한국군사과학기술학회지 제16권 제5호, pp. 675~682, 2013년 10월

학술논문 체계공학 부문

지지벡터회귀분석을 이용한 무기체계 신뢰도 예측기법

A Reliability Prediction Method for Weapon Systems using Support Vector Regression

나 일 용*
Il-Yong Na

ABSTRACT

Reliability analysis and prediction of next failure time is critical to sustain weapon systems, concerning scheduled maintenance, spare parts replacement and maintenance interventions, etc. Since 1981, many methodology derived from various probabilistic and statistical theories has been suggested to do that activity. Nowadays, many A.I. tools have been used to support these predictions. Support Vector Regression(SVR) is a nonlinear regression technique extended from support vector machine. SVR can fit data flexibly and it has a wide variety of applications. This paper utilizes SVM and SVR with combining time series to predict the next failure time based on historical failure data. A numerical case using failure data from the military equipment is presented to demonstrate the performance of the proposed approach. Finally, the proposed approach is proved meaningful to predict next failure point and to estimate instantaneous failure rate and MTBF.

Keywords : Support Vector Regression(지지벡터회귀분석), Support Vector Machine(지지벡터머신), Reliability Growth Analysis(신뢰도 성장분석), Reliability Prediction(신뢰도 예측), NHPP(비동질포아송과정)

1. 서 론

무기체계의 운용단계 신뢰도 분석은 주로 초도 배치 후 장비 운용 안정화 여부 식별, 형상변경, 예방점검시기 조정 등 군수지원성 최적화와 다양한 개선활동 효과 확인에 사용된다. 또한, 무기체계의 노화정도를 식별하여 창정비시기, 도태시점 등 정책적 판단이 요구되는 사안에 대한 의사결정에 활용되기도 한다. 미국 등 선진국에서는 이미 30여 년 전에 Duane[1], AMSAA 모형[2,3] 등의 다양한 신뢰도 성장분석 모형을 개발한 바 있으며, 지속적인 개량활동을 통해 다양한 획득 환경 하에서 관리, 평가 모형으로 사용 중이다. 민수 분야에서도 역시 시스템 차원은 아니나 핵심 구성품에 대해 신뢰도 성장 분석에 기반한 신뢰도 예측을 수행 중이다.

이러한 신뢰도 성장분석 및 예측 활동에는 몇 가지 통계적 기법을 사용해 왔다. AMSAA는 Weibull 고장강도함수(Failure Intensity Function)기반의 비동질 포아송과정(NHPP : Non Homogeneous Poisson Process)를 사용하였고, Ho.& Xie[4,5]는 시계열 분석모형인 ARIMA

† 2013년 6월 7일 접수~2013년 9월 13일 게재승인
* 국방기술품질원(DTaQ)
책임저자 : 나일용(mongguri@dtaq.re.kr)

와 MFNN(Multi-layer Feed-forward Neural Network)을, Wei-Chiang et al.[6]는 지지벡터머신(SVM : Support Vector Machine)을 사용한 바 있다. 가장 최근에는 Yuan Fuqing[7]이 지지벡터회귀분석(SVR : Support Vector Regression)을 이용하여 광산에서 운용하는 LHD (Load-Haul-Dump) 차량의 신뢰도와 미래고장시점을 예측하였다. 이러한 연구 결과는 신뢰도 예측분야의 추세가 기존의 확률과정, 시계열 등 전통적인 확률/통계적 방법에서 신경망이론, SVM 등 점차 인공지능(A.I. : Artificial Intelligence) 기법 분야로 확장되고 있음을 보여준다. 이에 본 논문에서는 비교적 최근 신뢰도 예측분야에 적용이 시도되고 있는 지지벡터머신과 지지벡터회귀분석을 이용한 신뢰도 예측모델을 제안하며, 궤도장비 고장자료에 적용하여 모형의 실효성을 확인한다.

2. 지지벡터회귀분석

지지벡터회귀분석은 Vapnic이 최초로 제안한 기법으로, 임의의 입력공간(Input Space, x)을 입력특성벡터(Input Feature Vector)에 대한 식 (1)과 같은 비선형 회귀식 $f(x)$을 통해 2개 이상의 공간으로 분할하는 곡선 탐색 과정이다. 이 과정에서 필요에 따라 커널(Kernel) 함수로 입력특성공간을 $\psi(x)$로 변환하여 공간분할 함수를 입력공간이 아닌 특성공간에 대해 산출하기도 한다. 지지벡터회귀분석은 회귀식 추정 시 Fig. 1과 같이 일정 허용오차(ϵ)를 고려, 오차 허용범위($f(x)\pm\epsilon$)를 벗어난 차이(ξ_i)에 대해서만 벌과금(Penalty Cost)을 부여하는 방식을 취한다. 이런 방식을 통해 공간분할선 추정시 유연성을 제공하며, 타 회귀분석과 차별화 된다. 추정된 $f(x)$는 입력공간을 몇 개의 공간으로 분할할 수 있기 때문에 판별분석이나 패턴인식 등에 널리 활용된다.

$$f(x)=w\psi(x)+b \tag{1}$$

지지벡터회귀분석과 지지벡터머신은 비선형회귀식 $f(x)$를 탐색하는 과정은 동일하나 지지벡터머신은 $f(x)$에 의해 분할된 공간에, 즉, $f(x)$를 이용한 판별분석에 관심을 두는 분석법이며, 지지벡터회귀분석은 임의의 입력값 z에 대한 결과값 $f(z)$ 활용에 관심을 두는 분석법이다. $f(x)$ 도출을 위한 원(Primal) 문제는 식 (2)과 같으나, 이를 쌍대문제(Dual)로 변환 할 경우, 훨씬 간단한 형태인 식 (6), 식 (7)로 문제를 풀 수 있다[9].

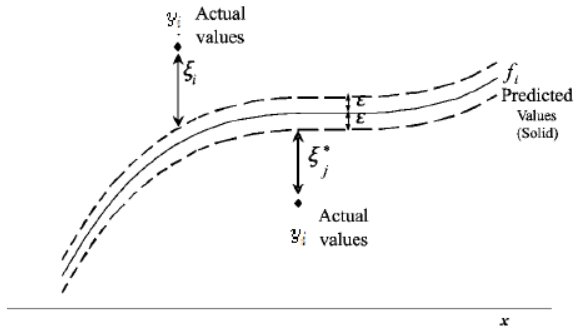


Fig. 1. The parameters used in SVM, SVR(Vojislav[9])

$$\begin{aligned}&\min \frac{1}{2}\|w\|^2+C\sum_{i=1}^{m}(\xi+\xi_i^*) \\ &s.t.\, f(x_i)-y_i \leqq \epsilon+\xi_i \\ &\quad y_i-f(x_i) \leqq \epsilon+\xi_i^* \\ &\quad \xi_i \geqq 0,\ \xi_i^* \geqq 0,\ i=1,\ 2,\ 3,\ ...,\ m\end{aligned} \tag{2}$$

C : 벌과금(penalty cost)
ϵ : 허용오차(loss function의 tube size)
y_i : x_i에 대한 실제 결과값(actual value)
ξ_i, ξ_i^* : distance between actual value and tube

이를 쌍대(Dual)문제로 변환하기 위해 라그랑지함수를 식 (3)과 같이 정의할 수 있고, 사용된 쌍대변수 α, α^*, η_i, η_i^*는 식 (4)를 만족해야 한다.

$$\begin{aligned}L=&\frac{1}{2}\|w\|^2+C\sum_{i=1}^{m}(\xi+\xi_i^*) \\ &-\sum_{i=1}^{m}\alpha(\epsilon+\xi_i-f(x_i)+y_i) \\ &-\sum_{i=1}^{m}\alpha^*(\epsilon+\xi_i^*-y_i+f(x_i))-\sum_{i=1}^{m}(\eta\xi_i+\eta^*\xi_i^*)\end{aligned} \tag{3}$$

$$\alpha_i,\ \alpha_i^*,\ \eta_i,\ \eta_i^* \geqq 0 \tag{4}$$

K.K.T Condition[10]에 의해 L의 w, b, ξ_i, ξ_i^*에 대한 편미분값은 모두 0이어야 하므로, 식 (5)를 얻을 수 있으며,

$$\frac{\partial L}{\partial b}=\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)=0$$
$$\frac{\partial L}{\partial w}=w-\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)\psi(x_i)=0 \quad (5)$$
$$\frac{\partial L}{\partial \xi^{(*)}}=C-\alpha_i^{(*)}-\eta_i^{(*)}=0$$

식 (5)를 식 (3)에 대입, 식 (6)의 쌍대문제를 얻는다.

$$\begin{aligned}\max\quad & W(\alpha,\alpha^*)= \quad (6)\\ & -\epsilon\sum_{i=1}^{m}(\alpha_i+\alpha_i^*)\\ & +\sum_{i=1}^{m}(\alpha_i^*-\alpha_i)y_i\\ & -\frac{1}{2}\sum_{i,j=1}^{m}(\alpha_i^*-\alpha_i)(\alpha_j^*-\alpha_j)\,\psi(x_i)\cdot\psi(x_j)\end{aligned}$$
$$s.t.\sum_{i=1}^{m}(\alpha_i-\alpha_i^*)=0$$
$$\alpha_i,\ \alpha_i^*\in[0,C]$$
$$i=1,\ 2,\ 3,\ \ldots,\ m$$

이 때, 쌍대문제에서 목적함수에 더 이상 η_i, η_i^*가 더 이상 나타나지 않음을 알 수 있으며, 식 (5)의 두 번째 조건과 식 (1)로부터 식 (7)을 얻을 수 있다. 즉, 비선형 회귀식은 식 (7)의 형태를 가지며, 식 (6)을 만족하는 α, α^*를 파라미터로 가지게 된다.

식 (6)과 식 (7)에서 $\psi(x_i)\cdot\psi(x_j)$를 Kernel함수 $K(x_i,\ x_j)$로 대체할 경우, 최종적인 회귀식은 식 (7)이 된다.

$$f(x)=\sum_{i=1}^{m}(\alpha_i^*-\alpha_i)K(x,\ x_i)+b \quad (7)$$

통상 커널(Kernel)함수는 다항식 커널(Polynomial Kernel, 식 (8))과 유클리드 거리에 기반한 가우시안 커널(Gaussian Kernel, 식 (9)) 두 종류가 있으나, 다항식 커널은 입력 특성벡터의 크기에 영향을 받으므로, 표준화(Nomalization)가 필요하고, 둘 중 하나의 그룹에 속하는지 여부를 판단하는 분류에 적합[11]하지 않은 단점이 있다.

$$K(x,x')=(\langle x,x'\rangle+1)^d \quad (8)$$
d : 임의의 정수

$$K(x,x')=\exp(-\frac{\|x-x'\|^2}{2\sigma^2}) \quad (9)$$
σ : width of gaussian function

이상의 논리로 기계학습 과정에서는 쌍대문제 최적해 α, α^*를 이용, 최적 파라미터 C, ϵ, σ를 결정한다.

3. Modeling

제안하는 모형은 Fig. 2와 같이 3단계로 구성한다.

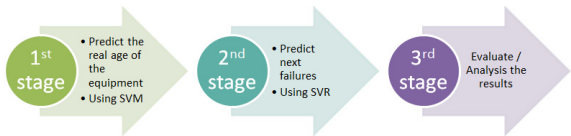


Fig. 2. Flowchart of reliability prediction

통상 동일한 시기에 배치된 장비는 같은 고장 추세를 나타냄을 가정하고 분석한다. 그러나 동일한 시기에 배치되었더라도, 운용환경, 사용자, 사용률 등에 따라 고장 추세가 다를 수 있으므로, 분석에 사용되는 장비들을 실제 동일집단으로 간주할 수 있는지 여부가 검토되어야 한다. 기존 AMSAA 모형에서는 이를 위해 고장추세 예측에 앞서 CVM(Cramer Von Mise) 검정[2,3]으로 분석에 사용될 장비의 동일집단 여부를 확인한다. 이와 유사하게 본 논문에서 제시될 모형도 대상장비의 고장추세 예측에 앞서 해당 장비가 어떤 환경에서 운용된 장비들과 유사한지 여부를 확인할 수 있는 방법이 필요하였다. 고장추세 예측에 사용되는 모형이 확률모형을 이용한 모형이 아니므로, 고장추세 구분은 운용환경, 고장 개소, 심각도 등을 입력특성벡터(Input Feature Vector)로 하는 지지벡터머신을 사용한다. 이 후 같은 장비그룹에 속한 장비들은 같은 고장발생 경향을 나타냄이 확인되었으므로, 해당 장비 집단의 특성을 이용하여 미래 고장시점과 신뢰도를 예측한다. 단계별 특성벡터의 구성, 학습 및 예측 방법은 '가.', '나.'와 같다.

가. 1단계(고장추세 구분단계)

본 단계에서는 지지벡터머신을 이용하여 분석 대상 장비의 고장자료가 실제로 어떤환경과 운용량을 가지는 장비군과 유사한 고장추세를 나타내는지 확인한다.

1) 특성벡터 구성

특성벡터는 식 (2)에서 x에 해당되는 부분이며, 분석에 사용될 특성벡터의 구성은 아래와 같다. 총 5개 분류의 Input과 1개의 Output으로 구성되며, Table 1과 같다. 입력 특성벡터는 장비의 운용 기간 중 발생한 고장정보를 종합한 것이며, 총 고장수, 총 운용기간, 총 운행거리, 고장개소별 고장수와 고장의 심각도에 따른 고장수로 구성된다. 이에 따른 출력 특성벡터(Output Feature Vector)는 장비군(운용부대, Deploy로 표현)으로 정의하였다.

Table 1. Feature vector for 1st stage(partial)

Input Feature									Output Feature
Total Fail	Total Op	Total Run	Failure Location			Mission Effectiveness			
			AA	BB	···	F1	F2	F3	
157	2175.96	4199.1	3	7	···	2	60	95	1
152	2126.06	4085.7	3	9	···	2	57	93	1
128	1923.00	3857.1	5	3	···	1	40	87	1

2) 지지벡터머신 학습 및 고장추세 구분

지지벡터머신 학습에는 '2.'절에서 설명한 바와 같이 제한점을 가지는 다항식 커널 대신 가우시안 커널을 사용하며, 식 (2)를 만족하는 최적 파라미터 C, ϵ, σ를 기계학습을 통해 도출한다. 이후 대상장비의 특성벡터를 학습된 지지벡터머신에 입력하여 입력된 고장추세와 가장 유사한 고장추세를 나타내는 장비군을 찾아낸다. C, ϵ, σ간의 상관관계의 복잡성으로 각 파라미터의 전역 최적화는 어려운(NP-hard) 문제임이 타 논문을 통해 확인된 바 있으므로[12] 학습은 임의의 범위 내에서 Grid Search 방법을 사용하였고, 결과가 입력된 C, ϵ, σ의 경계값(Boundary)이 아닌 경우 최적으로 판단한다.

나. 2단계(고장추세 예측단계)

본 단계에서는 지지벡터회귀분석을 사용하여 고장추세, 즉, 미래 고장시점들을 예측한다. 전술한 바와 같이, 동일 장비군의 장비는 동일한 고장 추세를 가지므로, 장비군 내에서 가장 많은 고장수를 가진 장비가 기준(Reference) 장비가 되며, 이 장비의 고장에 대해 학습된 지지벡터회귀선이 타 장비의 고장시점 예측에 사용된다. 모형 학습에 사용된 특성벡터와 학습방법은 아래와 같다.

1) 특성벡터 구성

분석에 사용될 자료는 고장시점의 시계열 자료로 1차원이므로, 입력 특성벡터로 사용가능하도록 동일한 차원을 가지는 특성벡터로 변환이 필요하다. i번째 배치된 장비 중 기준 장비의 고장시점을 t_{i1}, ..., t_{ik}, ..., t_{in}이라 할 때, Y. Fuqing[7]이 사용했던 것과 동일한 방식의 변환을 통해 동일 차원의 특성벡터 Set 생성이 가능하며, $lag = m$일 경우, Table 2와 같은 Feature Vector를 얻을 수 있다. Table 2에서 n^{th} 고장시점은 $(n-m)^{th} \sim (n-1)^{th}$고장시점을 Input으로 하는 값이며, 결국 전체 Feature Vector는 $s_{i,n-m}$으로 표기할 수 있다.

Table 2. N-step lagged input array and its corresponding output[11]

Series	x(input)					y(output)
s_{i1}	x_{i1}	x_{i2}	···	$x_{i,m-1}$	x_{im}	$x_{i,m+1}$
s_{i2}	x_{i2}	x_{i3}	···	x_{im}	$x_{i,m+1}$	$x_{i,m+2}$
···	···	···	···	···	···	···
$s_{i,n-m}$	$x_{i,n-m}$	$x_{i,n-m+1}$	···	$x_{i,n-2}$	$x_{i,n-1}$	x_{in}

2) 지지벡터 회귀식 학습

2단계 모형 학습은 기준 장비의 특성벡터 Set $\{s_{i1}$, s_{i2}, ..., $s_{ik}\}$을 입력으로 하며, 전체 자료의 2/3 지점부터 학습에 사용한다(단, k = 1, 2, ..., $n-m$). 자료의 개수가 많지 않은 경우 LOO(Leave-One-Out) 방법 혹은 n-fold CV(Cross Validation)을 사용하는 것이 일반적[13]이므로 본 논문에서는 5-fold CV 방법을 사용하여 각 고장시점까지의 MSE를 최소화 하여 $(n+1)^{th}$ 이후 고장을 추정할 수 있는 각 시점별 최적 파라미터 C, ϵ, σ를 도출한다. 최적값 탐색은 1단계와 동일한 Grid Search 방법을 사용한다.

3) 고장시점 예측

고장시점 t_{i1}, ..., t_{ik}, ..., t_{in}을 특성벡터 형태로 변환 후 '나. 2)'에서 학습된 고장시점별 지지벡터 회귀식 입력값으로 사용하여 $(n+1)^{th}$ 이후 고장 시점을 예측

한다.

다. 3단계(고장률, MTBF 예측단계)

만약 특정 시점 t에서의 순간 고장률(λ_{INST})과 순간 MTBF($MTBF_{INST}$)를 산출하고자 할 경우, 해당 시점 전, 후의 고장위치를 근거로, 식 (10), 식 (11)로 산출할 수 있다. 그림으로 표현하면, Fig. 3과 같다.

$$\lambda_{INST} = \frac{1}{t_{i+1} - t_i} \tag{10}$$

$$MTBF_{INST} = \frac{1}{\lambda_{INST}} \tag{11}$$

$$st.\ t_i < t < t_{i+1}$$

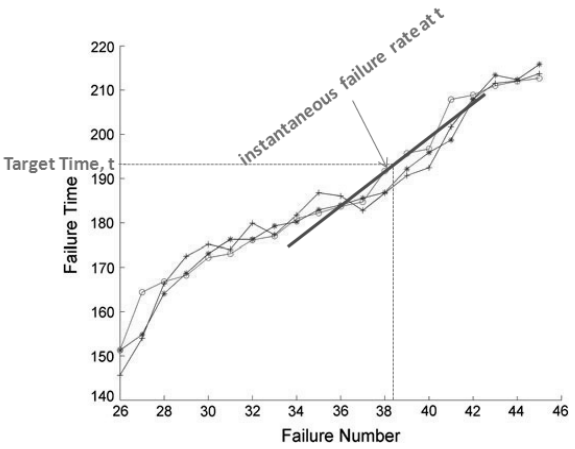


Fig. 3. Failure rate, MTBF prediction

4. 사례분석

분석에 사용된 자료는 현재 육군에서 운용 중인 궤도장비의 11년간 수집된 고장자료(16,419개)로, 군수지원요소 최신화를 위해 별도로 수집되었으며, 원 자료 형태는 Table 3과 같다.

1단계 모형에는 고장을 장비별로 종합하여 Table 1 형태로 변환, 입력 특성벡터로 사용하였다. 세부 구성요소로는 총 고장수, 총 운용거리, 총 운용시간 및 고장개소(동력장치, 송탄장치, 현수장치 등) 20위치별 누적고장수와 고장의 심각도(임무불가, 임무저하, 임무무관)별 누적고장수 3개 등 총 26개로 구성하였다. 2단계 모형에는 원 자료 중 장비호수, 배치일자, 고장시점 자료만 사용하며, 고장시점은 Table 2와 같이 변환하여 입력 특성벡터로 사용하였다. 분석에는 통계분석 SW R[14]의 지지벡터회귀분석 범용 라이브러리인 e1071을 사용하였다.

Table 3. Test data set(partial)

Machine #	Deploy	Fail Location	Fail Time	Mission Effectiveness
1	1	CC	73.3	F2
1	1	RR	76.5	F3
1	1	BB	81.8	F3
1	1	SS	94.5	F3
1	1	SS	100.0	F2
1	1	CC	111.3	F2

가. 1단계 모형 적용 결과

총 자료의 70 %를 임의 추출하여 학습에 사용하였으며, 나머지 30 % 자료를 모형 검증에 사용하였다. 최적 파라미터와 예측 성능은 Table 4와 같다.

Table 4. Test result for 1st stage

구 분		결 과
Parameter	C	1.000
	ϵ	0.010
	σ	0.363
F-measure	최소	48.7 %
	평균	69.5 %
	최대	100.0 %
Total Accuracy		67.8 %

예측력 확인을 위해 F-measure[15]와 Accuracy를 사용한 결과 평균적으로 2/3정도의 정확도를 나타내어 모형의 적용 가능성은 확인되나, 향후 실 적용을 위해서는 입력특성벡터 정교화 및 분석대상 장비수량 확대 등이 필요한 것으로 나타났다. 장비군 별로 F-measure 기준, 48.7 % ~ 100.0 % 수준을 나타내어, 일부 장비군에 대해서는 매우 높은 정확도를 나타내기도 한 것

으로 확인되었다.

F-measure의 형태는 식 (12)와 같으며, Precision과 Recall 조화평균(Harmonic mean) 형태이고, 1단계 모형과 같이 다중판별분석 시 각 그룹별 효율을 확인하기 위한 측정자로 사용된다.

$$F-measure_i = 2 \cdot \frac{Precision_i \cdot Recall_i}{Precision_i + Recall_i} \quad (12)$$

$$Precision_i = \frac{i\text{를 } i\text{로 예측한 개수}}{\text{실제 } i\text{의 개수}}$$

$$Recall_i = \frac{i\text{를 } i\text{로 예측한 개수}}{\text{예측된 } i\text{의 개수}}$$

나. 2단계 모형 적용 결과

동일한 장비군으로 식별된 5개의 장비 고장자료를 이용하여 2단계 모형의 성능을 확인하였다. 각 장비별 고장수는 Table 5에 제시된 것과 같으며, '3. 나. 2)'에서 제시한 방법에 따라, 1호기의 고장수가 가장 많으므로, 기준 장비는 1호기가 된다.

Table 5. # of failure for each machine

장비	고장수(EA)	운용시간(Hr)
1호기	157	**2175.96**
2호기	152	2126.06
3호기	128	1923.00
4호기	150	1923.50
5호기	140	1921.16

입력특성벡터는 lag = 3일 때의 값을 기준으로 하였고, 각 위치별 최적 파라미터 학습 결과는 Table 6과 같다.

Table 6. Parameter training result for 2nd stage

Fail point	σ	C	ϵ
106	0.745	70	0.026
107	0.526	10	0.002
...	...	...	...
157	0.665	10	0.011

Table 7의 결과를 이용하여 나머지 4개 장비의 고장위치를 추정하였고, 이를 신뢰도 성장분석 및 예측에 가장 많이 사용되는 모형인 AMSAA 모형[3]과 비교하였다.

AMSAA 모형은 식 (13)과 같은 형태의 고장 강도함수(Failure Intensity Function)를 가지는 NHPP 모형이므로,

$$\lambda(t) = \frac{\beta}{\theta}\left(\frac{t}{\theta}\right)^{\beta-1} \quad (13)$$

$[t_i, t_{i+1}]$에서의 고장수 추정값은 NHPP의 정의에 따라, 식 (13)을 적분한 평균값 함수 $\Lambda(t)$에 관한 식 (14)로 표현된다.

$$\Lambda(t_{i+1}) - \Lambda(t_i) = \left(\frac{t_{i+1}}{\theta}\right)^{\beta} - \left(\frac{t_i}{\theta}\right)^{\beta} \quad (14)$$

따라서 i^{th}고장시점을 알 때, $(i+1)^{th}$고장시점 t_{i+1}은 식 (15)로 추정된다.

$$t_{i+1} = \theta\left[1 + \left(\frac{t_i}{\theta}\right)^{\beta}\right]^{\frac{1}{\beta}} \quad (15)$$

성능 비교는 두 모형에 의해 추정된 고장시점과 실제 고장시점간의 차이에 대한 MSE와 NMSE(Normalized Mean Squared Error)를 사용하였다.

Table 7. Performance Comparison(SVR vs. AMSAA)

장비	MSE (SVR)	MSE (AMSAA)	NMSE (SVR)	NMSE (AMSAA)
2호기	704.77	636.94	4.58	4.19
3호기	2,021.26	830.78	28.84	11.69
4호기	521.70	698.07	4.30	5.81
5호기	1723.02	1763.57	2.82	2.87

NMSE : $\times 10^{-6}$

기준 장비인 1호기를 기준으로, AMSAA 모형과 SVR을 이용하여 나머지 호기의 고장위치를 추정한 결과는 Fig. 4 ~ Fig. 7에서 확인할 수 있고, MSE와 NMSE 차원의 성능은 Table 7과 같다. 측정자에 따른 차이는 거의 없었으며, 효율은 3호기의 경우 AMSAA 모형을 통해 예측한 결과보다 매우 낮은 결과를 나타

내고 있으나, 나머지 2,4,5호기의 경우 유사하거나 더 나은 결과를 나타내고 있다. 3호기의 경우 AMSAA 모형에 의한 예측결과도 타 장비에 비해 다소낮기 때문에, 비록 CVM test나 1단계 모형을 통과할 수준은 되나 나머지 호기와 동질성은 타 장비대비 낮으며 이에 따른 예측 성능도 낮아진 것으로 해석된다.

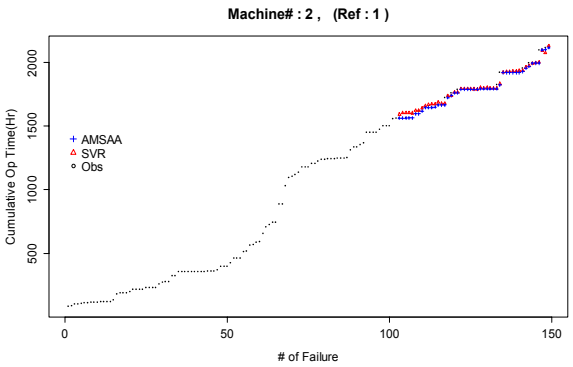


Fig. 4 Failure point prediction(Machine #2)

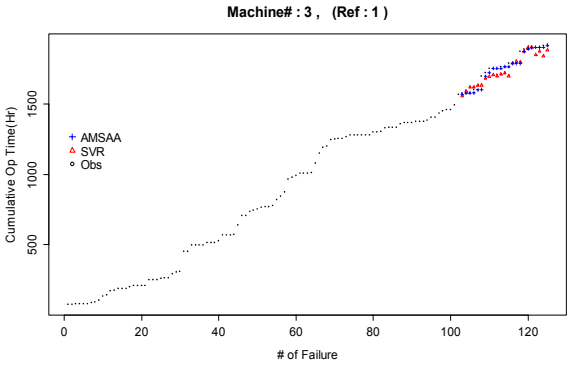


Fig. 5. Failure point prediction(Machine #3)

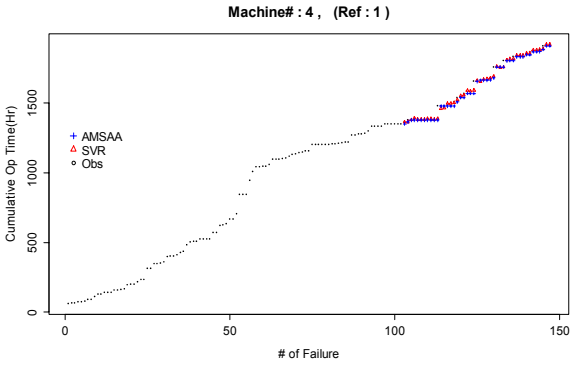


Fig. 6. Failure point prediction(Machine #4)

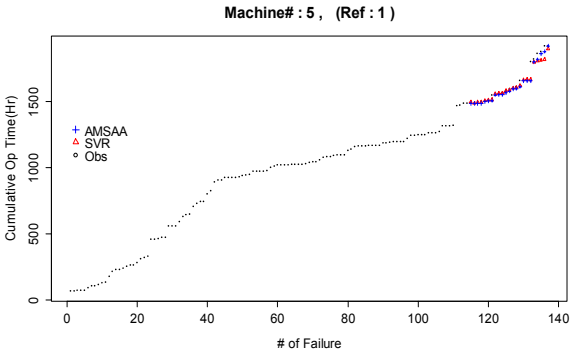


Fig. 7. Failure point prediction(Machine #5)

다. 3단계 모형 적용 결과

Fig. 4에서, 1,800시간에서의 순간 고장률 λ_{INST}와 순간 MTBF $MTBF_{INST}$를 추정하면, 1,800시간 내외에서의 고장시점 추정값은 각각 1781.008, 1834.805이므로, 식 (16), 식 (17)과 같다.

$$\lambda_{INST} = \frac{1}{1834.805 - 1781.008} = 0.019 \ \text{(회/시간)} \qquad (16)$$

$$MTBF_{INST} = \frac{1}{0.019} = 53.797 \ \text{(Hr)} \qquad (17)$$

5. 결 론

본 논문에서는 장비 고장위치, 장비 고장수 등의 고장자료로 미래 고장시점을 예측하는 방법을 제시하고 있다. 제시된 방법은 단순히 배치시점별로 장비를 그룹화 하여 분석하는 방식에서 탈피하고 있어 보다 현실성 있는 신뢰도 예측을 수행할 수 있다. 특히 SVR을 사용하기 때문에 방법론의 특성상 고장이 자주 발생되지 않는 장비의 고장시점 예측에도 무리 없이 적용이 가능한 점과, 일단 학습이 완료된 후에는 향후 배치될 장비 혹은 배치 초기 장비의 전반적인 고장 시점을 예측할 수 있는 점 등 기존 모형들과 유사한 기능을 제공할 수 있다.

모형의 성능 비교를 위해 가장 대표적인 신뢰도 장모형인 AMSAA 모형과 비교한 결과 역시 고장위치 예측력에 있어서는 AMSAA 모형과 유사함을 확인할 수 있었고 일부 자료에 대해서는 좀 더 작은 오차를 보여주기도 하는 등 향후 모형 정교화에 따라 기존

모델 이상의 성능을 보여줄 수 있는 가능성을 확인하였다.

다만 제안된 방법은 기계학습과정을 수반하여, 기준 장비 특성을 학습하는 단계에 많은 시간이 소요되고, 예측된 자료를 도시한 그래프 외에는 신뢰도 성장, 감소를 판단하는 것이 용이하지 않은 단점이 있다. 또한, 향후 1단계 모형의 타당성 확인에 추가적인 자료가 필요할 것으로 예측되는 점 등은 본 논문의 제한점으로 남는다. 그러나 현재 참고문헌으로 제시된 논문 외에도 인공지능 기법을 사용한 다양한 신뢰도 예측 방법이 연구되고 있음이 확인되는 바, 본 논문을 시작으로 군수분야의 신뢰도 성장분석 및 예측시 전통적인 확률/통계 기반의 모형 뿐만 아니라, 다양한 기법이 사용될 수 있게 되기를 기대한다.

References

[1] Duane J. T., “Learning Curve Approach to Reliability Monitoring”, IEEE Trans. Aero., AS-2, pp. 563～566, 1964.

[2] “Reliability Growth Management”, MIL-HDBK-189, 1981.

[3] AMSAA Reliability Growth Guide, Technical Report # TR-652, 2000.

[4] Ho SL, Xie M, “The Use of ARIMA Models for Reliability Forecasting and Analysis”, Computer Ind. Eng., Vol. 35, pp. 213～216, 1998.

[5] Ho SL, Xie M, “A Comparative Study of Neural Network and Box-Jenkins ARIMA Modeling in Time Series Prediction”, Comput. Ind. Eng., Vol. 42, pp. 371～375, 2002.

[6] Wei Chiang Hong, Ping Feng Pai, “Predicting Engine Reliability by Support Vector Machines”, Int J Adv Manuf Tech., Vol. 28, pp. 154～161, 2006.

[7] Y. Fuqing, U. Kumar, D. Galar, “Reliability Predicting Using Support Vector Regression”, Intl J of Systems Assurance Engineering and Management, Vol. 1, No. 3, pp. 263～268, 2010.

[8] Vapnic V., “The Nature of Statistical Learning Theory”, Springer, Berlin Heidelberg New York, 1995.

[9] Vojislav K., “Learning and Soft Computing : Support Vector Machines, Neural Networks and Fuzzy Logic Models”, MIT Press, Cambridge, MA, 2001.

[10] Avriel, Mordecai, “Nonlinear Programming : Analysis and Methods”, Dover Publications, 2003.

[11] Scho¨lkopf B, Smola A. J., “Learning with Kernels : Support Vector Machines, Regularization, Optimization and Beyond Adaptive Computation and Machine Learning”, MIT Press, Cambridge, 2002.

[12] Hochbaum D. S., “Approximation Algorithms for NP-hard Problems”, PWS Pub. Co, Boston, 1997.

[13] Cawley GC, Talbot NLC, “Efficient Leave-One-Out Cross Validation of Kernel Fisher Discriminant Classifiers”, Pattern Recogn, Vol. 36, No. 11, pp. 2585～2592, 2003.

[14] R Core Team, “R : A Language and Environment for Statistical Computing. R Foundation for Statistical Computing”, Vienna, Austria, ISBN 3-900051-07-0, 2012.

[15] Powers, D. M. W. “Evaluation : From Precision, Recall & F-measure to ROC, Informedness, Markedness & Correlation”, Journal of Machine Learning Tech., Vol. 2, Issue 1, pp. 37～63, 2011.