

論文

합류하는 두 항공기간 도착순서 결정에 대한 로지스틱회귀 예측 모형

정소연*, 이금진**

Prediction Model with a Logistic Regression of Sequencing Two Arrival Flows

Soyeon Jung* and Keumjin Lee**

ABSTRACT

This paper has its purpose on constructing a prediction model of the arrival sequencing strategy which reflects the actual sequencing patterns of air traffic controllers. As the first step, we analyzed a pair-wise sequencing of two aircraft entering TMA from different entering points. Based on the historical trajectory data, several traffic factors such as time, speed and traffic density were examined for the model. With statistically significant factors, we constructed a prediction model of arrival sequencing through a binary logistic regression analysis. With the estimated coefficients, the performance of the model was conducted through a cross validation.

Key Words : Sequencing Strategy(순서결정기법), Logistic Regression(로지스틱 회귀모형), Data-driven(데이터 기반), Machine Learning(기계 학습), Air Traffic Controller(항공교통관제사), Cross Validation(교차 검증)

1. 서 론

최근 항공교통수요의 지속적인 증가로 인하여 세계 주요공항 인근 공역의 항공교통량은 포화 상태에 이르렀다. 증가하는 항공교통량에 비해 공역 및 공항자원은 한정되어 있으며, 따라서 항공기의 출도착 지연, 조종사 및 관제사의 과도한 업무 부하 등으로 인한 항공기 운항의 비효율성 문제가 심화되고 있다. 이러한 문제점을 해결하기 위하여 항공교통관제사들의 관제 업무를 보다 효율화 하고, 출도착 지연을 최소화하며 항공교통흐름을 원활히 할 필요성이 어느 때

보다 높아졌다.

도착 접근 단계의 항공기에 대한 항공교통관리(ATM, Air Traffic Management) 업무 중 가장 중요한 것은 도착순서결정(Sequencing) 및 스케줄링(Scheduling)이다. 이는 다양한 방향의 경로에서 진입하는 항공기들에 대하여 착륙순서를 결정하고, 이들을 순서에 따라 일렬로 배열하여 질서 있게 공항에 착륙시키는 업무를 의미한다.

최근 항공교통관제사들의 도착순서결정 및 스케줄링 업무를 보조하기 위한 다양한 자동화 의사결정 지원 시스템이 개발되고 있다[1-3]. 특히 도착관리기법(AMAN, Arrival Management)에 대한 연구가 세계적으로 활발히 진행되어 왔다. Fig. 1에 나타난 바와 같이 도착관리기법은 특정 공항으로 진입하는 항공기들 간의 분리간격과 공항의 수용량을 고려해 도착계획시간(STA, Scheduled Time of Arrival) 및 지연 정보를 제공하는 도구이다[4,5].

2015년 12월 7일 접수 ~ 2015년 12월 19일 심사완료
게재 확정일 (2015년 12월 20일)

* 한국항공대학교 대학원 항공교통물류학과

** 한국항공대학교 항공교통물류학과 조교수

연락처, E-mail : soyeon.c.jung@gmail.com

경기도 고양시 덕양구 항공대학로 76

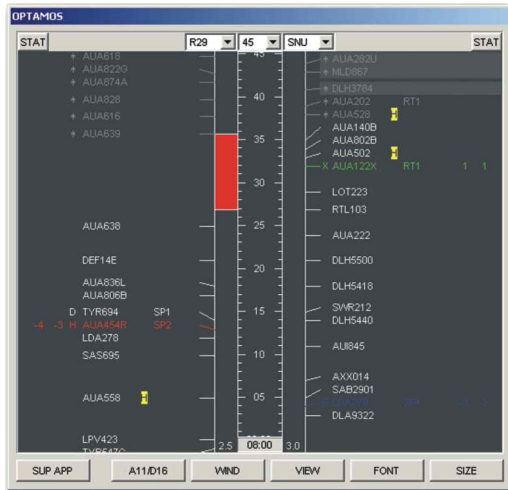


Fig. 1 AMAN user interface (OPTAMOS)[6]

도착관리기법은 항공기의 도착예정시간(ETA, Estimated Time of Arrival)을 기반으로 도착순서를 결정하는데, 도착예정시간은 항공기 성능에 대한 항공역학적인 모형을 통해 계산된다 [7,8]. 하지만 실제 관제사들에 의한 도착순서결정은 공역 및 항공로 구조, 항공교통 복잡도 등 다양한 요소들이 종합적으로 고려되어 이루어지며, 따라서 실제 관제사들의 행동 패턴이나 순서결정 전략에 대한 데이터 분석이 결여된 기존의 방식은 운용현장에서 바로 상용화되기에 많은 문제점이 있다.

본 연구는 접근 항공기들에 대한 항공교통관제사들의 도착순서 결정전략에 대한 데이터 기반 예측 모형을 제시하는 것을 목표로 한다. 이를 위하여 인천공항(RKSI) 15L/R와 16방향 활주로로 착륙한 항공기들의 항적 데이터를 기반으로 하여 도착순서 결정에 영향을 미치는 요인들을 분석하고, 이를 토대로 이항 로지스틱 회귀모형을 구축한 후 성능을 평가하였다.

2. 연구 대상

2.1 대상 공역 및 항로

Fig. 2는 서울접근관제구역(서울TMA)에 위치한 인천국제공항과 해당 공항의 15L/R, 16방향 활주로를 사용하는 표준계기도착절차(STAR)들의 모습이다. 서울TMA에 진입하여 인천국제공항 15L/R, 16방향 활주로로 착륙하는 모든 항공기는 REBIT 1N, OLMEN 1N, GUKDO 1N, KARBU 1N의 네 가지 STAR 절차 중 하나를

따라 인천공항에 접근한다. 실제 관제현장에서는 네 곳에서 진입한 항공기들에 대한 순서결정이 통합적으로 이루어진다. 하지만 본 연구에서는 절차를 통해 진입하는 모든 항공기 전체에 대한 통합순서결정을 분석하기에 앞서 두 절차 간의(pair-wise) 순서결정 패턴을 분석하였다.

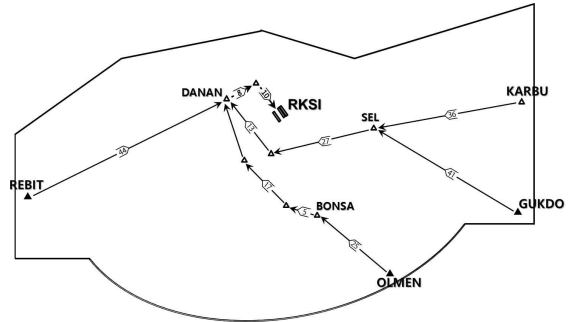


Fig. 2 Seoul TMA and STAR procedures of RCSI using RWY15L/R and RWY16[9]

네 개의 STAR 절차로부터 다양한 절차쌍 조합이 가능하나, 대표적으로 REBIT과 OLMEN 절차를 따르는 항공기들 간의 순서결정을 분석하였다(Fig. 3). 두 절차 간의 순서결정패턴에는 REBIT을 통과한 항공기(AC_{REBIT})를 먼저 착륙시킨 후 OLMEN을 통과한 항공기(AC_{OLMEN})를 착륙시키는 경우와 반대 순서로 착륙시키는 경우의 두 가지 패턴이 존재한다.

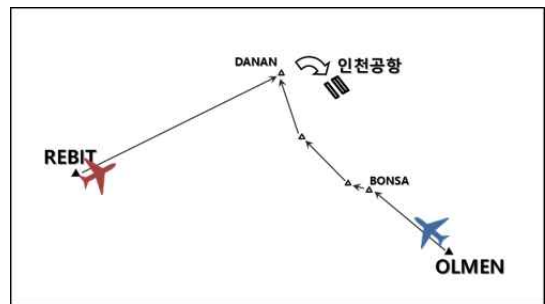


Fig. 3 Arrival aircraft to RCSI entering through REBIT and OLMEN[9]

2.2 대상 항공기

본 연구는 세 달 치의 과거 항적 데이터로부터 추출한 REBIT과 OLMEN 절차를 따라 인천국제공항에 착륙하는 2,541쌍의 항공기 데이터를 대상으로 한다. 전체 2,541쌍 중에서 AC_{REBIT} 을 먼저 착륙시킨 후 AC_{OLMEN} 을 착륙시키는 패

턴을 따르는 데이터는 1,434쌍이었으며, 반대의 패턴을 따르는 데이터는 1,107쌍이었다.

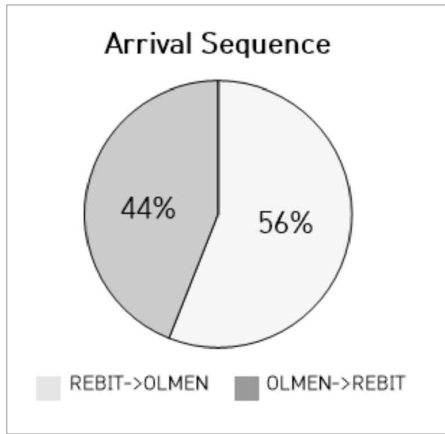


Fig. 4 Proportion of arrival sequences between REBIT and OLMEN

3. 도착순서결정 예측모형

3.1 로지스틱 회귀모형

본 연구는 이항 로지스틱 회귀모형(binary logistic model)을 사용하여 항공교통관제사의 순서결정패턴을 분석하였다. 로지스틱 회귀모형은 종속변수가 범주형(categorical) 질적변수인 경우에 개별 관측치들이 어느 집단으로 분류될 수 있는가를 분석하는 대표적인 이산선택모형(discrete choice model)이다[10,11]. 본 연구에서 사용할 이항 로지스틱 회귀모형은 종속변수가 두 가지 범주를 가지는 경우에 사용된다.

일반적으로 종속변수 y 는 1 또는 0의 값을 가지는 이진 변수로 정의한다. 이 때 1과 0은 y 가 각 범주에 속하는 경우를 대표하는 값이며, 이항 로지스틱 회귀방정식은 y 가 1의 값을 가질 확률을 계산한다. k 개의 독립변수를 입력하였을 경우 y 가 1의 값을 가질 확률은 다음과 같다.

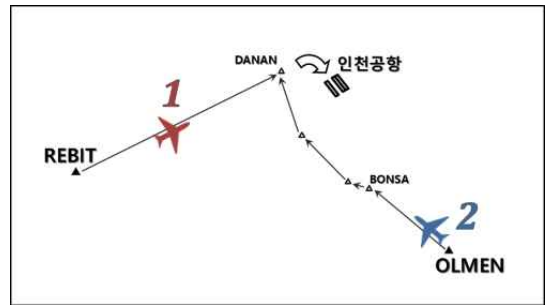
$$P(X) = \frac{e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}}{1 + e^{(\beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k)}} \quad (1)$$

$$= \frac{1}{1 + e^{-g(X)}}$$

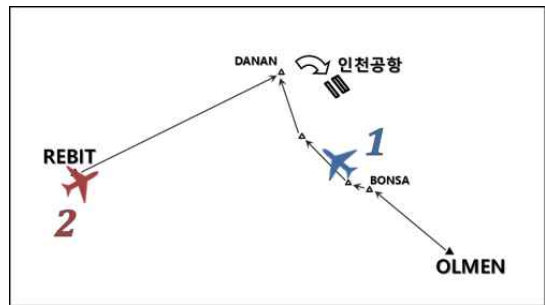
여기서 $X = [x_1 \ x_2 \ \dots \ x_k]$ 는 독립변수들이며, $B = [\beta_0 \ \beta_1 \ \dots \ \beta_k]$ 는 독립변수들의 계수를 나타

낸다. y 가 0일 확률은 $1 - P(X)$ 로 계산된다.

본 연구가 제시하는 도착순서결정 예측모형은 두 항공기의 도착순서(Seq)를 최종적인 종속변수로 취한다. Fig. 5과 같이 AC_{REBIT} 가 AC_{OLMEN} 보다 인천국제공항에 먼저 도착하는 경우를 $Seq=1$ 로 설정하고, 반대의 경우를 $Seq=0$ 로 설정하였다.



(a) Seq=1



(b) Seq=0

Fig. 5 Sequence strategies for arrival aircraft from REBIT and OLMEN[9]

3.2 독립변수의 선정

모형의 후보독립변수들로 두 항공기에 대한 관제사의 도착순서 결정에 영향을 미칠 것으로 예상되는 다섯 가지 요인을 선정하였다.

우선, FCFS(First-Come First-Served) 원칙에 따라 공항으로의 예상도착시간이 두 항공기의 도착순서 결정에 절대적인 영향을 미칠 것이라 판단된다. 따라서 예상도착시간에 영향을 미치는 두 항공기의 TMA 진입시간 차이(TD)와 진입속도 차이(SD)를 첫 번째와 두 번째 후보독립변수로 선정하였다. 또한 두 항공기의 예상도착시간이 비슷한 경우, 고도에 따라 우선순위가 결정될 것으로 생각되어 TMA 진입고도 차이(AD)를 세 번째 후보독립변수로 지정하였다. 여기서, 진입시간(T)과 진입고도(A)는 값이 작을수록 착륙 우

선순위가 높은 반면, 진입속도(S)는 값이 클수록 우선순위가 높다. 따라서 본 논문에서는 변수 TD , SD , AD 의 값이 커짐에 따라 착륙 우선순위가 높아지도록, Fig. 6에 나타난 바와 같이 각 변수를 계산하였다.

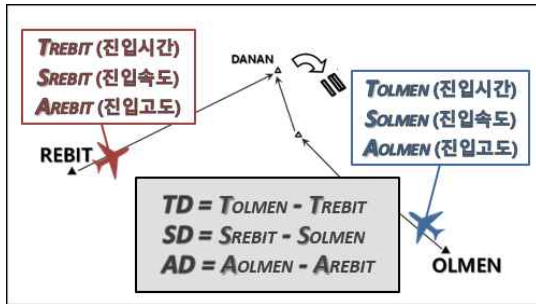


Fig. 6 Explanatory figure for TD , SD and AD variables

다음으로는, 두 항공기가 모두 TMA에 진입한 시점의 각 절차상 교통량($DEM1$, $DEM2$)을 네 번째와 다섯 번째 후보독립변수로 고려하였다. 단, 교통량의 경우 상대량 뿐 아니라 절대적인 수치가 관제사의 순서결정에 영향을 미칠 수 있는 것으로 판단되어, 각 절차상의 교통량을 별개의 독립후보변수로 선정하였다. 또한 절차상의 전체 구간보다는 진입 항공기 전방의 교통 상황이 중요할 것이라 판단된다. 따라서 Fig. 7에 나타난 바와 같이, 두 항공기 중 TMA에 늦게 진입한 항공기의 진입 시점을 기준으로 하여, 각 절차상 두 항공기를 포함한 전방의 교통량을 계산하였다.

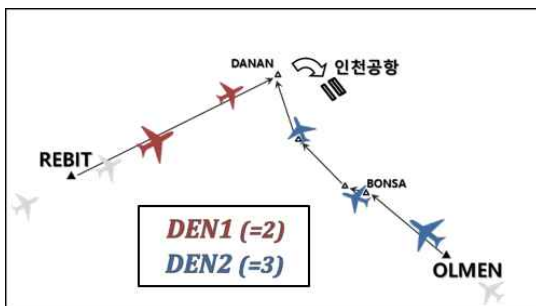


Fig. 7 Explanatory figure for $DEN1$ and $DEN2$ variables

마지막으로, 각 항공기의 직전 항공기와의 간격 또한 도착순서 결정에 영향을 줄 것이라 판단하였다. 간격은 진입시간의 차이로 계산하였

다. Fig. 8과 같이, 직전 항공기와의 간격이 충분히 넓은 항공기(AC_{REBIT})는 직전 항공기에 의한 영향을 받지 않는 반면, 직전 항공기와의 간격이 ‘일정 수준’ 이내인 항공기(AC_{OLMEN})의 경우, 관제사가 두 항공기를 하나의 묶음으로 인식하는 경향이 있을 것으로 예상하였다.

이 때, 관제사 인식의 기준이 되는 ‘일정 수준’은 다음의 방법으로 계산하였다. 항로에서 동일 고도의 항공기간 최소분리간격은 10NM이며, 세 달치 항적 데이터로부터 5,082대(2대*2,541쌍) 항공기들의 평균진입속도는 0.105nm/s(약 380KT)이다. 연속된 두 항공기가 최소분리간격을 유지하면서 평균진입속도로 비행할 경우, 약 95초의 시간 간격이 발생한다. 이를 토대로 본 연구에서는 직전 항공기와의 간격이 100초 이내인 경우 1의 값을 갖고, 그렇지 않은 경우 0의 값을 갖는 더미 변수(Dummy Variables)를 REBIT과 OLMEN에 대한 여섯 번째($INT1$)와 마지막($INT2$) 후보독립변수로 선정하였다.

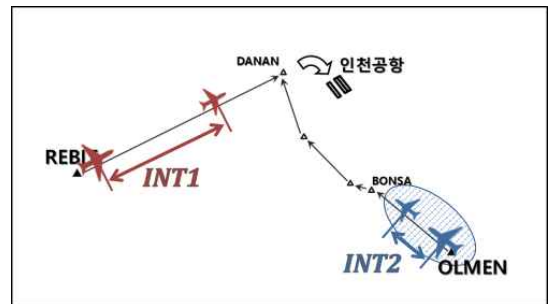


Fig. 8 Explanatory figure for $INT1$ and $INT2$ variables

4. 실험 결과

본 연구는 항공교통관제사의 도착순서결정을 예측하기 위해 로지스틱 회귀모형을 사용하였다. REBIT과 OLMEN 절차를 따르는 2,541쌍의 항공기 조합 데이터를 학습(training) 데이터로 사용하여 모형을 구축하였고, 10겹 교차 검증(10-fold Cross Validation)을 실시하였다. 모형 구축에 앞서 도착순서결정에 대한 후보독립변수들의 통계적 유의성을 검정하였고, 그 결과를 토대로 하여 이항 로지스틱 회귀모형을 구축하였다.

4.1 독립변수와 순서결정의 상관관계

모형에 적용할 독립변수들을 선별하기 위하여, 각 후보독립변수의 도착순서결정에 대한 통계적 유의성을 일변량분석(Univariate Analysis)을 통하여 검정하였다. 검정 방법으로는 독립변수 각각의 유의성을 검정하는 Wald-검정을 사용하였다. Wald-검정은 독립변수가 종속변수에 통계적으로 유의미한 영향을 미치지 않는다는, 즉, 독립변수들의 계수가 0이라는 귀무가설을 설정하고 이를 검정하는 방법이다[4,11]. Wald-검정의 검정 통계량인 W 는 식(2)와 같이 계산되며, 독립변수의 계수가 0이라는 귀무가설 하에서 표준정규분포를 따른다.

$$W = \frac{\hat{\beta}}{SE(\hat{\beta})} \quad (2)$$

이 때, $\hat{B} = [\hat{\beta}_0 \hat{\beta}_1 \dots \hat{\beta}_k]$ 는 추정된 계수들의 행렬값이며, $SE(\hat{\beta})$ 는 추정된 계수의 표준오차이다.

Table 1.은 각 후보독립변수의 계수 추정치($\hat{\beta}$), 표준오차($SE(\hat{\beta})$), W 값 및 p 값을 나타낸다. p 값을 통하여 유의수준 $\alpha = 0.05$ 하에서 TD , SD , $DEN1$, $DEN2$, $INT2$ 변수는 유의한 반면, AD , $INT1$ 변수는 유의하지 않음을 확인할 수 있다.

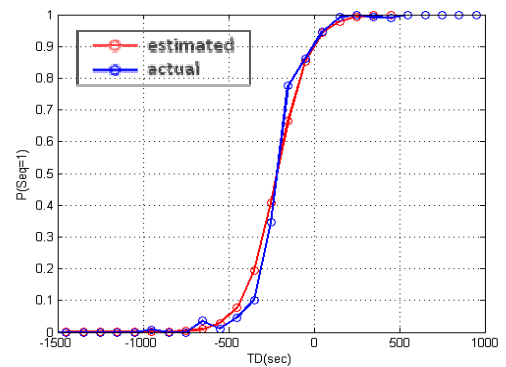
Table 1. Univariate associations of candidate independent variables and arrival sequences

Variable	Coeff.	Std.Err.	W	p-val
TD	0.011	<0.001	25.653	<0.001
SD	7.882	3.087	-2.554	0.011
AD	<0.001	<0.001	-1.277	0.202
$DEN1$	-0.818	0.040	-20.275	<0.001
$DEN2$	0.627	0.042	15.076	<0.001
$INT1$	-0.501	0.282	-1.776	0.076
$INT2$	0.6841	0.289	2.365	0.018

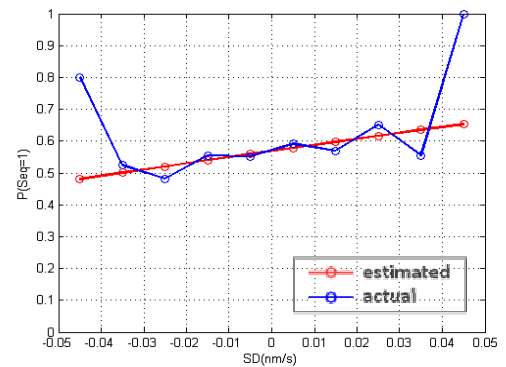
또한 Table 1.의 계수(Coeff.)들은 각 변수와 도착순서의 상관관계를 의미한다. 계수가 양의 값인 경우 해당 변수와 도착순서는 양의 상관관계를 가지며, 음의 값인 경우에는 음의 상관관계를 갖는다. 이는 Fig. 9의 그래프에서도 확인할 수 있다.

Fig. 9의 그래프들은 각 변수와 Seq=1인 확률간의 관계를 나타낸다. 청색 그래프는 실제 데이터

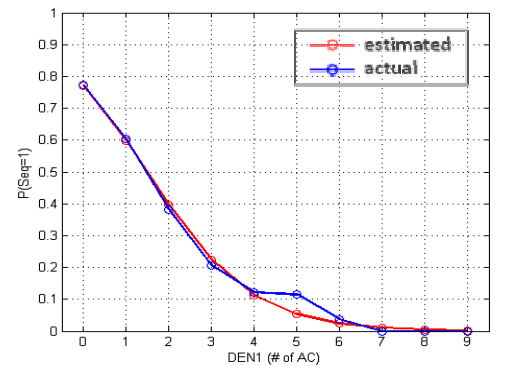
를 통해 도출된 관계식이며, 적색 그래프는 모형을 통하여 산출된 회귀식이다. TD , SD , $DEN2$ 는 $P(\text{Seq}=1)$ 와 양의 상관관계를 갖는 반면 $DEN1$ 은 $P(\text{Seq}=1)$ 와 음의 상관관계를 갖는다. 즉, AC_{OLMEN} 에 비하여 AC_{REBIT} 의 진입시간이 이룰수록, 진입속도가 빠를수록 관계사는 AC_{REBIT} 을 먼저 도착시키는 확률이 높다. 또한 두 항공기가 모두 TMA에 진입하였을 시점에 REBIT 절차상의 교통량이 많을수록 AC_{OLMEN} 를 먼저 도착시키고, 반대로 OLMEN 절차상의 교통량이 많을수록 AC_{REBIT} 을 먼저 도착시키는 확률이 높은 것으로 나타났다.



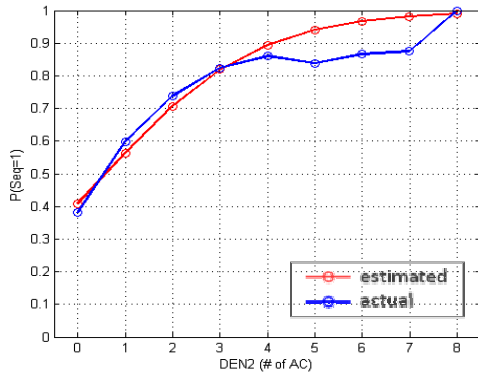
(a) TD variable



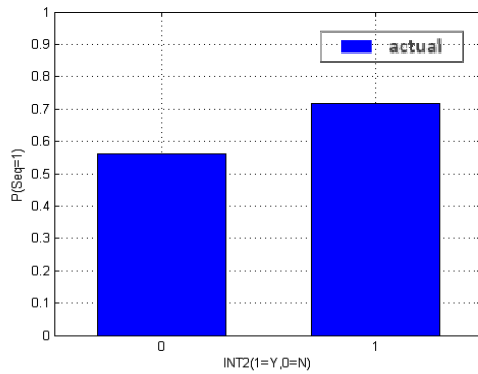
(b) SD variable



(c) $DEN1$ variable



(d) DEN2 variable



(e) INT2 variable

Fig. 9 Correlations between the arrival sequence and each independent variable

INT2 변수는 Seq=1의 확률과 양의 상관관계를 가지는 것으로 나타났다. INT2 그래프에서는 AC_{OLMEN} 가 직전 항공기와 100초 이내의 간격인 경우 $P(Seq=1)$ 이 71.67%이며, 100초 이상인 경우에는 56.07%인 것으로 나타났다. 즉, 특정 항공기가 직전 항공기와 100초 이내의 간격으로 진입한다면, 상대 항공기를 우선적으로 착륙시킨 후 해당 항공기를 착륙시키는 경향이 있다.

4.2 Pair-wise 순서결정 예측모형

4.1에서 일변량분석을 실시한 결과, 유의수준 $\alpha = 0.05$ 하에서 통계적으로 유의미한 TD , SD , $DEM1$, $DEM2$, $INT2$ 변수들이 모형의 최종적인 독립변수로 채택되었다. 과거 항적 데이터를 이용하여 선정된 독립변수들의 계수 추정치($\hat{\beta}_k$)는 Table 2와 같고, 이를 적용한 예측모형은 식 (3)과 같다. $X = [x_1 \ x_2 \ \dots \ x_k]$ 는 독립변수들이며, $\hat{B} = [\hat{\beta}_0 \ \hat{\beta}_1 \ \dots \ \hat{\beta}_k]$ 는 계수들의 추정치이다.

$\hat{P}(X)$ 는 변수값들이 주어졌을 때 Seq=1일 확률, 즉 AC_{REBIT} 를 먼저 도착시킬 예측 확률이다.

$$\hat{P}(X) = \frac{e^{(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k)}}{1 + e^{(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \dots + \hat{\beta}_k x_k)}} = \frac{1}{1 + e^{-\hat{g}(X)}} \tag{3}$$

$$\hat{g}(X) = -2.639 + 0.011 TD + 5.718 SD - 0.209 DEN1 + 0.069 DEN2 - 0.486 INT2$$

Table 2. Estimated Coefficients of the Logistic Regression Model

	Variable	Est.Coeff.
$\hat{\beta}_0$	Bias	2.639
$\hat{\beta}_1$	TD	0.011
$\hat{\beta}_2$	SD	5.718
$\hat{\beta}_3$	$DEM1$	-0.209
$\hat{\beta}_4$	$DEM2$	0.069
$\hat{\beta}_5$	$INT2$	-0.486

4.3 모형적합도 검증: 10겹 교차 검증

검증(Validation)이란 모형에 의하여 산출된 예측값을 새로운 데이터에 적용시켜 실제값과의 일치도를 검사하는 과정을 의미한다. 예측모형의 성능은 여러 방법으로 검증될 수 있다. 본 연구에서 사용한 검증 방법은 10겹 교차 검증(10-fold Cross Validation)이다. 이는 전체 데이터를 임의로 10등분하여, 총 10회에 걸쳐 서로 다른 9등분을 모형 구축을 위한 학습 데이터로 사용한 후, 나머지 1등분으로 검증을 실시하는 방법이다.

앞서 구축한 모형으로 10겹 교차 검증을 실시한 결과, 각 회차에서 모형의 정확도는 Table 3과 같이 나타났다. 여기서 정확도란 전체 검증 데이터 중 실제값과 예측값이 일치하는 비율을 의미한다. 본 논문에서 제한한 예측모형을 통해 결과적으로 본 연구에서 제안하는 모형을 사용해 관제사가 두 항공기 간 순서를 어떻게 결정하는지를 95.24%의 정확도로 예측할 수 있는 것으로 나타났다.

Table 3. 10-fold cross validation

trials	accuracy(%)	trials	accuracy(%)
1st	96.85	6th	95.28
2nd	95.59	7th	94.98
3rd	95.62	8th	94.31
4th	97.79	9th	92.44
5th	95.35	10th	97.29

5. 결 론

본 연구는 항공교통관제사의 행동 패턴을 반영한 도착 항공기에 대한 순서결정 예측모형을 개발하였다. 인천공항에 도착한 항적 데이터를 기반으로 서로 다른 절차상의 두 항공기에 대한 순서결정을 분석하였다. 관제사의 도착순서 결정에 영향을 미치는 요인들을 분석한 결과, 두 항공기의 진입시간 차이 및 진입속도 차이 뿐 아니라 항공기들이 진입한 시점의 절차상 교통량도 통계적으로 유의미한 것으로 나타났다. 변수들에 대한 분석을 토대로 도착순서 예측모형을 이항 로지스틱 회귀모형을 이용하여 구축한 후, 교차 검증을 실시하여 모형의 적합도를 검증하였다.

본 연구에서 제안한 pair-wise 모형은 향후 공역 내의 모든 접근 항공기의 전체적인 순서결정기법에 대한 연구에 유용하게 활용될 것으로 기대된다. 향후 연구에서는, 기상조건 등 도착순서 결정에 영향을 미치는 기타 요인들에 대한 추가적인 분석이 필요하다. 또한 본 연구에서 사용한 변수들에 대한 심층적인 검토가 필요하다. 예를 들어, INT1, INT2의 기준값을 100초로 설정한 것이 타당한지에 대한 연구가 필요할 것이다. 또한 대상 항공기의 전방 뿐 아니라 후방 교통상황이 순서결정에 미치는 영향에 대하여 추가적으로 살펴볼 필요가 있다. 더불어, 로지스틱 회귀모형 이외의 다양한 모형을 시도해봄으로써 데이터를 가장 정확히 반영하는 예측모형을 개발할 수 있을 것이다.

후기

본 연구는 국토교통과학기술진흥원의 “항공기 출발 및 도착 통합 관리 기술 연구”과제의 일환으로 수행되었으며 지원에 감사드립니다.

참고문헌

- 1) Erzberger, H, and Nedell, W., “Design of Automated System for Management of Arrival Traffic,” NASA TM-102201, 1989.
- 2) Saraf, A. P., and Slater, G. L., “Optimal Dynamic Scheduling of Aircraft Arrivals at Congested Airport,” Journal of Guidance, Control, and Dynamics Vol. 31, No. 1, 2008, pp. 53-65.
- 3) Hu, X. B., and Chen, W. H., “Receding Horizon Control for Aircraft Arrival Sequencing and Scheduling,” IEEE Transactions on Intelligent Transportation Systems, Vol. 6, No. 2, 2005, pp. 189-197.
- 4) 김민지, 홍성권, 이금진, “항공교통관제사의 항공기 합류순서결정에 대한 확률적 예측모형 개발”, 한국항공운항학회지 제22권 제3호, 2014.
- 5) 오은미, 김현경, 은연주, 전대근, “도착관리시스템 Queue 예측 모듈의 성능 개선을 위한 Queue 예측 정확도 분석 방법 연구”, 한국항공운항학회지, 제23권 제3호, 2015.
- 6) Eurocontrol, "AMAN Status Review 2010", 2010.
- 7) 오은미, 은연주, 전대근, “BADA를 활용한 4-D 경로 모델링법 개발”, 한국항공운항학회지, 제20권 제2호, 2012.
- 8) Eurocontrol, "User Manual For The Base of Aircraft Data (BADA) Revision 3.9", EEC Technical/Scientific Report No. 11/03/08-08, 2011.
- 9) Soyeon Jung, and Keumjin Lee, “Analysis of Air Traffic Controller’s Sequencing Strategy for Arrival Traffic Flow, AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, submitted, Washington D.C., 2016.
- 10) Train, K. E., *Discrete Choice Methods with Simulation*, Cambridge University Press, New York, 2003.
- 11) Hosmer, D. W. and Lemeshow, S., *Applied Logistic Regression*, John Wiley & Sons, Inc, New York, 2000.