

# 앙상블 기법을 활용한 무인항공기 수리부속 수요예측 연구

김재동<sup>1</sup>, 유지훈<sup>1</sup>, 최정호<sup>2\*</sup>

<sup>1</sup>한국국방연구원, <sup>2</sup>광운대학교 방위사업학과(박사과정)/육군 소령

## A demand forecasting using ensemble technique for UAV spare parts

Jae-Dong Kim<sup>1</sup>, Ji-Hoon Yu<sup>1</sup>, Jung-Ho Choi<sup>2\*</sup>

<sup>1</sup>Korea Institute For Defense Analyses

<sup>2</sup>Department of Defense Business, KwangWoon University(PhD Candidate), Republic of Korea Army

**요약** 최근 전략무기체계인 무인항공기에 대한 세계적 관심이 높아지고 있다. 이러한 세계적 흐름에 발맞춰 한국군도 무인항공기 개발을 통해 전투준비태세를 유지하기 위해 노력하고 있다. 무인항공기를 효과적으로 운용하기 위해서는 수리부속 보급 및 정비가 필수적이다. 이를 위해서는 수리부속 수요를 정확히 예측하는 것이 중요하다. 그동안 한국군은 장비정비정보체계의 데이터를 활용하여 몇 가지 시계열 방법을 기반으로 수행하여 왔으나 수리부속 수요예측의 정확성을 향상시키기 위해 개선된 모델이 필요하다. 본 연구에서는 머신러닝 방법 중 앙상블 방법을 활용하여 수리부속 수요예측의 정확도를 높일 수 있는 모델을 제안했다. 머신러닝 방법은 시계열 방법에 비해 많은 정보를 활용하므로 수리부속 수요예측의 정확도를 높일 수 있다. 이를 위해 6년간의 무인항공기 수리부속 소모 데이터를 수집하여 수요예측 모델을 제시하였다.

**Abstract** A number of states have focused considerable attention on the strategic value of UAV as a military weapon system. South Korean military has also been making efforts to maintain combat readiness through the development of unmanned aerial vehicles. However, proper and accurate forecasting of maintenance accessories procurement is required for UAV development and deployment, which can eventually contribute to combat readiness. For a long time, the ROK army has used the data set in the equipment maintenance information system to forecast the demand of maintenance subordinates estimated by the several time series method. Nevertheless, more accurate forecasting methods need to be established and implemented. In this study, we applied the ensemble method for military industry demand forecasting, which is performed by the maintenance department. Additionally, we proposed a model to improve the accuracy of the prediction. Since the machine learning method utilizes more information than the time series method, it is expected to improve the accuracy of spare part future demand forecasts. To verify the effectiveness of this novel method, we implemented the demand forecasting classification model utilizing the consumption data of UAV maintenance accessories collected for a period of six years.

**Keywords** : Spare Parts, Demand Forecast, Data Mining, Deep Learning, Logistics

---

\*Corresponding Author : Jung-Ho Choi(KwangWoon University, Republic of Korea Army)

email: junghochoi486@gmail.com

Received December 19, 2022

Revised January 11, 2023

Accepted February 3, 2023

Published February 28, 2023

## 1. 서론

드론(Drone) 또는 무인항공기(UAV: Unmanned Aerial Vehicle)는 실제 조종사가 직접 탑승하지 않고 지상에서 사전에 프로그래밍 된 경로에 기반을 두거나, 원격으로 통제하여 자동 또는 반자동으로 비행하는 비행장치를 일컫는다[1]. 이와 같은 특징으로 드론의 중요성은 점점 증대되고 있다. 따라서 무인항공 장비는 고장 정비에 따른 신속한 수리부속 교체 및 수리가 필요하며, 정비사는 적시에 부대 및 야전정비를 실시하여 장비의 고장을 사전에 예방함으로써 전투준비태세를 제고하여야 한다[2]. 수리부속 확보를 위해서는 X-1년에 해당 수리부속에 대한 수요예측을 통해 해당 수리부속 조달을 위한 계획 수립 및 예산확보가 선행되어야 한다. 이를 위해 한국군은 장비정비정보체계(DELIIS: Defense Logistics Integrated Information System 이하 DELIIS)의 소모 데이터를 활용하여 수요예측을 수행하고 있다[3].

현재 한국군은 시계열 방법을 중심으로 산술평균법, 이동평균법 등 5~8개 방법을 활용하여 필요한 수리부속에 대한 수요예측의 정확도를 높이기 위해 노력하고 있다[4]. 여러 해 동안 관리되고 있는 소모 데이터를 활용하고 인간의 방법을 벤치마킹하여, 목표 가동률을 제고시킬 수 있다. 2장에서는 예측 방법론 관련 연구를 알아보고 3장에서는 데이터 수집 및 변수 추출과 수리부속 수요예측 모델을 설명하고 4장에서는 방법별 결과에 대한 비교 및 분석 결과를 제시하며 마지막 5장에서는 논문 결과와 향후 연구 방향을 제시한다.

## 2. 문헌연구

### 2.1 시계열 방법

시계열 데이터는 시간의 흐름에 따라 식별된 데이터로 시계열 분석을 통해 과거에 관측된 데이터를 바탕으로 미래의 데이터 예측이 가능하다. 시계열 방법은 산술평균법(AMM: Arithmetic Mean Method 이하 AMM), 단순이동평균법(SMA: Simple Moving Average 이하 SMA), 가중이동평균법(WMA: Weighted Moving Average 이하 WMA), 선형이동평균법(LMA: Linear Moving Average 이하 LMA), 최소자승법(LSM: Least Square Method 이하 LSM) 등의 5가지 방법을 활용한다[4].

AMM은 수요를 예측할 때 과거자료를 이용하여 발생한 수요를 모두 평균하여 예측하는 것으로 과거자료가 충분히 많고, 균등하게 형성될 경우에 활용되는 방법이다. LSM은 전체 자료의 추세를 이해하는데 가장 많이 사용되며 관측치와 오차를 최소화하여 예측하는 방법이다[5].

시계열 예측이 활용되는 사례는 아래와 같다. 기존 연구에서는 한국의 국내총생산 시계열 자료를 분석하여 2년간의 미래 예측에서 비관측요인모형이 보다 우수함을 확인하였고[6] 시계열 데이터를 활용하여 코로나19 동향을 예측하여 미래의 질병에 대해서 방역과 방역수칙의 중요함을 강조하였다[7]. 또한 여러 가지 시계열 모형을 활용하여 최대 전력수요를 예측하고 정확도를 높이는 방법을 연구하는 등 공공분야에서도 선행연구가 이루어지고 있다[8].

### 2.2 머신러닝 방법

기계 학습(Machine Learning)은 측정된 데이터를 컴퓨터가 학습하고 학습된 데이터를 분석하는 용도로 많은 분야에 적용되고 있다. 머신러닝은 데이터의 입력 및 출력정보가 모두 있는 지도 학습과 입력학습데이터만 있는 비지도 학습으로 분류하며 다양한 방법이 있다[9].

먼저, 의사결정나무(DT: Decision Tree 이하 DT)는 의사결정 규칙을 나무구조로 도표화하여 효과적인 분류와 예측을 가능케 하는 데이터 마이닝 방법 중의 하나이다. 기존 연구에서는 기업들의 효율적인 열처리 선택을 위해 DT 모델을 골격으로 델파이 방법과 AHP(Analytic Hierarchy Process) 방법을 사용하여 열처리선택 의사결정 모델을 개발하였다[10].

랜덤 포레스트(RF: Random forest 이하 RF)는 과적합(overfitting) 문제와 같이 기존의 나무 모델(Tree Model)의 한계를 개선하기 위해 앙상블 방법으로, 기존의 훈련 데이터에서 추출한 부트스트랩 샘플(Bootstrap Samples)로 다양하게 나무 모델을 형성하는 방법이다. 즉, 여러 나무 모델을 형성할때 임의성(Randomness)을 부여하여 상호 연관성이 없는(Uncorrelated) 나무들로 구성하고 이들의 분류 결과나 회귀 결과를 종합하여 예측력을 향상시키는 개념이다[11]. 기존 연구에서는 기상청에서 제공되는 강수량과 도시 유역에 대한 2차원 침수에 대한 결과, 그리고 기계학습의 하나인 회귀모델을 활용하여 실시간으로 도시 유역에 대한 침수지도를 예측할 수 있는 방법을 제시하였다[12].

서포트 벡터 머신(SVM: Support Vector Machine 이하 SVM)은 데이터를 고차원(High Dimension)의 형상 공간(Feature Space)으로 사상(Mapping)하여 최적의 의사결정 영역(Decision Boundary)을 구해내는 방법이다. 최근 연구에서는 교각의 안전진단을 위한 데이터를 머신러닝 알고리즘으로 학습시켜 혼동행렬을 통해 5개의 머신러닝을 비교하여 안전진단 예측에 가장 적합한 알고리즘을 도출하였다. 그 결과 SVM이 가장 적합한 알고리즘으로 제시되었다[13].

선형회귀(LR: Linear Regression 이하 LR)는 연관성 있는 각 변수들이 제시되면 변수들의 관련성을 분석한 후 결과를 예측하는 방법이다[14]. 기존 연구에서는 무기체계 학습모형으로 예측되는 우선순위와 성능개량 형상의 주요 수리부속 400개에 대해 품목담당관들이 판단한 우선순위를 비교하여 학습모형에 대한 예측 정확도를 평가한 결과, 80개의 오차범위를 기준으로 Ridge 모형에서 86%가 일치하는 것을 확인했다[15].

머신러닝에서 양상블의 의미는 여러 개의 분류 모델을 생성하고, 그 예측을 결합하여 더욱 정확한 예측 결과를 도출하는 방법이다[16]. 여러 개의 모델은 같거나 다른 알고리즘을 사용할 수 있고, 동일한 데이터셋(data set)을 중복 허용 샘플링 또는 중복 불가 샘플링으로 학습을 시킬 수 있다[17]. 양상블 방법 중에서 다수결분류는 다른 알고리즘을 조합하여 투표를 통해 결정하는 방식이다. 학습기의 예측 결과 중 가장 많이 선택되는 클래스를 예측 결과로 사용하는 하드 보팅(HV: Hard Voting 이하 HV) 방식과, 예측 결과들의 평균을 최종 예측 결과로 사용하는 소프트 보팅(SV: Soft Voting 이하 SV) 방식으로 구분된다[18].

### 3. 수리부속 수요예측 모델 제안

#### 3.1 데이터 수집 및 변수 추출

군에서 활용되고 있는 DELIIS는 2009년에 개발을 시작하여 2020년 재개발된 시스템으로 다음의 Table 1과 같이 보급과 정비에 대한 자산을 관리하고 있다. 본 시스템을 이용하여 편성 부대로부터 국방부에 이르기까지 정보관리시스템을 통합 운영하고 있다. 본 연구에서는 육군의 전략자산 중 대대급 주력기인 무인항공기를 연구 대상 장비로 선정하였다. 육군 DELIIS의 정비테이블에서 53,422개 트랜잭션 데이터를 수집하였고, 수집된 데이

Table 1. Army DELIIS Table Example

Equipment Name	Equipment NSN	Equipment registration Number	maintenance date	NSN	troops
Equipment A	A0001	2001A1	20130514	A54321	X LSC
Equipment C	C0001	2001C1	20210320	D54321	X battalion
...	...	...	...	...	...
Request Unit	consumption Unit	consumption	worker	number of workers	recoverability
EA	EA	10	0 0 0	3	N
EA	EA	5	0 0 0	3	Y
...	...	...	...	...	...

터에는 정비날짜, 정비당일 소모된 수리부속품목의 개수, 대당 구성수, 표준단가, 조달기간, 비고(정비내역) 등 22개 항목이 포함되어 있다.

본 연구의 목적은 무인항공기의 수리부속별 정확한 수요예측이고 관련 수리부속 품목수는 총 146개였다. 이 과정에서 수리부속별로 2017년부터 2022년까지 6개년 간의 연도별 데이터로 정형 데이터 변수를 추출하였다. 먼저 정형 데이터는 소모 개수를 추출하였고, 최종적으로 활용한 변수는 아래의 Table 2와 같이 11개의 변수이다. 목표변수는 품목별 2017년 수요발생 여부 데이터로서 146개의 품목 중 10,510건에 대해서 수요가 발생하였다. 수리부속 전체 품목 수는 146품목이며, 수요가 발생한 품목이 91품목, 수요가 발생하지 않은 품목이 55품목이다. 데이터 불균형을 해소하기 위해 오버샘플링(Oversampling) 방법을 활용하여 기존 수요가 발생하지 않은 품목에 36품목을 포함하였다. 따라서 수리부속 발생/미발생 품목 개수의 합은 182품목을 활용하여 실험하였다.

Table 2. Feature Descriptions

	Feature	Means
independent variable	Maintenance Troop 1	Binary [0,1] data (Whether or not maintenance occurs in the maintenance unit)
	Maintenance Troop 2	
	Maintenance Troop 3	
	Maintenance Troop 4	
	Maintenance Troop 5	
	2017 year Consumption	Spare Parts By Item/Unit/Year Sum of consumption
2018 year Consumption		
2019 year Consumption		
2020 year Consumption		
2021 year Consumption		
dependent variable	Sum of Consumption	Consumption from 2022 year (Binary)

### 3.2 모델 수립

본 연구에서는 아래의 Fig. 1과 같이 앙상블 모델을 기반으로 소모 데이터 모델과 부대를 포함한 데이터 모델 두가지를 제시하였다. 모델 1은 연도별 소모 개수만을 적용한 모델이며, 모델 2는 연도별 소모개수와 정비 부대에서의 수리부속별 발생여부 정보를 포함하였다. 분류방법은 시계열 방법, 머신러닝 방법을 활용하였다. 시계열 방법은 AMM, SMA, WMA, LMA, LSM의 5가지 방법을 활용하였으며, 머신러닝 방법은 DT, RF, KNN, NB, SVM, LR, HV를 사용하였다. 분석 프로그램은 Python 3.7.4를 사용하였으며, 활용된 프레임워크는 scikit-learn이다. 보다 신뢰성있는 값을 제시하기 위하여 데이터셋을 무작위(random)로 추출하여 모델을 구축하여 실험하였다.

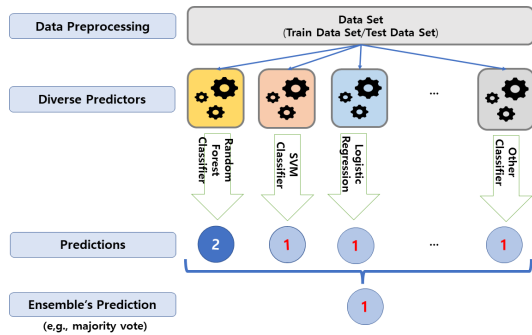


Fig. 1. Ensemble Model Concept

## 4. 실험 분석 및 결과

### 4.1 모델 수립

모델 평가 방법은 Table 3과 같이 오차행렬(Confusion Matrix)을 사용하였다. 따라서 세부 평가방법은 Accuracy, Recall, Precision, F1-Score 정확도를 측정하였다. 혼

Table 3. Confusion matrix

		Predicted	
		Y (Positive)	N (Negative)
Actual	Y (Positive)	A (True Positive)	B (False Negative)
	N (Negative)	C (False Positive)	D (True Negative)
Accuracy = $\frac{A+D}{A+B+C+D}$		Recall = $\frac{A}{A+B}$	
Precision = $\frac{A}{A+C}$		F1-Score = $\frac{2(Precision \cdot Recall)}{(Precision + Recall)}$	

련 데이터 셋(training data set)과 테스트 셋(test data set)은 7:3의 비율로 나누어 10겹 교차 검증(10-fold cross validation) 방식을 사용하였다.

### 4.2 모델 결과

모델 1과 모델 2의 정확도 기준으로 모델 성능을 평가한 결과 아래의 Table 4와 Table 5 및 Fig. 2와 Fig. 3과 같다. 모델 1은 시계열과 머신러닝 방법 결과를 제시하였다. 먼저, 시계열 방법에서는 AMM과 SMA이 다른 방법에 비해서 72%로 더 높은 정확도를 보였다. 머신러닝 방법에서는 DT가 72%로 가장 높은 정확도를 보였다. 시계열 방법보다 머신러닝 방법이 전체적으로 높은 정확도를 보였다.

Table 4. Model 1 Time series and machine learning results

		Accuracy	Recall	Precision	F1-Score
time series	AMM	72%	85%	72%	75%
	SMA	72%	85%	72%	75%
	WMA	68%	89%	68%	75%
	LMA	71%	88%	74%	74%
	LSM	62%	89%	62%	77%
Machine Learning	DT	72%	80%	73%	74%
	RF	69%	85%	70%	74%
	KNN	70%	85%	71%	74%
	NB	70%	85%	71%	74%
	SVM	62%	100%	63%	77%
	LR	62%	100%	63%	77%
	HV	69%	85%	70%	74%

모델 2는 모델 1과 달리 정비부대 정보가 들어가 있어서 DT를 제외한 모든 모델이 우수한 결과를 보였다. DT를 제외한 모든 모델이 71%를 넘었으며, HV모델은 가장 높은 75%의 정확도를 보였다. 이는 VOTE 분류기 학습이 다른 방법들 보다 모든 Accuracy, Recall, Precision, F1-Score의 정확도 측면에서 우수하였다. 이는 VOTE 방법론의 특성인 각 분류기의 결과들을 집계하여 가장 많은 표를 얻은 클래스를 최종 예측값으로 분류하는 방식의 특징이 잘 적용된 모델이라고 볼 수 있다.

모델 1과 모델 2의 결과를 비교해보면, 모델 1은 단순히 연도별 소모 데이터만을 적용하였기 때문에 시계열 데이터 분석 방법을 적용한 결과나 머신러닝 방법을 사용한 결과가 유의미한 차이를 보이지 않는다. 그러나 모델 2에서는 연도별 데이터와 함께, 연도별 소모된 정비 부대 발생 정보를 추가하여 머신러닝 방법을 사용한 데

이터에서는 DT를 제외한 모든 결과에서 정확도가 1~10% 정도 향상된 결과를 얻을 수 있었다. 모델 2에서 시계열 방법의 결과가 제외된 이유는 시계열 방법은 변수를 추가하여도 결과의 변화가 없으므로 모델 1의 결과와 동일하기 때문이다.

결과적으로 실제 정비 데이터가 추가됨에 따라 머신러닝에서 정확도가 제고되었음을 확인할 수 있었고 그 중에서 양상블 기법이 가장 좋은 결과를 나타냄을 확인할 수 있었다.

Table 5. Model 2 Machine learning result using oversampling technique

	Accuracy	Recall	Precision	F1-Score
DT	69%	89%	69%	75%
RF	71%	88%	71%	76%
KNN	72%	88%	72%	76%
NB	71%	88%	71%	76%
SVM	72%	88%	72%	76%
LR	72%	88%	72%	76%
HV	75%	87%	75%	78%

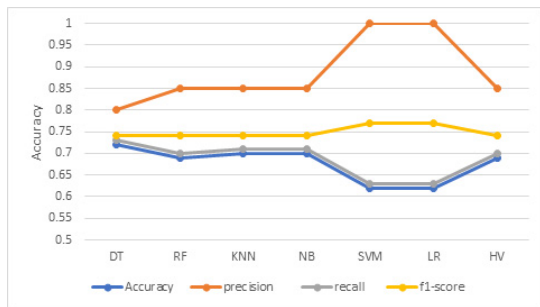


Fig. 2. Comparison of Model 1 Results

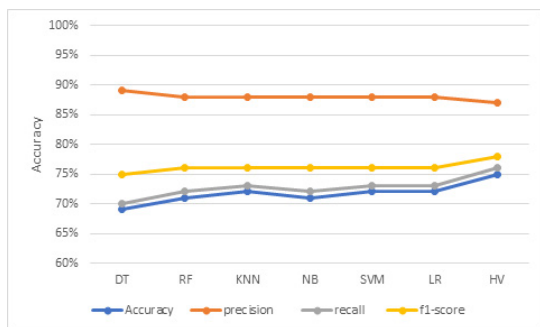


Fig. 3. Comparison of Model 2 Results

## 5. 결론

본 논문에서는 육군에서 운용 중인 대대급 무인 항공기의 수리부속 수요예측 정확도 제고를 위하여 지난 6개년의 수리부속 수요 데이터를 분석하고 머신러닝 방법을 활용하여 수리부속 수요예측 모델을 제안하였다. 그 결과 여러 머신러닝 방법론 중 HV 다수결 분류 방법이 75%로 가장 높은 정확도를 보였다. 제안한 모델은 기존 시계열 방법중 AMM과 SMA 72%에 비해서 약 3% 향상된 정확도를 나타냈다는 점에서 논문의 의미가 있다. 하지만 본 방법론은 수요예측 발생여부인 분류 문제로 정확도를 제시하였으므로 한계가 존재한다. 따라서 수량 예측이 가능한 모델로 추후 연구가 필요하다. 또한, 정비 특성 정보나 장비의 운항거리, 사용연수 등의 정형 및 비정형 데이터를 활용한다면 본 연구 결과보다 높은 수요예측 정확도를 기대할 수 있다.

## References

- [1] H. J. Kim, "A Monitoring Platform that Recommends Customized Flight Paths for UAVsin a Mixed Reality", Dankook University, pp. 4, Dec. 2021. <http://www.riss.kr/link?id=T16076119>
- [2] J. S. Kim and D. K. Koh and J. W. Chung, "Accuracy Improvement Research for Lumpy Aircraft Spare Parts Demand Forecast Based On Machine Learning", Korean Society of Management Science, pp. 1-11, Sep. 2019. DOI: <https://doi.org/10.7737/KMSR.2019.36.3.001>
- [3] J. D. Kim and H. J. Lee, "A study on predictive model for forecasting anti-aircraft missile spare parts demand based on machine learning", Journal of the Korean Data & Information Science Society, Vol.30(3), pp. 587-596, May 2019. DOI: <https://doi.org/10.7465/jkdi.2019.30.3.587>
- [4] Cryer, Jonathan D. Time series analysis. Vol. 286. Boston: Duxbury Press, 1986.
- [5] H. T. Kim and S. H. Kim, "Data mining based army repair parts demand forecast", Journal of the Korean Data & Information Science Society, Vol. 30(2), pp. 429-444, Mar. 2019. DOI: <https://doi.org/10.7465/jkdi.2019.30.2.429>
- [6] S. B. Chon and S. G. Lee, "Covid19 trends predictions using time series data", Journal of the Korean Data & Information Science Society, Vol. 22(5), pp. 829-837, Sep. 2011.
- [7] J. H. Kim and J. Y. Kim, "Covid19 trends predictions using time series data", Journal of the Korean institute of

information & Communication engineering, Vol. 25(7), pp. 884-889, Aug. 2021.

[8] H. T. Kim and S. H. Kim, "Time Series Predictions for Electric Power Demand", Chungbuk National University, pp. 1~3, Aug. 2021.

[9] H. S. Yim, "SOH estimation of LiFePO<sub>4</sub> battery management systems using a linear regression analysis", Ajou university, Feb. 2022.  
<https://dspace.ajou.ac.kr/handle/2018.oak/20709>

[10] S. Y. Park, "A Study on the Determination of the Heat Treatment Methods of STD11 Using a Decision Tree Model", Dong-A University, Feb. 2021.

[11] S. H. Jeong, "Estimation of the office rent using the machine learning methods : focusing on the use of random forests, artificial neural networks, support vector machines", Dept. of Applied Economics Graduate School of Hanyang University ,PP. 8~30, Feb, 2020.  
<https://repository.hanyang.ac.kr/handle/20.500.11754/123619>

[12] H. I. Kim, Y. S. Lee and B. H. Kim, "Real-time flood prediction applying random forest regression model in urban areas", JOURNAL OF KOREA WATER RESOURCES ASSOCIATION, .vol. 54, pp. 1119-1130, Dec. 2021.  
DOI: <https://doi.org/10.3741/JKWRA.2021.54.S-1.1119>

[13] D. H. Yoon, "Algorithm for prediction pier safety diagnosis using machine learning", University of science&technology, Sep. 2021.

[14] Lee, S. J., "Development on AI Education Program of Prediction System Based on Linear Regression for Elementary School Students", Daegu Nation University of Education, 48, Aug. 2022.

[15] S. C. So, "Estimation of acquisition priority for concurrent spare parts using shrinkage regression", Department of Defense Fusion Engineering The Graduate School of Yonsei University, Dec. 2018.

[16] Y. J. Hwang, "Verification of the Validity of MRC Models Using Deep Learning Methodology", Namseoul University, pp. 13, Jun. 2021.

[17] K. M., "Gas collection and methods for application of machine learning and ensemble techniques for simulating gas diffusion", Korea Polytechnic University, pp. 5, Jun. 2020.

[18] J. M. Lee, "An Integrated Ensemble Technique for Improving Performan", Kumoh National Institute of Technology, pp. 3~4, Dec. 2018.

김 재 동(Jae-Dong Kim)

[정회원]



- 2005년 2월 : 한양대학교 산업공학과 (공학사)
- 2007년 2월 : 한양대학교 산업공학과 (공학석사)
- 2020년 2월 : 고려대학교 산업경영공학과 (박사수료)
- 2007년 5월 ~ 현재 : 한국국방연구원 연구위원

<관심분야>

데이터마이닝, 최적화, 시뮬레이션, 국방정책

유 지 훈(Ji-Hoon Yu)

[정회원]



- 2006년 12월 : 미국 해군대학원 안전보장학과 (안전보장학석사)
- 2013년 8월 : 미국 시라큐스대 정치학과 (정치학박사)
- 2021년 1월 ~ 2021년 12월 : 해군사관학교 군사전략학 교수
- 2022년 1월 ~ 현재 : 한국국방연구원

<관심분야>

해양안보, 한미동맹, 외교/안보, 데이터마이닝

최 정 호(Jung-Ho Choi)

[정회원]



- 2004년 3월 : 육군3사관학교 환경공학과 (공학사)
- 2013년 2월 : 경희대학교 경영학과 (경영학석사)
- 2017년 8월 : 광운대학교 방위사업학과 (박사수료)
- 2004년 3월 ~ 현재 : 육군 소령

<관심분야>

드론, 방위사업, 데이터마이닝