. 그림 3에서 가장 진한 초록색은 총 체류 시간이 1 시간 미만인 여객을 의미하며 노란색은 6 시간 이상의 총 체류 시간을 갖는 여객을 의미한다.

총 체류 시간의 관점에서 전체 여객의 60% 정도는 공항 내 총 체류 시간이 평균 약 2 시간 미만인 첫 번째 클러스터로 구분된다. 나머지 40% 정도의 여객은 약 3 시간가량의 총 체류 시간을 갖는 두 번째 클러스터로 구분되며, 아주 미미하지만 전체 여객 중 1%는 평균 체류 시간이 11 시간이 넘고 3 시간 이상의 큰 편차를 갖는 세 번째 클러스터로 구분되는 것을 알 수 있다

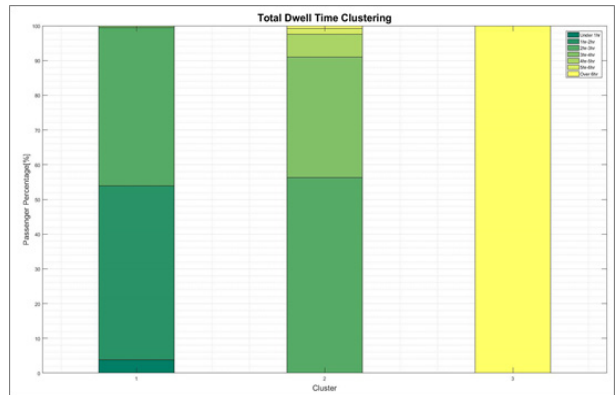


그림 3. 총 체류 시간 클러스터  
Fig. 3. Clusters for total dwell time.

표 3. 총 체류 시간 클러스터링 결과  
Table 3. Clustering results for total well time.

	Cluster 1	Cluster 2	Cluster 3
Percentage	61%	38%	1%
Mean	1hr 54min	3hr 7min	11hr 22min
Standard Deviation	28min	40min	3hr 10min



#### 4-2 에어사이드 체류 시간 클러스터링 결과

여객의 에어사이드 체류 시간의 경우도 여객의 총 체류 시간의 클러스터링 결과와 비슷하게, 클러스터에 따라 여객의 에어사이드 체류 시간이 증가하는 것으로 나타났다. (그림 4, 표 4 참고) 하지만 앞선 결과와는 다르게 세 번째 클러스터의 평균 에어사이드 체류 시간은 두 번째 클러스터의 체류 시간과 큰 차이가 없는 것으로 보아 여객의 총 체류 시간에 따라 세 번째 클러스터가 구분되는 것으로 보인다.

#### 4-3 클러스터링 분석 결과 요약

클러스터링 분석 결과 여객은 체류 시간에 따라 크게 3 개의 클러스터로 구분되었다. 전체 여객의 61%에 해당하는 첫 번째 클러스터에 속하는 여객들은 평균 2 시간미만의 총 체류 시간과 1 시간 30분 정도의 에어사이드 체류 시간을 갖는 것으로 나타났다. 첫 번째 클러스터에 속한 여객은 대부분의 시간을 에어사이드에서 소비하는 것으로 분석되었다.

두 번째 클러스터에 속하는 여객은 평균 약 3 시간가량의 총 체류 시간을 갖고, 이 가운데 2 시간 40 분가량을 에어사이드에서 보내는 것으로 나타났다. 첫 번째 클러스터의 여객보다는 상대적으로 공항에 일찍 도착한 승객으로, 그에 따라 랜드사이드, 에어사이드에서 모두 첫 번째 클러스터의 여객보다 많은 시간을 소비하는 것으로 나타났다.

마지막으로 전체 여객의 1%에 해당하는 세 번째 클러스터에 속하는 여객은 11 시간 이상의 시간을 총 체류 시간으로 갖

지만 에어사이드 체류 시간은 두 번째 클러스터의 여객과 비슷한 것으로 나타났으며 총 체류 시간의 차이가 세 번째 클러스터를 결정하는 요인으로 작용한 것으로 보인다. 세 번째 클러스터에 해당하는 여객은 주로 에바 항공 이용 여객이었으며, 항공사에서 제공하는 이른 체크인 서비스에 따라 먼저 체크인을 한 후 별도의 일정을 마치고 출발 예정 시각에 맞추어 에어사이드로 들어오는 것으로 보인다.

### V. 결 론

#### 5-1 연구 결과

본 연구에서는 공항 터미널 계획, 공항 수익성 측면에서 중요하고 PFM 구축에 직접적으로 영향을 주는 공항 내 여객 체류 시간에 대한 분석으로 하였다. 공항에서 실시간으로 수집되는 여객 데이터를 바탕으로 체류 시간 분석으로 수행하였으며, 많은 양의 데이터를 보다 효율적으로 분석하기 위하여 데이터 마이닝 기법 중 하나인 클러스터링을 이용하여 여객을 체류 시간에 따라 구분하였다.

클러스터링 결과, 인천 국제 공항을 이용하는 여객은 총 체류 시간과 에어사이드 체류 시간에 따라 크게 3 가지 클러스터로 구분되는 것으로 나타났다. 첫 번째 클러스터는 총 출발 여객에 61%에 해당하는 클러스터로서 2 시간 이하의 짧은 체류 시간을 갖고 에어사이드에서 압도적으로 많은 시간으로 보내는 여객 클러스터로 나타났다. 반면, 두 번째 클러스터의 여객은 첫 번째 클러스터에 비해 총 체류 시간, 에어사이드 체류 시간 모두 길었으며, 에어사이드에 체류하는 시간이 더 길지만 첫 번째 클러스터에 비해 랜드사이드에 체류하는 시간 역시 긴 것으로 나타났다. 마지막으로 세 번째 클러스터의 경우 대부분 특정 항공사를 이용하는 여객으로 이른 체크인에 따라 굉장히 큰 총 체류 시간을 갖는 클러스터로 구분되었다.

그동안 여객 데이터 수집의 어려움으로 여객의 단기적인 행동 특성에 대한 분석을 어려운 것으로 간주되어 왔고 그에 대한 연구는 미미하였다. 하지만 스마트 공항 구축의 일환으로 방대한 양의 여객 데이터가 수집되고 있으며, 축적된 데이터를 활용하여 스마트 공항을 실현하기 위한 논의가 활발하게 이루어지고 있다. 그런 관점에서 본 연구는 실제 공항에 수집된 데이터를 활용하여 여객의 행동 특성을 분석하였다는 데 그 의의가 있다.

#### 5-2 추후 연구

본 연구는 실제 여객 데이터를 활용하여 여객의 체류 시간을 분석하였다는 데 의의가 있지만 여전히 한계점 존재하며 그에 따른 추후 연구가 필요하다.

첫째, PFM을 실현하기 위해서는 여객의 터미널 도착 시간에 대한 정보 수집이 필요하다. 본 연구에서는 공항에서 여객의 정

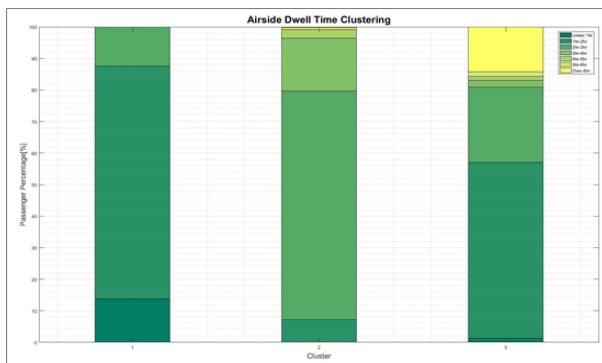


그림 4. 에어사이드 체류 시간 클러스터  
Fig. 4. Clusters for airside dwell time

표 4. 에어사이드 체류 시간 클러스터링 결과  
Table 4. Clustering results for airside well time

	Cluster 1	Cluster 2	Cluster 3
Percentage	61%	38%	1%
Mean	1hr 30min	2hr 40min	2hr 55min
Standard Deviation	25min	37min	2hr 30min

보가 처음으로 수집되는 체크인카운터를 기준으로 하여 여객의 체류 시간을 분석하였지만, 향후 여객의 공항 도착 시간에 대한 정보를 활용한 여객의 행동 특성 분석이 수행되어야 한다.

둘째, 본 연구는 PFM에 활용될 수 있는 기초 연구로 여객의 체류 시간을 분석하였다. 따라서 해당 결과를 토대로 하여 성별, 연령, 국적 등과 같은 여객의 인구통계학적 정보와 항공기 운항 정보 등을 결합하여 실제 공항에서의 여객 흐름 예측이 가능한 모형 구축이 요구된다.

이러한 한계에도 불구하고, 본 연구는 실제 여객 데이터를 활용하여 여객의 체류 시간을 분석하였다는 점에서 PFM의 기초자료로 활용될 수 있을 것으로 기대된다.

## References

- [1] Frost & Sullivan, Global airport passenger flow management market, forecast to 2025, Frost & Sullivan, San Antonio: TX, Report ASDR-431586, 2017.
- [2] R. de Neufville and A. R. Odoni, *Airport Systems: Planning, Design and Management*, 2nd ed, New York, NY: McGraw Hill Education, pp. 497-611, 2013.
- [3] A. Graham, *Managing Airports*, 4th ed, London: Routledge, pp. 188-219, 2014
- [4] E. Torres, J. S. Dominguez, L. Valdes, and R. Aza, "Passenger waiting time in an airport and expenditure carried out in the commercial area," *Journal of Air Transport Management*, Vol. 11, Issues 6, pp. 363-367, Nov. 2005.
- [5] Z. Huang, "Extensions to the k-means algorithm for clustering large data sets with categorical values," *Data Mining and Knowledge Discovery*, Vol. 2, Issues 3, pp. 283-304, Sept. 1998.
- [6] S. Raschka, *Python Machine Learning*, 1th ed, Birmingham: Packt Publishing Ltd., 2015
- [7] D. Arthur and S. Vassilvitskii, "K-means++: the advantages of careful seeding," in *Proceeding of the 18<sup>th</sup> annual ACM-SIAM Symposium on Discrete Algorithms*, New Orleans: LA, pp. 1027-1035, Jan. 2007.
- [8] S. Theodoridis, A. Pikrakis, K. Koutroumbas and D. Cavouras, *Introduction To Pattern Recognition: A Matlab Approach*, Cambridge, MA: Academic Press, 2010



### 안 덕 배 (Deok-Bae An)

2017년 3월 ~ 현재 : 한국항공대학교 일반대학원 항공교통물류학부  
 ※ 관심분야 : 항공교통, 시뮬레이션, 최적화, 공항운영



### 김 휘 양 (Hui-Yang Kim)

2011년 : 한국항공대학교 (경영학석사)  
 2000년 2월 ~ 2018년 2월 : 국토교통부  
 2018년 3월 ~ 현재 : 한국항공대학교 항공교통물류학부 조교수  
 ※ 관심분야 : 항공교통, 공역 및 비행절차, 수용량, 항행계획



### 백 호 종 (Ho-Jong Baik)

2000년 : Virginia Tech (공학박사)  
 2007년 8월 ~ 2010년 8월 : 미주리 주립대 조교수  
 2010년 8월 ~ 현재 : 한국항공대학교 항공교통물류학부 교수  
 ※ 관심분야 : 항공교통, 시뮬레이션, 최적화, ATM