

# 데이터마이닝을 이용한 공군 무기정비병의 조기 숙달을 위한 배속방안 연구

(An Effective Recruits' Assignment Method for Early Job Adaptation of  
 Air-munition Maintenance Airmen Using Datamining Technique)

강 규 영(Kew Young Kang)\*, † 윤 봉 규(Bong Kyoo Yoon)\*\*

## ABSTRACT

Recently, the military service period has been shortened continuously. Meanwhile, more skilled airmen are needed as the complexity of weapon systems increase. This phenomenon could lead to a disastrous result such as deteriorating the level of the readiness and the fighting power. We suggest a method to improve recruit's maintenance capability rapidly by assigning airmen to jobs appropriate to their characteristics using Datamining methods (K-menas and CART). We focus on the assigning method for air force's air-munition maintenance airmen since they are requested more skilled than other airmen. Grouping airmen with k-means method and devising classification rule with CART algorithm, we found that airmen's proficiency arrival period could be shortened by 1.79 months when they are assigned in the suggested way.

**Keywords :** The Period of Service, Assignment Method, Data Mining, K-means, CART

논문접수일 : 2011년 1월 26일 심사(수정)일 : 2011년 2월 25일 논문게재확정일 : 2011년 3월 8일

\* 공군 군수사령부 항공기술연구소

\*\* 국방대학교 운영분석과 부교수

† 교신저자

## 1. 서론

군 복무기간과 관련된 논의가 최근 활발하게 이루어지고 있다. 이 논의의 핵심은 여러 가지 안보 현실을 고려할 때 현재의 복무기간으로 안보 위협에 대처하기 적당한 수준의 전투력을 유지할 수 있는가와 군 복무로 인한 국민들의 기회비용을 고려할 때 현재의 복무기간이 적정한가라는 두 가지 측면이 상충관계로 얽혀 있는 문제이다. 이 문제는 간단히 대답하기에는 어려운 측면이 있는 복잡한 문제이므로 결국 이에 대한 답은 두 측면을 동시에 고려한 신중한 접근이 필요하다. 그러나 현역병의 복무기간을 현재보다 늘이는 것이 쉽지 않은 현실을 고려한다면 현재 복무기간 단축에 대한 사회적 압력이 거센 것은 부정할 수 없는 사실이다.

복무기간과 관련된 주요 고려사항의 하나인 병사들의 전투력 수준 저하에 대한 대책 없이 복무기간이 단축된다면 장기적으로 안보상의 위협을 내부에서 키우는 결과로 이어질 가능성이 있다. 복무기간 단축으로 인한 공백을 메우기 위한 방안으로 무기체계 향상을 통한 전투력 제고, 효과적인 교육훈련 방법 도입 등 다양한 방안이 있을 수 있으나, 본 연구에서는 병사들의 특성을 파악하여 파악된 특성에 부합하는 업무에 병사들을 배치함으로써 전투력 수준을 제고할 수 있는 방안을 제시하고자 한다. 특히 본 연구에서는 기술군의 특징이 강해 숙련된 병사를 양성하는데 타군에 비해 시간과 비용이 많이 들고, 숙련도가 전투력 수준에 중요한 역할을 할 가능성이 높아 복무기간 단축으로 인한 영향이 큰 공군 무기 정비병들의 숙련도 제고 방안에 대해서 살펴본다.

공군 병사의 복무기간은 6.25전쟁 직후 36개월이었으나, 이후 지속적으로 단축되어 2010년에는 24개월까지 단축되었다. 2014년까지 21개월로 가파른 단축 계획을 가지고 있었으나 천안함 사태와 연평도 포격 이후 복무기간 단축에 대한 논의는

잠시 중단된 상태이다. 그러나 복무기간 단축에 대한 사회적 압력은 다시 높아질 가능성이 있다.

복무기간 단축에 대해 군에서는 안보환경의 큰 틀이 변하지 않는 현 상황에서 더 이상의 복무기간 단축은 병사정원 감소 및 숙련도 저하로 전투력 저해요인이 된다는 의견을 내고 있는 한편, 전투력 공백을 극복하기 위해 무기체계 과학화, 병사정원 확대 등 다양한 수단을 강구하고 있다. 그러나 무기체계 과학화는 단기간에 해결할 수 없고 많은 추가예산을 필요로 하는 문제점이 있다. 또한 병사 정원 확대 방안은 최근 극심한 출산을 감소로 인한 대상자 부족으로 시행에 차질이 예상되고 있다. 한편, 무기체계 과학화는 병사들의 숙련도 향상이 수반되어야 하나 복무기간 단축은 숙련도를 저하시키는 측면이 있어 문제를 더욱 복잡하게 만들고 있다. 이를 해결하기 위해 도입한 지원병 및 간부 정원을 확대하는 방안은 막대한 유인예산 마련을 수반해야 하는 단점이 있다.

이와 같이 병사의 복무기간이 단축된 현실에 대비한 뚜렷한 해결책을 제시할 수 없는 상황에서 본 연구에서는 한정된 복무기간동안 병사 활용을 극대화할 수 있는 방안으로 병사가 조기에 숙련도가 높은 ‘숙달급’에 도달하도록 병사의 특성을 고려하여 최적의 근무지에 배속하는 방안을 제시한다.

군 복무기간은 여러 차례에 걸쳐 안보환경 변화와 복무대상자 정원의 증감에 따라 늘어나거나 줄어들었다. 복무기간은 복무대상자 정원 증감뿐만 아니라 안보 위협의 수준, 무기체계 등을 종합적으로 고려해서 판단해야 하는 문제다. 이런 차원에서 권희면(1992)과 정주성(2003)은 복무기간 단축의 전제조건과 적정복무기간에 대한 연구를 수행했다. 권희면(1992)은 현역병 복무기간 단축 조건을 북한 위협감소, 병역자원 증가, 간부 및 지원병 확보, 첨단 무기체계 확보로 들었다. 정주성(2003)은 육군 병사의 최소복무기간을 병과별로 구체적으로 제시했는데, 보병과 포병은 22개월, 기갑과 정비는 25개월이 최소복무기간으로 필요

하다는 결과를 제시했다.

이와 같은 연구결과에도 불구하고 군 복무기간은 사회일각에서의 지속적인 단축요구, 선거 때마다 정치권의 복무기간 단축공약, 세계적으로 징병제 국가의 복무기간 단축 및 지원병제로 전환 추세에 따라 지속적으로 단축 압력을 받고 있다.

한편, 신성한 병역의 의무 수행을 위해 제공된 국민의 귀중한 시간을 효과적으로 활용한다는 차원과 단축된 복무기간의 제약조건 하에서 활용도를 극대화하는 차원에서 병역의무 수행 효과를 극대화할 수 있는 방안에 대한 연구가 수행되었다. 이에 대한 연구는 병사의 인성에 초점을 맞추어 복무환경을 개선하고 만족도를 높일 수 있는 방안을 제시하는 연구와 병사를 목적에 맞게 할당하는 방안을 제시하는 연구로 크게 구분할 수 있다. 병역의무 수행 기간 동안의 복무 효과를 높일 수 있도록 환경을 개선하는데 주안점을 맞춘 연구는 손봉기(2009), 최은영 (2009)이 다양한 관련 연구를 소개했으므로 이를 참고하기 바란다.

병사들을 적정 보직에 할당하는 방법을 통해 병역 의무 수행 효과를 극대화하는 연구는 최적화 기법을 활용한 연구와 데이터마이닝을 활용한 연구로 구분할 수 있다. Klingman and Phillips (1984), Sweetser et al. (2001)은 최적화 기법을 활용하여 다양한 상황에서 설정된 목표를 달성할 수 있는 최적 인원 할당계획을 도출할 수 있는 방법론을 제시했다. 한편 민규식 외 (2004), 정재범 (2009)은 데이터마이닝 기법으로 군의 특기를 분류하는 방법론과 특정 부대에 관심병사가 집중되지 않도록 신병을 배치하는 방법을 제안했다. 한편 데이터마이닝에 대한 소개와 다양한 기법은 박우창 외 (2000)과 Shmueli et. al. (2006)에 자세히 소개되어 있다

본 연구는 데이터마이닝 기법을 활용하여 병사들을 특성에 따라 분류하고 이를 통해 특정 업무에 적합한 병사를 배치하여 숙련도를 높일 수 있는 방안을 제시한다. 특히 높은 숙련도가 요구되

는 항공정비병에 대해서 숙련도를 높일 수 있는 방안을 제시하여 본 연구의 활용성을 높이고자 한다. 이를 위해 2장에서는 항공정비병 양성 체계와 연구 방법론인 데이터마이닝에 대해서 살펴보고, 3장에서는 분석 대상에 대한 자료 수집 절차를 소개하고 연구 대상의 특징을 제시한다. 4장에서는 분석결과를 토대로 효과적인 무기정비병 배치 알고리즘을 제시한다.

## 2. 항공무기정비병과 분석방법론에 대한 이해

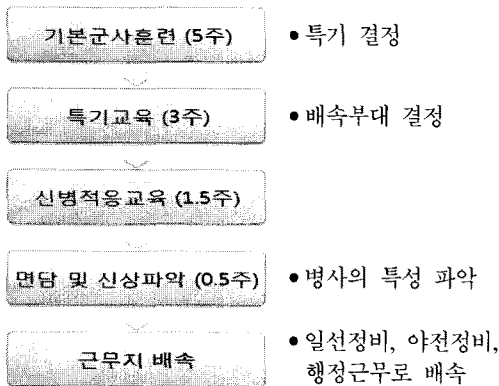
### 2.1 병사의 근무지 배속절차

공군을 지원한 병사의 근무지 배속 절차는 <그림 1>에 요약되어 있다. 이를 자세히 살펴보면 다음과 같다.

공군을 선택한 자원은 공군교육사령부 기본군사훈련단에서 5주간의 기본 군사훈련을 받게 된다. 공군병사의 특기는 군악병 등 특수한 분야에 한해 해당 분야 자격 구비자의 입대전 지원에 의해 결정되며, 이외에는 기본 군사훈련 수료시 자격증, 대학 전공과목, 적성검사(훈련 3주차)를 통해 결정된다. 기본 군사훈련을 수료하면 부여받은 특기별로 해당 특기학교에서 기초 특기지식을 배우게 되는데 특기학교 수료시 기본군사훈련 교육 성적(50%)과 특기학교 교육성적(50%)을 합산하여 성적순으로 희망부대에 배속된다. 해당부대에서는 전입 신병의 스트레스 해소와 부대 적응을 돕기 위해 신병들을 전입자 생활관에 수용하여 1~2주 동안 해당부대의 역사와 전통을 비롯한 간단한 소개와 종교시설, BX 등 편의시설, 면회절차 등에 대한 설명과 오리엔테이션 등을 수행하게 된다. 이상과 같은 단계를 거치면 신병들은 본격적으로 해당 대대 및 부서로 넘겨져, 지휘관과 면담 후 제대할 때까지 근무할 부서에 배속된다.

항공무기정비병의 근무지 배속절차도 이와 동일하며 주로 통신전자, 기계 전공자 등 공학계열

전공자가 항공무기정비병 특기로 분류된다. 항공무기정비병의 특기교육은 기술학교 항공무기정비 특기병 과정에서 이루어지며 항공기 및 무장계통에 관한 기초 업무지식과 실습, 안전교육을 3주 동안 수행한다. 기술학교 교육을 수료한 항공무기정비 특기 병사들은 99%가 비행단 및 비행전대로 배치된다.



〈그림 1〉 군무지 배속 절차

## 2.2 항공무기정비병 근무 유형

항공무기정비병의 주요 근무지는 근무환경 및 임무의 유사성에 따라 크게 일선정비, 야전정비, 행정근무로 나눌 수 있다.

일선정비는 항공기에 각종 폭탄, 확산탄, 로켓, 미사일, 기관포 등의 장/탈착과 무장발사계통 기능 점검 임무를 담당한다. 야전정비는 항공기에 장착되는 무장발사계통을 점검한다. 따라서 일선정비와 야전정비 임무는 항공기 무장계통에 적용되는 전기, 기계적 원리에 관한 지식과 도표, 기술 도서의 사용과 판독지식이 요구되며, 장비 및 공구작동 경험이 도움이 된다. 하지만, 일선정비는 실외에서 이루어지기 때문에 한여름 더위와 한겨울 매서운 추위를 견디며, 시끄러운 항공기 소음 속에 작업하여야 하고, 무거운 장비와 공구를 다루고 직접 폭탄을 들어야 하는 육체적 피로가 심한 작업이며 대부분의 임무는 4인 1조로 수행하

기 때문에 협동성과 희생정신이 요구된다. 반면, 야전정비는 실내에서 작업을 수행하여 더위와 추위 및 소음으로부터는 어느 정도 벗어나 있는 환경이지만, 작업등 아래에서 세심하게 항공기에 장착된 무장계통 주기 점검을 반복적으로 수행해야 하는 작업으로 다소 지루하게 느껴지는 면이 있으나 작업특성상 집중력을 유지해야 하는 작업이다. 행정근무는 간부를 도와 행정업무를 수행하는 업무로써 전산장비 활용능력과 함께 개인업무를 할당하여 수행하므로 책임감과 상하 동료간의 원만한 관계형성이 필요한 업무 특성을 갖고 있다.

근무지별 특성에 가장 적합한 병사를 배속하기 위해 전입병사의 성향을 신상자료, 면담 등을 통해 신중히 파악하여 가장 적합한 근무지에 배속할 필요가 있다. 특히 한 차례 근무지 배속이 결정되면 병사는 그곳에서 군 복무를 마쳐야 하므로 병사 배치는 해당 부서는 물론 병사 개인에게도 중요한 일이다.

## 2.3 데이터마이닝 기법

본 연구에서 방법론으로 활용한 데이터마이닝(Datamining)이란 자동화되고 지능을 갖춘 데이터베이스 분석 기법으로 90년대 초반부터 지식발견(KDD: Knowledge Discover Databases), 정보발견(Information Discovery), 정보수확(Information Harvesting) 등의 이름으로 소개되어 왔는데 일반적으로 ‘대량의 데이터 집합으로부터 유용한 정보를 추출하는 것’이라 정의된다 (Shmueli et al, 2006).

데이터마이닝은 데이터에 들어있는 의미를 파악해서 다른 사람들이 알 수 없는 정보를 해석하고 새로운 기회를 창출하는 역할을 하고 있으며, 이런 기능을 담당하는 사람들을 넘버크런처(NumberCruncher)라고 부르기도 한다. 데이터마이닝 기법으로 무장한 넘버크런처들은 이미 구글, 아마존 등의 기업에서 접속하는 사람들의 구매성향을

파악하고 이를 통해 구미에 맞는 광고나 상품을 소개하는 것에서부터 대본 단계에서 제작될 영화의 흥행여부를 사전에 예측하는 것까지 다양한 분야에서 활동하고 있다 (이언 에어즈, 2009).

군에서도 데이터마이닝 기법을 활용하고 있는데, 이라크에서 거리에 설치된 폭탄을 탐지하기 위해 촬영된 자료를 데이터로 변환하여 이전에 촬영을 통해 획득한 데이터와 비교한 후 변화를 분석하여 폭탄 설치 여부를 예측하는 시스템을 활용하고 있다 (Economist, 2010). 이렇듯 데이터마이닝 기법은 다양한 분야에서 활용되고 있으며, 그 유용성이 점차 증명되고 있다.

대부분의 상용화된 데이터마이닝 소프트웨어들에 포함되어 있거나 데이터마이닝 문헌에 자주 등장하는 데이터마이닝 기법들은 연관성 규칙발견 (Association Rule Discovery), 군집 분석(Cluster Analysis), 연결 분석(Link Analysis), 의사결정나무(Decision Tree), 인공신경망(Artificial Neural Network), 유전자알고리즘(Genetic Algorithm) 등이 있다.

이들 기법들은 목표변수가 존재하는가, 그렇지 않은가에 따라서 지도예측과 자율예측으로 분류된다. 지도예측은 목표변수가 존재하는 경우로 판별분석, 회귀 분석, 로지스틱 회귀분석, 시계열 분석 등이 포함된다. 자율예측은 목표값이 존재하지 않는 기법을 의미하는데, 군집 분석과 연관성 규칙 발견이 가장 대표적인 기법이라고 할 수 있다.

본 연구에서는 병사들의 유형을 분류할때 유형에 대한 사전정보가 명확하지 않으므로 다양한 가능성에 대해 고려할 수 있는 자율예측기법을 선택했다. 또한 자율예측의 대표적인 기법인 군집분석과 의사결정나무 분석을 사용했다. 군집분석을 통해 병사들을 특성이 유사하도록 군집화했고, 군집별로 어느 군무지에서 숙련기간이 가장 단축되는가를 비교하여 숙련급에 빨리 도달하는 군무지를 최적 군무지로 선택했다. 의사결정나무 분석은 병사의 몇 가지 특성만으로 해당 병사가 어느 군집

에 분류되는지를 확인할 수 있는 방안을 제시할 때 활용했다.

본 연구에서 활용된 군집분석과 의사결정 나무 기법의 알고리즘을 살펴보면 다음과 같다.

### 2.3.1 K-means 알고리즘

물리적 혹은 추상적 객체들을 비슷한 객체들의 클래스로 그룹화하는 것을 군집화 (Clustering)라고 한다. 군집은 같은 군집 내의 객체들과는 비슷하고, 다른 군집의 객체들과는 상이한 데이터 객체들의 집합이다. 군집화에 의해 자료의 전체적인 패턴과 데이터 속성들 사이의 흥미있는 상관관계를 찾을 수 있다.

군집분석 방법의 하나인 k-means 알고리즘은 입력 값으로 분류 그룹의 수를 k개로 취하고, 군집내 유사성은 높고 군집간의 유사성은 낮게 되도록 n개 객체들의 집합을 k개의 군집으로 분해하는 방법이다. 군집 유사성은 군집에서 군집의 무게중심으로 볼 수 있는 객체들의 평균값으로 측정한다.

k-means 알고리즘은 다음과 같이 진행된다. 우선 군집 평균이나 중심을 나타내는 값으로 객체들에서 k개를 임의로 뽑는다. 남겨진 객체는 군집 평균에 기초하여 가장 유사한 군집에 할당된다. 그리고 각 군집에 새로운 평균을 구한다. 이 과정을 반복하여 수행하게 되면 기준함수가 수렴하게 되는데, 일반적으로 식 (1)과 같이 정의되는 제곱오차(squared error) 기준이 기준함수로 사용된다.

$$E = \sum_{i=1}^k \sum_{p \in C_i} |p - m_i|^2 \quad (1)$$

$E$ 는 데이터베이스에서 모든 객체들의 제곱오차를 합한 것이다. 이때  $C_i$ 는  $i$ 번째 군집 집합을 나타내며,  $p$ 는 주어진 객체를 나타내는 공간의 점이고,  $m_i$ 는 군집  $C_i$ 의 평균이다 (박우창 외, 2000).

k-means 기법은 군집의 평균이 정의되어 있을 때에만 적용할 수 있고 데이터가 범주형의 특성을 포함하고 있는 때는 적용하기 어려울 수 있으며, 사용자가 군집의 개수 k개를 미리 정해야 하는 단점이 있으나, 알고리즘의 효율성이 뛰어나 군집분석에 널리 활용되는 기법이다.

### 2.3.2 CART(Classification And Regression Tree)

의사결정나무는 의사결정규칙을 나무구조로 도표화하여 분류와 예측을 수행하는 분석방법이다. 의사결정나무는 속성의 분리기준을 포함하는 내부마디와 최종분류를 의미하는 잎으로 구성된다. 의사결정나무는 비모수적 모형으로 선형성이나 정규성 등의 가정을 필요로 하지 않으며 이상치에 민감하지도 않다 (Shmueli et al, 2006).

본 연구에서는 의사결정나무 분석을 위한 알고리즘으로 CART(Classification And Regression Tree)를 사용했다. CART는 지니지수(Gini Index, 범주형 목표변수인 경우 적용) 또는 분산의 감소량(연속형 목표변수인 경우 적용)을 이용하여 이진분리를 수행하는 알고리즘이다.

지니지수는 불순도 (Impurity)를 측정하는 하나의 지수이다. 지니지수는 n개의 원소 중에서 임의의 원소를 추출했을 때, 추출된 원소가 다른 그룹으로 오분류 될 확률을 의미한다.

$P(i)$ 를 각 마디에서 한 개체가 i 번째 범주에 속할 확률이라고 하고, r을 분류하고자 하는 목표변수의 범주 개수라고 하면 지니지수는 다음 식 (2)과 같이 정의된다.

$$G = \sum_{j=1}^r P(j)(1-P(j)) = 1 - \sum_{j=1}^r \left(\frac{n_j}{n}\right)^2 \quad (2)$$

(n : 마디에 포함되어 있는 관찰치 수,  $n_j$ : 목표변수 j번째 범주에 속하는 관찰치 수)

일반적으로 CART는 범주형 목표변수에 대해

서는 지니지수를 분리기준으로 사용한다. 즉 지니지수를 가장 많이 감소시켜주는 입력변수와 그 때의 최적분리 기준에 의해서 자식마디가 형성된다.

CART 알고리즘은 지니지수를 최대로 감소시켜주는 예측변수와 그 변수의 최적분리를 자식마디로 선택하여 의사결정 나무를 구성한다.

## 3. 자료 수집 및 분석 자료 특성

### 3.1 자료 수집 및 분석 절차

분석에 사용된 데이터는 2010년 7월 현재, 공군비행단 및 전대에 근무하고 있는 항공무기정비특기 병사들 중 숙련급에 도달하였다고 판단되는 병사 395명과 그 병사들과 중대이하 단위에서 가까이 근무하는 관리자들을 대상으로 설문조사하였다. 한편 숙련급 여부는 지휘관의 판단에 근거하여 분류했다.

관리자들은 2년 동안 병사와 함께 임무를 수행하면서 자연스럽게 병사의 성격과 장단점 등을 파악할 수 있다. 하지만 기본 군사훈련을 마친지 얼마 되지 않은 신병의 특성을 단기간내에 파악하는 것은 쉽지 않다. 따라서 관리자가 갖 전입해 온 신병의 특성을 근무지 배속 전까지 면밀히 파악하기 힘들 것이라 여기고 병사의 특성에 관한 설문은 병사가 답하도록 하였다. 반면 병사의 숙련급 도달기간과 특기별로 병사들이 숙련급에 도달하는 평균 기간은 관리자가 답하도록 하였다.

입력변수에 대한 문항 선정은 키, 형제유무 등 객관적 자료를 위주로 하였으나, 좀 더 정확한 병사 특성을 파악하기 위해 병사의 숙련급 도달기간에 영향을 미치는 요인을 다음과 같은 기준으로 선별하였다.

첫째, 전투병과의 경우 숙련도와 학력과의 상관관계를 입증하기 어려우나, 전투지원 병과의 경우는 다소 상관관계를 보인다 (김경수, 2005)는 사실에 근거하여 병사들의 전공학문과 입대전 업무

경험 모두를 입력변수에 포함하여 조사해 보았다.

둘째, 육체적인 업무가 많은 항공무기정비병의 업무 특성상, 신체조건이 우수한 병사가 유리할 것으로 생각되어, 키와 몸무게, 과거 질병경험 등을 조사대상에 포함하였다.

셋째, Robert (2010)는 ‘각 근무지의 유형은 구별되는 특징을 지니는 성격 유형’이라 하였다. 즉, 근무지별 직무를 수행하는 데 어떤 과제가 요구되는지를 안다면, 그 직업을 수행하는 데 필요한 성격 유형도 알 수 있고, 마찬가지로 수행해야 할 주요 과제가 무엇인지 안다면, 과제와 관련된 성격 유형을 지닌 직원을 둘 수 있다고 하였다. 따라서 각 근무지에 적합한 다양한 성격을 파악하기 위해 외향성, 낙천성, 친밀성, 협조성, 치밀성, 상하관계를 묻는 질문을 포함하였다.

넷째, 형제 수가 사회성과 상관관계가 있을 수 있는 가능성과 ‘혈액형에 따른 적합한 직업’ 등의 논의에 근거하여 병사들의 형제자매 유무와 혈액형을 변수로 채택하였다.

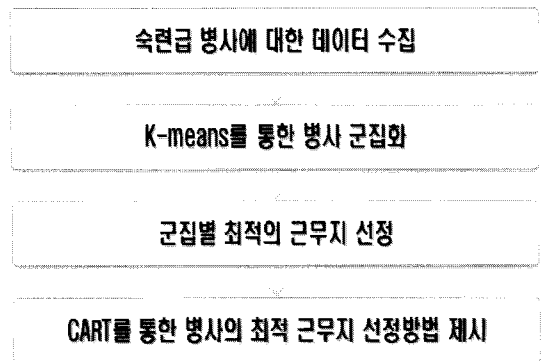
다섯째, 자존감은 성격이나 성취와 밀접한 상관관계가 있으므로, 병사들이 본인을 어떻게 평가하는지와 타인이 자신을 어떻게 평가할 것이라고 느끼는지를 파악하였다.

이외에 입대 전 공군에 대한 정보를 얼마나 알고 있는가? 공군 입대동기는 무엇인가? 거주지역의 도시형태는 어떠한가? 여자친구가 있는가? 흡연을 하는가? 훈련단에서 기본군사훈련을 마친 심정은 어떠한가? 지금(근무지 결정전)의 심정은 어떠한가? 로 구성된 38가지 항목을 설문에 포함하였다.

목표변수는 대상자의 근무지와 업무 숙련급 도달기간으로써 도달기간 측정은 감독자 평가를 통해 이루어졌다. 즉, 대상 병사를 가까이에서 지켜봐온 직속 상관이 측정토록 하였으며, 평가자에 따라 기준이 흔들리지 않도록 부대별로 1명이 측정토록 하였고 평가자가 대상자들을 평가하면서 평가기준이 바뀌지 않도록 하기 위해 평가자들은

공군에서 정하고 있는 교육도달목표를 참조토록 하였다.

이렇게 수집된 자료를 토대로 다음과 같은 절차로 자료를 분석했다. 먼저, 공군 항공무기정비병의 데이터를 수집하여 병사들을 입력변수별 특성에 따라 여러 군집으로 나누고, 각 군집에 속한 병사들이 어느 근무지에서 가장 빨리 숙련급에 도달하는지를 비교하여 숙련급에 가장 빨리 도달하는 근무지를 각 군집별 최적 근무지로 선택하였다. 다음으로 395명을 자신이 속한 군집별 최적 근무지에 배속하였다면 숙련급 도달기간이 현재보다 얼마나 단축될 수 있는지 계산하였다. 이런 결과를 바탕으로 실무에서 전입 신병을 최적 근무지에 배속하는 데 활용할 수 있도록 의사결정 나무를 구성하였다. <그림 2>는 본 연구에서 수행한 자료분석 절차를 나타낸 것이다.



<그림 2> 자료 분석 절차

한편, 데이터 분석은 R을 활용하였다. R은 통계적 컴퓨팅과 그래픽스를 위해 개발된 언어이다. R에는 모든 분야의 자료를 처리할 수 있는 통계적 분석도구가 있다. SAS나 SPSS와 같은 통합 통계 패키지에도 다양한 통계 분석 도구가 있어서 통계 분석을 할 수 있다. 그러나 다른 통계분석 소프트웨어와 비교할 때 R은 사용자 인터페이스가 다소 떨어지는 측면이 있지만 자료 처리 속도나 공개된 소스를 활용한 분석의 편의성이 높아 R을 활용했다.

### 3.2 수집자료의 특성

앞 절에서 설명한 자료를 토대로 분석을 수행한 결과는 다음과 같다. 공군비행단 및 비행전대에 근무하는 항공무기정비병 중 숙련급에 도달되었다고 판단되는 병사로서 총 395명의 특징은 <표 1>과 같다. 조사대상자의 근무지별 분포는 일선정비 233명, 야전정비 65명, 행정근무 97명이며, 근무지별 숙달소요기간은 일선정비가 14.12개월로 가장 길고, 전체 조사대상자들의 평균은 13.52개월이었다.

표 1. 숙련급 도달 기간

구분	일선정비 (233명)	야전정비 (65명)	행정근무 (97명)	전체 (395명)
평균	14.12개월	13.45개월	12.11개월	13.52개월

395명의 자료를 토대로 다음과 같은 단계에 따라서 군집분석을 수행했다. 먼저 군집 수  $k$ 를 결정하고, 각 군집에 군집중심을 1개씩 할당한다. 다음 단계로 모든 데이터를 가장 가까운 군집 중심에 배속시켜서 초기 군집을 구성하며, 다음 단계로 각 군집에 배속된 데이터로 새로운 군집 중심을 계산하고, 군집 중심이 변화가 없을 때까지 군집을 계속 바꾸는 과정을 반복한다.

일반적으로 군집 분석시에 군집수 결정은 주어진 학습 데이터와 사전경험에 의해 분석가가 주관적으로 결정한다. 그러나 데이터에 대한 정확한 정보가 없는 상태에서 다수의 개체수와 속성수를 갖는 데이터에 대해 주관에 의존해 군집수를 결정하는 것은 쉬운 작업이 아니다. 이에 대한 해결책으로 군집수가 증가함에 따라 거리의 합이 어떻게 감소하는지 알아보는 방안을 직관 대신 활용할 수 있다 (박민재 외, 2003). 따라서 수집된 데이터에서 군집수 변화에 따라 거리의 합이 어떻게 변하는지를 살펴보았다.  $k$ -means 기법의 객체간 거리는 식 (3)과 같이 정의된 유클리드 거리 (Euclidean Distance)를 사용하며, 군집내 객체간의 거리평균은 군집 중심으로부터 군집내 모든 객체간 거리의 합을 평균한다.

dean Distance)를 사용하며, 군집내 객체간의 거리평균은 군집 중심으로부터 군집내 모든 객체간 거리의 합을 평균한다.

$$d(C_i, C_{i'}) = \sqrt{\sum_{j=1}^p (x_{ij} - x_{i'j})^2} \quad (3)$$

( $x_{ij}$ :  $i$  군집의  $j$ 번째 속성의 좌표)

<표 2>를 보면 군집수가 증가함에 따라 군집내 객체간 평균거리가 감소함을 알 수 있다. 이는 군집수가 증가함에 따라 군집내 객체간의 유사성이 더 증가하는 것으로 해석할 수 있다. 그리고 군집수가 3에서 4로 증가할 때, 객체간 평균거리가 다른 군집수 변화에 비해 급격히 줄어드는 현상을 관찰할 수 있고 이후 변화량이 작음을 확인할 수 있다. 즉 군집수로 4를 선택하는 것이 3을 선택하는 것보다 군집분류의 정확성을 크게 향상시키지만, 이후로는 군집수를 늘리더라도 군집내 객체간의 유사성이 크게 증가하지는 않음을 알 수 있다.

표 2. 군집수 변화에 따른 거리

군집수	2	3	4	5	6	7
평균 거리합	0.904	0.883	0.805	0.792	0.788	0.781

군집수를 크게 잡는 것은 결과 해석에 어려움이 있고 불필요하게 군집이 많아져 분석의 효율성이 떨어진다. 가령 병사들을 고졸과 대학 재학생으로 분류하려면 2개 군집이 적당한데, 3개 군집으로 하여 고졸, 서울소재 대학 재학생, 지방소재 대학 재학생으로 구분하는 것은 비효율적이다.

<표 2>의 결과를 토대로 군집수는 4 또는 5에서 선정하는 것이 바람직하다. 그러나 추가 분석 결과 군집수를 4로 선택하면 숙련급 도달기간 단축이 1.4개월이지만, 군집수를 5로 선택하면 숙련급 도달기간 단축이 1.8개월로 나타나 군집수를 5로 선택했다.



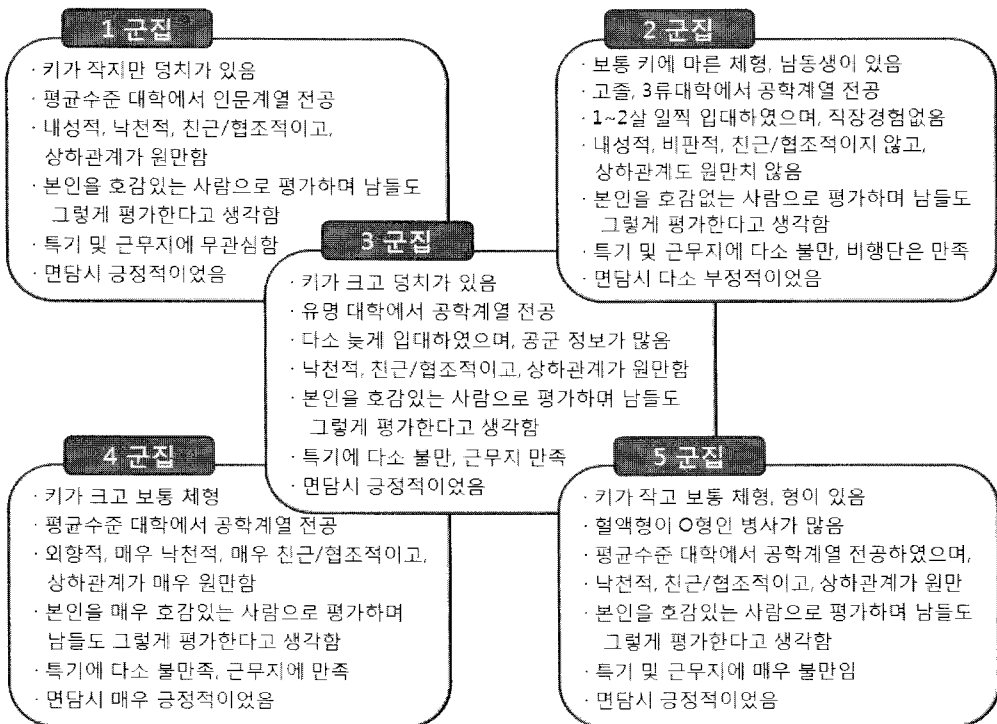
#### 4. 그룹(군집)별 특성과 최적 배치 알고리즘

앞 장에서 설명한 바에 따라 5개 집단을 기준으로 분석한 결과 도출한 각 군집별 특성은 <그림 3>에 제시되어 있다. <그림 3>의 군집 특성은 군집분석의 군집별 특징을 정성적(Qualitative) 언어로 표현한 것이다.

1군집은 키가 작고 인문계열 전공이며, 내성적이지만 친근하고 상하관계도 원만하다. 2군집은 내성적이고 비판적이며 친근하거나 협조적이지도 않다. 또한 본인을 호감이 없는 사람으로 부정적인 평가를 하였다. 3군집은 키가 크고 유명대학에서 공학계열을 전공하였고 긍정적이다. 4군집은 키가 크고 외향적이며 매우 긍정적이고 대인관계도 매우 원만하다. 5군집은 키가 작고 낙천적이며 긍정적이다. 특히 혈액형이 O형인 병사가 69.7%로 많았다.

각 군집에 속하는 일선정비, 야전정비, 행정에 근무하는 병사들의 숙련급 숙달속도를 평균속도를 기준으로 비교한 결과가 <표 3>이다. 즉, 전체 평균에 비해서 숙련급 도달 기간이 더 걸리면 양의 값을 가지고 짧으면 음의 값을 가진다.

<표 3>에서 1군집에 속한 병사 중 일선정비에 근무하는 병사의 숙련급 도달기간은 평균보다 0.91개월 빠르고, 야전정비에 근무하는 병사는 0.36개월 늦고, 행정에 근무하는 병사는 0.5개월 늦다는 것을 확인할 수 있다. 따라서 1군집에 속한 병사들은 일선정비에 배치했다면 다른 곳에서 근무한 것에 비해 빠르게 숙련급에 도달하였을 것이라고 예상할 수 있다. 마찬가지로 2군집은 행정, 3군집은 야전, 4, 5군집은 일선에 배치하였다면 숙련급 도달기간을 단축할 수 있었을 것이다.



<그림 3> 군집의 특성

표 3. 군집별 숙련급 도달 기간

(단위 : 개월)

부서	1군집 (65명)	2군집 (51명)	3군집 (127명)	4군집 (86명)	5군집 (66명)
일선	-0.91	1.64	1.95	-2.90	-0.83
야전	0.36	3.19	-2.38	-1.69	-0.50
행정	0.50	-2.35	0.13	-1.45	1.94

병사들의 군집별 특성인 <그림 3>을 항공무기 정비 장교(대위 14명) 및 부사관(원/상사 11명)에게 보여주고, 각 군집별로 어느 근무지에 배속하는 것이 가장 적합한가를 설문한 결과, 1군집은 일선에 배속하는 것이 바람직하다는 응답이 가장 많은 48%로 나타났으며, 이는 <표 3>에서 1군집은 일선에서 가장 빨리 숙련급에 도달한다는 연구결과와 부합되는 결과다. 마찬가지로 4군집에 대한 적절한 근무지도 일선으로 응답하였으며, 이 또한 분석결과와 부합된다. 관리자들에게 물어본 군집특성 별 최적 배치에 관한 설문 결과는 <표 4>에 요약되어 있다.

1, 4 군집과 달리 2, 3, 5군집에 대한 설문결과는 연구결과와 상당한 차이가 있었다. 응답자들은 적극적이고 명량한 병사는 일선에 배치하고, 내성적이고 대인관계가 좋지 않은 병사는 야전에 배속하는 것이 바람직하다고 응답했다. 따라서 2군집 병사들은 내성적이고 대인관계가 좋지 않은 성향을 나타내므로 야전에 배속하는 것이 적절하다고 응답하였다. 그러나 분석 결과는 2군집 병사 중 야전에 근무하는 병사들은 평균보다 3.2개월 늦게 숙련급에 도달하는 것으로 나타났다.

이처럼 설문에 응답한 관리자들은 병사의 적정 근무지 배속에 분석결과와 인식 차이를 나타내고 있었다. 특히, 내성적이고 대인관계가 좋지 않은 병사들을 업무가 떨어지는 것으로 인식하고 ‘업무능력이 떨어지므로 야전으로 보내야 한다’는 선입견을 가지고 있었다. 이런 선입견은 야전에 배속된 병사들을 더욱 위축시켜 업무 숙달 속도가 더

욱 늦추는 요인으로 작용할 가능성도 있다. 왜냐하면 같은 2군집 병사들이 야전정비에 배속되지 않고 행정부서에 배속된 경우는 다른 어느 군집에 속한 병사보다 더 빨리 숙련급에 도달하는 것으로 나타나기 때문이다.

<표 4> 관리자 시각에서 본 최적 배치

구 분	1군집	2군집	3군집	4군집	5군집
일선 배속	48%	16%	60%	52%	22%
야전 배속	8%	72%	4%	16%	39%
행정 배속	44%	12%	36%	32%	39%

<표 5>는 군집별로 병사들을 숙련급에 가장 빨리 도달할 수 있는 근무지에 배속시켰을 때, 숙련급 도달기간을 산정한 것이다. 즉, 일선정비에 1, 4, 5 군집 병사를 배치하고, 야전정비에는 3군집, 행정근무는 2군집을 배치하여 숙련 도달 시간을 살펴보면 일선정비는 1.55개월, 야전정비는 2.38개월, 행정근무는 2.35개월 기간이 단축된다. 그 결과 전체적으로 설문대상 항공무기정비병의 숙련급 도달기간은 1.79개월 단축되어 11.73개월로 나타났다.

<표 5> 최적 배치시 숙련급 도달 기간

(단위 : 개월)

구 분	일선정비	야전정비	행정근무	전체
적용 전	14.12	13.45	12.11	13.52
적용 후	12.57	11.07	9.76	11.73
단축	-1.55	-2.38	-2.35	-1.79

본 연구를 통해 항공무기정비특기 병사의 특성이 다양함을 알 수 있었다. 이러한 다양한 특성을 지닌 병사들은 비슷한 특성에 따라 5개 군집으로 나눌 수 있었고, 각 군집별로 어느 근무지에서 근무하느냐에 따라 숙련급 도달기간이 차이가 남을 발견하였다. 그 결과를 활용하여 신병 전입시 병사의 특성을 파악하여 5개 군집 중 해당 군집에

분류하고 그가 속한 군집에 따라 근무지를 배치하면 현재보다 개선된 병사의 숙련 수준을 유지할 수 있으며, 이는 전투력 수준 향상으로 이어질 수 있다. 그러나 군집분석의 단점인 해석의 어려움으로 인해, 전입 신병이 어느 군집에 해당되는가를 명확하게 구분하는 것은 쉽지 않다. 따라서 그 대안으로 의사결정 나무 기법을 신병의 최적 근무지를 결정할 수 있는 규칙을 제시하기 위한 방안으로 활용했다.

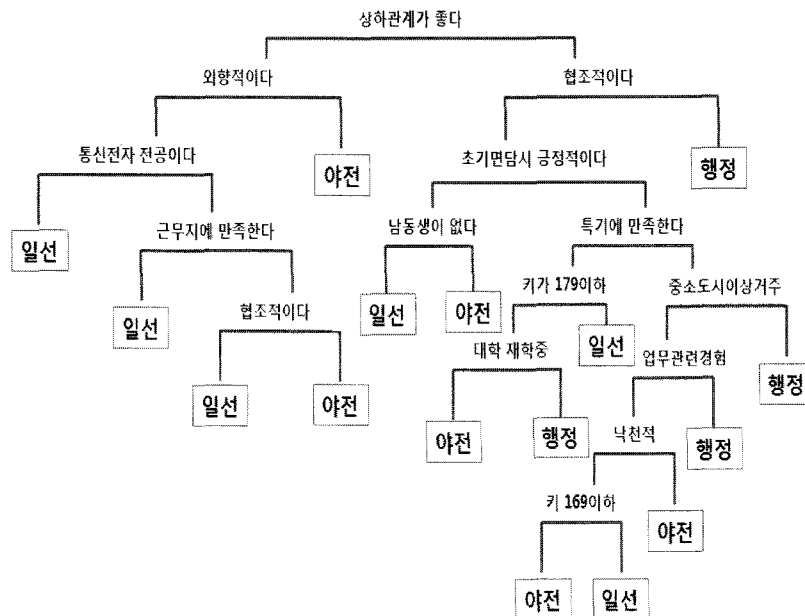
의사결정나무의 입력변수는 군집분석의 입력변수와 같고, 목표변수를 각 군집으로 하면 쉽게 결과를 얻을 수 있다. 의사결정나무는 3장에서 설명한 CART 알고리즘을 활용했으며, 분리기준으로 지니 지수 (Gini Index)를 활용했다. 의사결정나무 형성은 395명의 데이터 중 무작위 방법으로 90%의 Train data로 모형을 만들고, 나머지 10%는 Test data로 활용하여 만들어진 모형을 검증하였다. 이를 100회 반복하여 검증한 결과, 오분류율은 6.7%였다. 즉, 의사결정나무를 통해서 신병의 특성에 따른 해당군집으로 할당하였을 때, 100

명중 7명 정도가 잘못 분류될 가능성이 있다는 것이다. 이렇게 만들어진 모형은 <그림 4>와 같다. 모형에서 보듯이 최대 8단계만에 신병의 근무지를 결정할 수 있다. 즉, 병사를 분류할때 상관관계를 기준 분류하고 그 다음에 성격이 외향적이나 또는 협조적이나 여부로 분류하여 차례로 병사의 특성을 파악해서 분류한다면 병사에게 적합한 근무지가 일선인지 야전인지 행정인지 여부를 판단할 수 있다.

<그림 4>에서 제시된 의사결정나무를 신병배치 전 면담을 통해 확인한 정보와 더불어 활용한다면 신병의 특성(군집)을 결정할 수 있으며, 이를 기초로 신병을 배치한다면 숙련급 도달 시간을 현재 보다 단축시킬 수 있다.

## 5. 결론

공군 항공무기정비병 395명의 현재 숙달급 도달기간은 13.52개월이다. 본 연구에서는 이를 병사들의 특성에 따라 5개 군집으로 나누어 분석한



<그림 4> 의사결정 나무 분석 결과

결과 각 군집은 근무지별로 숙련급에 도달하는 기간이 달랐다. 따라서 각 군집의 최적 근무지를 숙련급에 가장 빨리 도달하는 곳으로 선정한다면 숙련급 도달 시간을 현재보다 단축할 수 있다는 것을 제시했다. 또한 신병 전입시 신병의 주요 데이터들만을 활용하여 신속히 최적 근무지에 배속할 수 있도록 의사결정 나무를 만들었다.

본 연구는 지금까지 수행되지 않았던 병사들의 최적 근무지 선정에 관한 과학적인 접근으로 병사의 숙련급 도달 기간을 단축시킬 수 있는 대안을 제시하였고, 이는 부대 내 숙련급 병사의 비율을 상승시켜 전투력 향상과 병사들의 개인 특성에 맞는 근무지 배속으로 자신감과 성취감을 고취시켜 군 복무 만족도를 자연스럽게 증가시킬 것으로 예상된다.

본 연구에서는 숙련급 도달 병사를 대상으로 설문조사를 수행하여 병사의 특성을 파악하고 연구하였다. 그러나 본 연구는 신병의 근무지 배속에 관한 연구이므로 신병을 대상으로 하는 자료를 수집하고 이들에 대한 추적연구를 수행한다면 보다 정확한 연구결과를 도출할 수 있을 것이다. 또한 본 연구에서는 군집분석은 k-means 방법, 의사결정나무는 CART 방법을 활용하였으나, 데이터 마이닝의 다른 기법들을 활용하고 결과를 비교해 본다면 강건한 (Robust) 의사결정 방안을 제시할 수 있을 것으로 판단된다.

## 참고문헌

- [1] 권희면 (1992), “현역병 복무기간의 적정성 여부 검토”, 한국국방연구원.
- [2] 김경수, “징집병 복무기간의 결정요인에 관한 연구”, 석사학위논문, 국방대학교.
- [3] 정주성 (2003), “한국 병영정책의 바람직한 진로”, 한국국방연구원.
- [4] 민규식, 정지원, 최인찬 (2004), “데이터 마이닝 기반의 군사특기 분류방법론 연구”, 한국국방경영분석학회지: 30(1), 1-14.
- [5] 박민재, 전성혜, 오경환 (2003), “붓스트랩 기법과 유전자 알고리즘을 이용한 최적 군집 수 결정”, 퍼지 및 지능시스템학회지: 13(1), 12-17.
- [6] 박우창, 승현우, 용환승, 최기현 역 (2000), 데이터 마이닝 개념 및 기법, 자유아카데미.
- [7] 손봉기 (2009), “성격적 강점이 군 생활만족도에 미치는 영향”, 석사학위논문, 서울대학교.
- [8] 이언 에어즈 (2009), Super Cruncher, 북하우스.
- [9] 정재범 (2009), “통계적 클러스터링을 활용한 새로운 신병 분류 방법”, 석사학위논문, 국방대학교.
- [10] 최은영 (2009), “신세대 병사의 군 부적응에 관한 연구”, 석사학위논문, 국방대학교.
- [11] Klingman, D. and V. Phillips (1984), "Topological and Computational Aspects of Preemptive Multi-criteria Military Personnel Assignment Problem," Management Science:30(11).
- [12] Hogan, R.(2007), Personality and the Fate of Organization, Taylor & Francis Group LLC.
- [13] Sweetser, A., K. Olson, M. Grabski, K. Miller, A. Seise and S. Sanvorn, “An Analysis of Personnel Distribution Options for the Chief of Staff, Army,” Phalanx: 34 (3), 12-15.
- [14] Shmueli, Galit and Patel, Nitin R. and Bruce, Peter C. (2006), Data Mining for Business Intelligence, John Willey & Sons.
- [15] Economist, "Special Report: Data Deluge", Economist Feb 27th 2010.

---

■ 저자 소개 ■

**강 규 영**(E-mail: kakeyo@hanmail.net)

2000 공군사관학교 졸업(학사)  
2010 국방대학교 운영분석학과 졸업(석사)  
현재 공군군수사령부 항공기술연구소  
관심분야 Datamining, Queueing Theory, Military O.R

**윤 봉 규**(E-mail: bkyoon1@gmail.com)

1996 연세대학교 졸업(경영학사)  
1998 한국과학기술원 산업공학 석사  
2002 한국과학기술원 산업공학 박사  
현재 국방대학교 운영분석학과 부교수  
관심분야 Performance Optimization & Innovation, Stochastic Models  
& Queueing Theory, Military O.R

<주요저서 / 논문>

- Analysis of the Queue-length Distribution for the discrete-time batch-service Geo/G/1/K Queue(E.J.O.R., To appear)
- On the Queue Length Distribution for the GI/G/1/K/Vm Queue(Stochastic Analysis and Applications, 2004.4)
- An Invariance in the Priority Queue with Generalized Server Vacations and Structured Batch Arrivals(Stochastic Analysis and Applications, 2003.7)
- Waiting-time Distribution of a Discrete-time Multiserver Queue with Correlated Arrivals and Deterministic Service Times : D-MAP/D/k System(Operations Research Letters, 2002.7)