

PSO법을 응용한 확률적 시뮬레이션의 최적화 기법 연구

김선범¹ · 김정훈¹ · 이동훈^{1*}

A Study on Modified PSO for the Optimization of Stochastic Simulations

Sunbum Kim · Junghoon Kim · Donghoon Lee

ABSTRACT

This paper describes the method to solve the optimization problems for stochastic simulation which is represented by military simulations. For this reason, the test fitness function reflecting the characteristics of military simulations, complex and stochastic results, is defined and PSO is used to solve the test fitness function. To control the known weak point of PSO for stochastic simulations, this paper proposes a technique which reevaluates the value of global optimum. By using the technique, the result shows notable improvements. From the simulation results, interactions among the calculation conditions which affect the accuracy and speed of optimization are analyzed. And the strategy for the optimization of stochastic simulations is proposed.

Key words : PSO (Particle Swarm Optimization), Stochastic simulation, Optimization

요 약

일반적으로 최적화 문제에서 군사 시뮬레이션과 같이 결과가 확률적으로 나타나는 경우를 계산할 때에는 문제를 모델링 하여 일반적인 최적화 기법을 적용하는 것에 어려움이 있다. 본 논문에서는 이러한 군사 시뮬레이션의 특징을 반영하는 복잡한 반응 표면을 가진 확률적 평가 함수를 정의하였다. 그리고 이러한 확률적 시뮬레이션에 대해 기존의 PSO법이 가진 약점을 보완하는 기법을 제안하였다. 제안한 기법을 이용해 평가 함수에 대한 최적화를 시행하였으며 최적화의 속도와 정확도에 영향을 미치는 계산 조건들의 상호작용을 분석하였다. 이를 통해 본 논문에서 제안한 확률적 시뮬레이션의 최적화 전략을 제시하였다.

주요어 : PSO법, 확률적 시뮬레이션, 최적화

1. 서 론

최적화 문제는 함수의 출력을 최대화 시키는 입력을 찾는 문제로 정의할 수 있다. 이러한 최적화 문제는 등반 알고리즘 등을 기반으로 하고 있으나, 출력의 형태가 복잡하거나 반응표면이 미분 불가능해 일반적인 알고리즘을 사용하기 어려울 경우, 유전 알고리즘이나 Particle Swarm Optimization(PSO법)과 같은 휴리스틱을 적용할

수 있다. 하지만 휴리스틱을 이용하는 최적화 문제에 대한 대부분의 연구는 반응표면이 결정적인 경우를 다루고 있기에 군사 시뮬레이션과 같이 결과가 확률적으로 나타나는 경우(특히 성공 또는 실패로 나타나는 경우)에는 적용하기가 어렵다. 이에 본 연구는 최적화 문제를 군사적 관점의 확률적 시뮬레이션에 적용하기 위한 방법론을 제안하고자 한다. 일반적인 국방 시스템은 이산 사건을 기반으로 한 객체들의 상호작용으로 복잡하게 모의된다. 복잡한 시스템의 시뮬레이션은 출력을 통해 평가가 가능하며 이러한 시스템의 시뮬레이션 출력은 명중률, 지속 시간 등 다양한 형태로 표현 될 수 있다. 본 연구에서는 해석의 용이함을 위해 해석 대상인 시스템을 시뮬레이션 출력이 성공과 실패로 구분되는 경우로 한정한다. 어뢰의 명중 여부, 함정의 회피 성공 여부, 전쟁의 승패 여부 등

접수일(2013년 10월 22일), 심사일(2013년 10월 30일),
게재 확정일(2013년 11월 12일)

¹⁾ 국방과학연구소

주 저 자: 김선범

교신저자: 이동훈

E-mail; dhlee@add.re.kr

에 대한 시물레이션이 이에 해당된다. 이 경우 시물레이션의 반응 표면이 베르누이 시행(Bernoulli trial)을 따르는 확률 변수가 됨으로 일반적으로 적용되는 등반 알고리즘(Hill climbing algorithm)등을 사용하기 위한 조건인 반응 표면이 단일 피크(peak)를 가지고 있으며 또 단일 표면이 미분 가능해야 한다는 조건을 충족하지 못하므로 진화적인 알고리즘의 고려가 필요하다.

위와 같은 연구의 일환으로써 유전 알고리즘을 이용한 방법(1999, 이동훈)이 확률적 시물레이션에 적용된 바 있으며, 본 논문에서는 새로운 대안으로써 진화적 알고리즘의 하나인 PSO법(Particle Swarm Optimization)을 사용하였다. PSO법은 새떼와 물고기떼 등의 사회적 행동양식을 모방한 최적화 기법으로써 각 해들인 Particle과 해들의 전체 집합인 군집 Swarm으로 표현되어 탐색공간에서의 이동을 통해 문제의 최적치를 찾아낸다. 유전 알고리즘과 PSO법은 아래와 같은 공통적인 특징을 지닌다(1989, Goldberg).

1. 한 점이 아닌 다수의 점으로 이루어진 군집을 사용해 탐색공간을 전체적으로 탐색한다.
2. 평가함수 값 이외의 미분 혹은 다른 종속적인 변수들을 계산에 이용하지 않아 복잡한 반응 표면을 가지는 시물레이션에 유리하다.
3. 계산과정 자체에 확률적인 요소를 포함한다.

때문에, PSO법은 확률적이고 복잡한 반응 표면을 가지는 군사 시물레이션 최적화 문제의 해결에 적합하며, 그 원리가 간단해 구현이 용이한 이점이 있다.

본 논문에서는 PSO법을 변형해 확률적 문제에 대한 탐색성능을 향상시킨 알고리즘을 제안하였다. 또한 본 논문에서 해석 대상으로 하는 군사시물레이션의 특징을 충분히 반영한 평가함수를 정의하고, 그 평가함수에 대해 여러 가지 계산조건에 따른 탐색 능력의 변화를 분석하여 탐색 전략을 제시하였다.

2. 확률적 시물레이션을 위한 평가 함수의 정의

1장에서 논의한 국방 분야의 시물레이션의 최적화의 적용성을 알아보기 위해서는 실제 문제와 유사한 형태의 적절한 평가 함수를 정의하는 것이 중요하다. 위에서 논의한 국방 분야의 시물레이션에 적용되어야 할 평가 함수의 특징은 다음과 같다.

1. 반응 표면의 형태가 복잡함.
2. 여러 개의 국부 피크치를 가진.
3. 출력의 형태가 베르누이 확률 변수로 나타남.

이러한 평가함수의 탐색 알고리즘의 효과도 비교를 위한 테스트 함수는 여러 가지가 알려져 있으나, 본 논문에서는 다음과 같은 평가함수를 정의하였다.

$$f(X) = \frac{1}{m} \sum_{i=1}^m Z(X)$$

$Z(X)$ 는 탐색공간의 한 지점 X 에서의 확률 $p(X)$ 인 베르누이 시행이며 1 혹은 0의 값을 가진다. 그러므로 $f(X)$ 는 평균 $p(X)$ 이고 표준편차 $\sqrt{p(1-p)/m}$ 인 정규분포에 가까워지고 표준편차가 0으로 수렴하여 결정적 문제에 가까워지게 된다. 여기서 설정한 $p(X)$ 는 다음과 같다.

$$p(X) = 0.05 + 0.9 \prod_{i=1}^D [|\sin(x_i)| \cdot s(x_i)]$$

$$s(x_i) = \begin{cases} 1 & \text{if } x_i > 0 \\ 0.9 & \text{otherwise} \end{cases}$$

$$-\pi < x_i < \pi, i = 1, \dots, D$$

D 는 탐색공간의 벡터의 크기를 나타내며 변수의 개수를 의미한다. $p(X)$ 는 탐색공간의 모든 변수의 값이 $\pi/2$ 인 지점 $\mathbf{0} = (\pi/2, \pi/2, \dots, \pi/2)$ 에서 전역적 피크치인 0.95를 가지게 되며 한 개 이상의 변수가 $-\pi/2$ 일 경우에는 0.95보다 작은 국부 피크치를 가지도록 설정되어 있다. 이러한 국부 피크치의 개수는 2^{D-1} 개이다. 그러므로 벡터의 크기 D 가 증가할 경우 국부 피크치의 개수는 기하급수적으로 증가하게 된다. Fig. 1에 변수의 개수가 2개, 국부 피크치의 개수가 4개인 $p(X)$ 가 나타나 있다.

이러한 확률 $p(X)$ 를 따르는 확률적 평가함수 $f(X)$ 를 정의해 본 논문의 해석 대상으로 함으로써 본 논문에서 다루는 확률적 시물레이션의 최적화 문제는 평가함수 $f(X)$ 를 최대화 시키는 입력조건을 찾는 것으로 귀결된다. 본 논문에서는 탐색을 위한 효과적인 계산 조건 설정을 위해 변수의 개수 D 를 8개로 제한하였다. 이에 따라 평가함수 $f(X)$ 의 국부 피크치는 총 $2^{8-1}=512$ 개가 된다.

이렇게 변수의 수를 조절하는 것으로 국부 피크치를 자유롭게 조절할 수 있는 점은 최적치 탐색 기법의 성능을 시험하기에 유리한 이점이라고 할 수 있다.

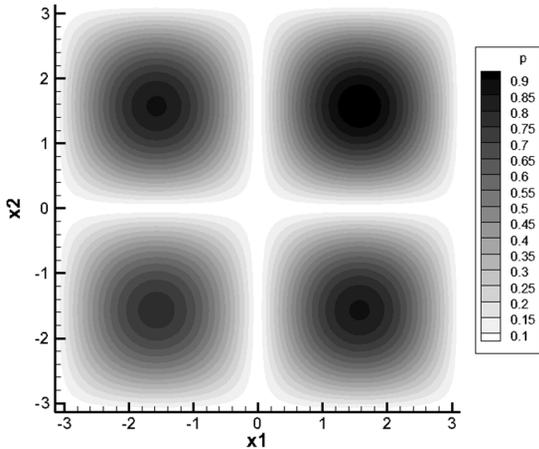


Fig. 1. Expected value of fitness function $p(X)$ ($D = 2$)

3. PSO법 및 비결정적 시뮬레이션을 위한 PSO법의 응용

3.1 PSO법

PSO (Particle Swarm Optimization)법은 전산 최적화 기법의 일종인 휴리스틱(Heuristic)에 속하는 방법이다. 휴리스틱은 알고리즘이나 반복법과는 달리 수렴치 등의 특정한 종료조건을 필요로 하지 않기 때문에 일반적으로 정해진 반복횟수 등이 만족될 경우 종료된다. Kennedy와 Eberhart(1995)에 의해 제안된 PSO법은 개체의 군집이론을 바탕으로 Global Optimization 문제를 수행한다.

PSO법의 알고리즘은 다음과 같다. D 차원 탐색 공간 안의 임의의 위치 $X_i^n = (x_{i1}, x_{i2}, \dots, x_{iD})$ 에 위치한 i 번째 Particle은 자신이 경험했던 가장 좋은 해의 위치 $P_i^n = (p_{i1}, p_{i2}, \dots, p_{iD})$ 와 군집이 경험한 가장 좋은 해의 위치 $G^n = (g_1, g_2, \dots, g_D)$ 를 이용해 다음 단위시간의 위치 X_i^{n+1} 를 결정한다. 즉,

1. D 차원의 N 개의 Particle들의 위치 X_i^n 과 속도 V_i^n 의 위치를 초기화 한다.
2. 각 Particle들의 평가 함수치를 계산한다.
3. 현재의 평가치와 과거의 최적치를 비교해 최적치를 갱신한다.
4. 현재의 평가치와 군집 전체의 최적치를 비교해 갱신한다.
5. 각 Particle들에 대해 각각의 최적치와 군집의 최적치를 이용해 속도를 계산한다.
6. 속도를 이용해 각 Particle들의 위치를 갱신한다.
7. 종료조건(일반적으로 반복 횟수)가 만족되었을 경우

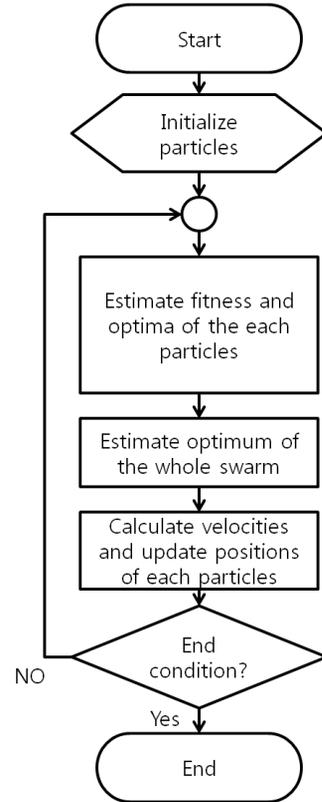


Fig. 2. Flowchart of the PSO

종료하며 만족되지 않았을 경우 2로 돌아가 반복한다. 이 알고리즘의 Flowchart는 Fig. 2과 같으며 식으로는 다음과 같이 표현된다.

$$V_i^{n+1} = wV_i^n + c_1r_1(P_i^n - X_i^n) + c_2r_2(G^n - X_i^n)$$

$$X_i^{n+1} = X_i^n + V_i^{n+1}$$

여기서 $i = 1, 2, \dots, N$, N 은 군집 전체의 크기를 나타내며, w, c_1, c_2 는 양의 실수로서 각 항의 가중치, r_1, r_2 는 0과 1사이의 실수 난수 그리고 n 은 현재 계산 스텝 수이다.

PSO법은 Particle이 경험한 최적치의 위치 P_i^n 과 군집 전체가 경험한 최적치의 위치 G^n 을 공유함으로써 Particle 주변의 국부적인 탐색과 공간 전체에 대한 전역적인 탐색을 동시에 수행하게 된다. 가중치 w 를 변화시킴으로써 전역적 탐색 및 국부적 탐색에 대한 가중치가 달라진다. w 가 큰 경우, 전역적인 탐색의 비중이 커지게 되고 w 가 작은 경우, 국부적인 탐색의 비중이 커진다.

또한, 본 연구에서는 탐색공간의 유효숫자 크기와 Sampling

수에 비례해 증가하는 계산량을 효율적으로 처리하기 위해, 변수를 2진수로 변환해 탐색공간을 유한개의 변수로 표현하는 DPSO (Discrete binary version of PSO)법 (1997, Kennedy and Eberhart)을 사용하였다. 변수의 범위를 충분히 크게 확보함으로써 실수공간인 탐색공간을 효과적으로 이산화해 효율적으로 계산할 수 있으며 차후 유전 알고리즘 등과의 비교 또한 용이해진다.

여기서 속도의 발산을 막기 위해 속도의 최대값 V_{max} 와 V_{min} 을 설정할 필요가 있다. 또한, 가중치 c_1, c_2 는 PSO법이 국부적인 최적치에 수렴하는 것을 방지하는 역할을 한다. Kennedy(1998)에 의해 $c_1 = c_2 = 2.0$ 이 제안되었다.

본 논문에서는 계수들에 대해 $V_{max} = 6.0, V_{min} = -6.0, w = 1.0, c_1 = c_2 = 2.0$ 을 적용하였다.

3.2 비결정적 시뮬레이션을 위한 PSO법의 응용

본 논문에서 해석 대상으로 하고 있는 평가함수 $f(\mathbf{X})$ 와 같은 확률적 시뮬레이션에서는 종래의 PSO법을 사용하는 것 뿐 만 아니라 확률적으로 발생하는 과대평가치를 처리하는 방법이 도입되어야 한다.

예를 들어, 기존의 PSO법을 사용해 평가함수 $f(\mathbf{X})$ 를 해석할 경우, Fig. 3에서와 같이 발생한 국부적 최적치 주변의 과대평가치의 위치 G_{fake1} 혹은 전역적 최적치 주변의 과대평가치의 위치 G_{fake2} 가 군집 전체가 경험한 최적치의 위치로서 인식되어질 수 있다. 기존의 PSO법에서는 군집 전체가 경험한 최적치의 위치 G 는 한번 결정된 경우, 위치가 변하지 않는 한 각 계산 스텝에서의 갱신은 이루어지지 않는다. 그러므로 이러한 과대평가치의 위치 G_{fake} 가 계산 종료 시까지 지속되는 경우, 군집의 속도 벡터의 계산의 영향을 미칠 뿐만 아니라 최종적인 최적치의 위치로 선택되어 계산의 정확도에 악영향을 미칠 수 있다. 이러한 과대평가치의 선택을 피하기 위해 본 논문에서는 최적치의 위치 G 에 대하여 각 계산 스텝에서 그 최적치를 갱신하는 방법을 통해 확률적 과대평가치의 발생을 제어

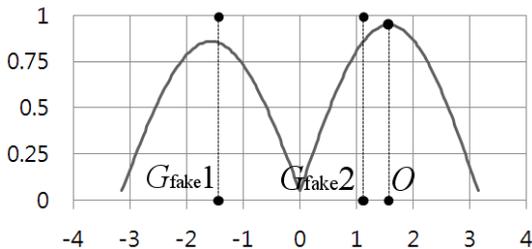


Fig. 3. Occurrence of the overestimated value G_{fake}

하였다. 상세한 방법은 다음과 같다.

1. 현재의 군집 전체의 최적치 G_{best}^n 가 지속된 횟수 C 를 계산한다.
2. 현재 최적치의 위치 G^n 에서 의 평가치 $f(G^n) = G_{best}^*$ 를 새로 계산한다.
3. G_{best}^n 에 C 를 가중치로 곱한 뒤 G_{best}^* 를 더해 $C+1$ 로 평균한 뒤 군집의 최적치로서 PSO법의 계산에 대입한다. 식으로는 다음과 같이 표현된다.

$$G_{best}^n = (CG_{best}^n + G_{best}^*) / (C + 1)$$

본 방법은 군집 전체의 최적치 G_{best}^n 와 최적위치에서의 재평가치 G_{best}^* 를 가중치를 주어 평균함으로써 확률적으로 발생하는 과대평가치를 평가함수의 기댓값으로 낮추는 역할을 한다. 이렇게 피크 주변에서 발생할 수 있는 과대평가치를 억제함으로써, 과거의 최적치를 기억하는 PSO법에 의해 개체가 확률적으로 최적치가 아닌 최적치 주변의 점으로 잘못 유도되어 질 수 있는 가능성이 낮아져 피크에서 수렴한 뒤의 fine-tuning이 이루어지는 것을 기대할 수 있다.

본 방법을 기존의 PSO법과 비교했을 때의 결과는 다음과 같다. Table 1은 $X = (X_1, X_2, \dots, X_8)$ 에 대해 Particle 수 200, Sampling 수 200, 계산 스텝 수 250으로 10번 시행한 결과이다. 평가함수 $f(\mathbf{X})$ 의 이론상의 최적위치인 $O = (\pi/2, \pi/2, \dots, \pi/2)$ 와 계산상의 최적치의 위치인 G 의

Table 1. Preparation of average value of Distance (10 trials) between PSO and modified PSO (Particle No. : 200, Sampling No. : 200, Steps : 250)

	PSO	Modified PSO
Trial 1	0.1592	0.0937
Trial 2	0.2934	0.1353
Trial 3	0.162	0.1236
Trial 4	0.2009	0.1122
Trial 5	0.3118	0.1442
Trial 6	0.2523	0.0874
Trial 7	0.2204	0.1189
Trial 8	0.2284	0.073
Trial 9	0.1299	0.1161
Trial 10	0.1475	0.0537
Average Value	0.21058	0.10581

거리를 $Distance = |G - O|$ 라고 정의하였다. G 가 Global Optima에서 가까운 경우, Distance가 작으면 작을수록 이론상 최적치에 근접하게 된다.

Table 1에 나타난 바와 같이 기존의 PSO법에 비해 매 스텝 최적치를 갱신한 방법의 경우 Distance의 평균값이 약 1/2로 감소해 상당히 개선된 결과가 나타남으로써 피크치 수렴 뒤의 fine-tuning이 실제로 적용되었음을 알 수 있다.

4. 각 계산 조건들과 비결정적 시뮬레이션의 최적화도의 상관관계

3장에서 설명한 해석법을 바탕으로 평가함수 $f(X)$ 를 이용한 시뮬레이션의 최적화 전략을 고찰하였다. 일반적으로 PSO법을 이용한 최적화는 다음과 같이 수행되어진다. Fig. 4는 Particle 수 100, Sampling 수 100에 대해 250 스텝을 계산한 결과이다.

Fig. 4에 나타난 바와 같이, 군집 전체의 최적치(Best Fitness)는 계산 초기 스텝의 증가와 함께 증가하여

수렴한다. 계산상의 최적치의 위치와 이론적 최적치의 위치의 거리(Distance) 또한 감소하는 것을 알 수 있다. 여기서 나타난 수렴 속도 및 수렴한 최적치의 정확도와 각 계산 조건들의 상관관계를 알아내는 것은 시뮬레이션 최적화 전략과 직결된다.

시뮬레이션의 계산시간은 Particle 수와 Sampling 수,

계산 스텝 수에 의해 결정된다. 수식으로는 다음과 같이 표현 된다.

$$\text{계산시간} = \text{Particle 수} \times \text{Sampling 수} \times \text{계산 스텝 수}$$

Particle 수와 Sampling 수 그리고 계산 스텝 수는 각 각 크면 클수록 계산의 정확성을 증가시키지만, 위 식과 같이 계산시간이 급격히 증가하므로 해석에 현실적인 어려움을 가진다. 그러므로 이러한 계산 조건을 얼마나 효율적으로 결정해 배치하느냐가 계산시간과 계산의 정확도에 결정적인 영향을 미치게 된다.

본 논문에서는 보다 정확한 비교를 위해 Particle 수와 Sampling 수의 곱이 일정한 12가지의 경우를 가지고 각 계산 조건들과 계산의 정확성의 상관관계를 고찰하였다. 계산 조건들은 Table 2과 같다. 계산의 일반성을 확보하기 위해 각 계산 조건들에 대해 100번의 계산을 수행한 뒤, 평균을 내어 조사하였다.

4.1 계산 조건과 최적치 크기와의 상관관계

먼저 Sampling 수와 평가된 최적치 크기와의 상관관계에 대해 해석해 보았다. Fig. 5는 각 계산 조건들에 의해 계산된 최적치의 평균을 나타낸 그래프이다. 최적치의 크기와 계산 스텝수와의 상관관계를 알아보기 위해 각 50 스텝 마다 최적치들의 평균의 측정이 이루어 졌다.

Sampling 수가 작은 경우(500 이하)에서 평가된 최적치의 크기가 이론적 최적치인 0.95를 넘어 1.0에 가까워

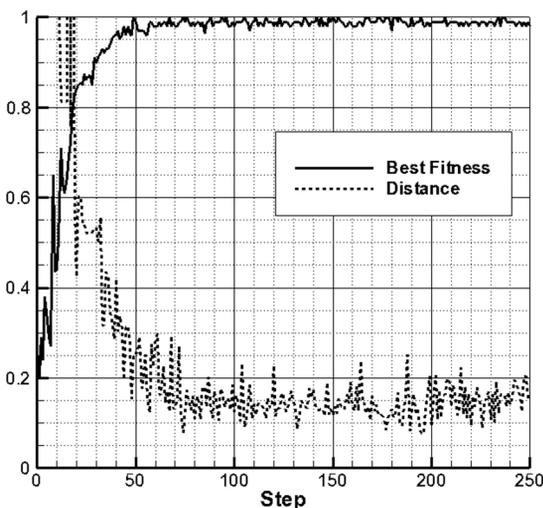


Fig. 4. Optimization of the fitness function by using modified PSO

Table 2. Number of cases of calculation conditions (Particle No.×Sampling No. = 10,000)

	Particle No.	Sampling No.
Condition 1	1,000	10
Condition 2	500	20
Condition 3	400	25
Condition 4	250	40
Condition 5	200	50
Condition 6	100	100
Condition 7	50	200
Condition 8	40	250
Condition 9	25	400
Condition 10	20	500
Condition 11	10	1,000
Condition 12	5	2,000

지는 것을 알 수 있다. 이것은 기댓값이 0.95인 최적치 근방에서의 베르누이 시행이 모두 성공한 경우이다. 예를 들어 Sampling 수가 10, 성공확률이 0.9인 지점에서 평가치가 1.0이 나올 확률은 약 0.35로서 상당히 높은 확률로 과대평가치가 발생하는 것을 알 수 있다. 3장에서 설명한 바와 같이 이러한 과대평가치의 발생은 이론적 최적치의 수렴을 방해하므로 어느 정도 이상의 Sampling 수를 설정함으로써 이러한 과대평가치의 발생을 억제시켜 이론적 최적치인 0.95에 수렴시키는 것이 바람직하다.

하지만 단순히 Sampling 수를 증가시키는 것이 이론적 최적치의 수렴을 보장하지는 않는다. Fig. 5에서 Sampling 수가 1000보다 큰 경우에는 계산상의 최적치가 이론치보다 점점 낮아지는 것을 알 수 있다. 이는 Sampling 수와 반비례하는 Particle 수와 관계가 있다.

Particle 수는 한 계산 스텝에서 평가되는 지점의 개수를 의미한다. 즉 시뮬레이션에서 최종적으로 평가되는 지점의 개수는 다음과 같다.

$$\text{총 평가된 지점 개수} = \text{Particle 수} \times \text{계산 스텝 수}$$

즉, Sampling 수를 증가시켜서 과대 평가치의 발생을 억제한다고 해도 평가된 지점의 총 개수가 줄어들어서 계산의 정확도가 떨어지는 현상이 발생하는 것이다. 예를 들어 Sampling 수가 2,000이고 Particle 수가 5인 조건 12가 100 스텝 동안 탐색한 지점의 총 수는 500 인 것에 반해, Sampling 수가 100이고 Particle 수가 100인 조건

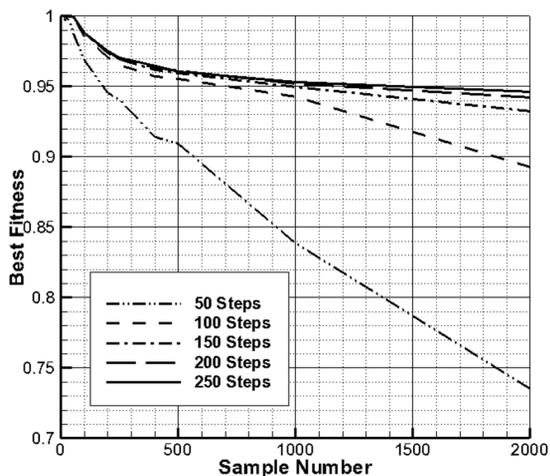


Fig. 5. The average value of Best fitness calculated on each conditions

6의 탐색지점 수는 같은 스텝동안 10,000에 달한다. 이러한 탐색 지점 수가 계산의 정확도에 미치는 영향 또한 Fig. 5에 잘 나타나 있다. Particle 수는 Sampling 수에 반비례하므로 Particle 수가 줄어들어 따라 평가되는 최적치의 값이 감소하는 경향을 보이고 있다.

여기서, 큰 Sampling 수로 인한 탐색 지점수의 부족을 보완하기 위한 방법의 하나로써 계산 스텝 수를 증가시킬 수 있다. 계산 스텝을 증가시킴으로써 평가되는 지점의 개수를 늘려 보다 큰 최적치를 탐색하는 것이 가능해진다.

4.2 계산 조건과 Distance와의 상관관계

하지만 확률적 시뮬레이션의 최적화에서는 단순히 최적치의 값이 큰 결과가 좋은 결과임을 보장하지 않는다. 3장에서 설명한 바와 같이 과대평가치가 발생하므로 이론적 최대치인 0.95보다 큰 값이 발생할 가능성이 있으며 또한 계산으로 얻어진 0.95의 최적치가 이론적 최적치의 위치인 θ 에 근접해 있다고 보기도 어렵기 때문이다.

이를 위해 3장에서 정의한 이론적 최적치와 계산상의 최적치의 거리인 $\text{Distance} = |G - \theta|$ 의 비교를 통해 탐색된 최적치가 이론적 최적치에 어느정도 가까운지 평가할 수 있다. Fig. 6는 각 계산 조건들에 의해 계산된 Distance의 평균을 나타낸 그래프이다.

Fig. 6에 나타난 바와 같이, 스텝 수가 50의 경우 최적화 계산이 발달하지 않아 불안정한 곡선을 나타내고 있다. 그리고 Sampling 수가 커짐에 따라 Distance가 감소했다가 다시 증가하는 변곡점이 전 스텝 수에서 일률적으로

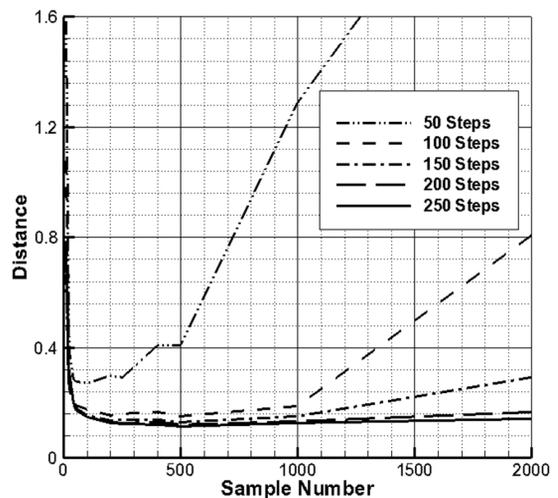


Fig. 6. The average value of Distance calculated on each conditions

나타나는 것을 알 수 있다. 이것은 PSO법을 응용한 방법을 통한 확률적 시뮬레이션의 최적화에서 Sampling 수에 대한 최적 개수가 존재함을 의미한다.

Fig. 6에서 같은 Sampling 수에서 y축에 평행한 각 스텝간의 차이는 최적화의 속도와 반비례한다. Sampling 개수가 500보다 큰 경우, 각 스텝간의 차이가 점점 커지는 것을 알 수 있다. 이것은 Sampling 개수가 큰 경우, 부족해지는 Particle 수로 인해 최적치의 탐색이 느려지는 것이라고 생각되어진다. 이러한 최적화의 속도는 스텝수가 증가할수록 증가해 일정치로 수렴하는 경향을 보이고 있다. 그러므로 이러한 경우, 스텝수를 증가시키는 것으로 부족한 Particle 수로 인한 오차를 해소 할 수 있다고 말할 수 있다.

그리고 Sampling 개수가 500보다 작은 경우 Sampling 수가 작아질수록 Distance가 증가해 발산하는 것을 알 수 있다. 이것은 과대평가치의 발생에 의한 최적치 탐색의 둔화라고 생각되어진다. 즉 Sampling 수의 부족으로 인해 충분한 베르누이 시행의 기댓값을 계산하지 못해 확률 공간 $p(\mathbf{X})$ 가 $f(\mathbf{X})$ 에 반영되지 못한 것으로 보인다. 이것은 실제로 Fig. 5에서 Sampling 수가 작을 때 나타난 1.0에 가까운 최적치와 관련이 깊다. 이러한 오차는 스텝 수를 증가시켜도 해소되지 않는 것을 알 수 있다.

각 스텝 간의 최적화의 속도의 수렴한 최적치의 위치를 종합적으로 고려할 때, 본 평가 함수에서의 최적화를 위해서는 Sampling 수가 200에서 500사이의 경우 가장 효과적인 것으로 판단된다. 이는 본 평가 함수의 국부 피크치의 개수인 512에 대해 어느 정도의 규모의 Particle의 수(50~20)를 확보하여 총 평가 지점 수를 충분하게 잡은 뒤, 비교적 많은 Sampling을 통해 평가 함수의 분산을 감소시켜 비결정적 문제에 근접시켜 최적화함으로써 최적화의 속도와 정확도가 향상되었다고 생각할 수 있다. 이것으로 볼 때 어느 정도의 Particle 수가 확보된다면 Sampling 수를 가능한 한 크게 설정하는 것이 본 평가 함수를 이용한 확률적 시뮬레이션의 최적화에 대해 유리하다고 판단된다.

또 Fig. 7은 각 계산 조건들에 의해 계산된 최적치의 기댓값의 평균, 즉 확률 $p(\mathbf{G})$ 의 평균을 나타낸 그래프이다. 이 그래프는 Fig. 6의 역수와 매우 유사한 형태를 보여준다. 이것은 Distance가 평가 함수의 최적화의 척도로써 효과적으로 작용하고 있음을 보여주는 한 예라고 생각된다.

Fig. 7에서는 스텝수를 증가시켜도 이론적 최적치인 0.95에 도달하지 못하고 있는 것을 알 수 있다. 이것은 수

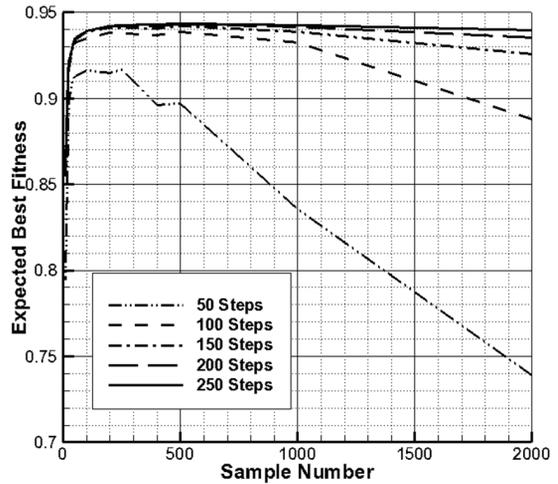


Fig. 7. The average value of Expected best fitness calculated on each conditions

렴 후의 fine-tuning과 관계가 있다. 추후 응용 PSO법의 파라미터인 V_{max} , V_{min} , w , c_1 , c_2 를 미세 조정하는 연구를 통해 이론적 최적치에 더욱 근접하게 할 수 있을 것으로 생각된다.

5. 결론

본 논문에서는 군사 시뮬레이션의 특징을 반영해 베르누이 시행을 출력하는 복잡한 반응 표면을 가진 확률적 시뮬레이션을 위해 변수의 개수가 D 일 때, 2^{D+1} 의 국소 최적치를 가지는 확률 공간 $p(\mathbf{X})$ 를 반영한 평가 함수(\mathbf{X})를 정의하였다.

그리고 확률적 시뮬레이션에 대해 기존의 PSO법이 가진, 군집 전체의 최적치가 과대평가치일 경우 최적화 계산의 정확도에 악영향을 미치는 약점을 보완하기 위해 각 계산 스텝에서 군집 전체의 최적치를 재평가해 가중치 평균을 내어 계산에 사용하는 기법을 제안하였다. 또 이 기법을 사용함으로써 확률적 시뮬레이션의 최적화에 대한 계산 결과가 개선됨을 보여주었다.

또 위의 기법을 도입해 평가 함수에 대한 최적화를 시행하였다. 최적화의 속도와 정확도에 영향을 미치는 세 가지의 계산 조건인 Particle 수, Sampling 수 그리고 계산 스텝 수의 상호작용을 분석하였다. 보다 정확한 비교를 위해 Particle 수와 Sampling 수의 곱을 10,000으로 제한한 12개의 경우에 대해 계산을 100번씩 시행한 뒤 평균값을 취해 계산 결과의 일반성을 확보하였다.

이러한 계산 조건들과 탐색되는 최적치의 크기, 그리고 계산 조건들과 이론적 최적치와 탐색된 최적치의 거리인 Distance의 상관관계를 조사한 결과, 어느 정도의 Particle 수를 확보한 뒤 가능한 한 Sampling 수를 크게 설정하는 것이 최적화의 속도와 정확도 면에서 유리하다고 판단할 수 있었다.

그러나 본 논문에서 나타난 바와 같이, 최적화가 수렴된 뒤의 fine-tuning등에 관련된 PSO법의 파라미터의 미세조정을 통해 최적화 기법의 정확도는 더욱 개선될 여지를 가지고 있으며, 국소적 최적치의 개수에 대해 요구되는 최소한의 Particle 수의 관계, Sampling 수와 Particle 수의 보다 상세한 상관관계 등이 밝혀지지 않은 채로 남아있다. 또한 PSO법이 아닌 다른 기법과 최적화 성능을 비교해야하고 테스트 평가 함수가 아닌 실제 문제에 적용해 이론을 검증해야하는 등 추후에도 확률적 시뮬레이션의 최적화에 대한 다양한 후속 연구가 필요할 것으로 생각되어진다.

References

1. Lee, D.H. and Huh, S.P., Genetic algorithm and clustering technique for optimization of stochastic simulation, Journal of the KIMST, Vol. 2, No. 2, pp. 90-100, 1999.
2. J. Kennedy and R. Eberhart, Particle swarm optimization, Proc. IEEE Int. Conf. Neural Networks, pp. 1942-1948, 1995.
3. J. Kennedy, "The behavior of particles", Evolutionary Programming VII, Lecture Notes in Computer Science Vol. 1447, pp. 579-589, 1998.
4. David E. Goldberg, "Genetic Algorithms in Search, Optimization, and Machine Learning", Addison-Wesley Professional, 1989.
5. J. Kennedy and R. Eberhart, A discrete binary version of the particles swarm algorithm, Proceedings of the IEEE international conference on Systems, Man and Cybernetics, pp. 4104-4108, IEEE press, 1997.



김 선 범 (kimsunbum@add.re.kr)

2011 Osaka University 지구종합공학과 선박해양공학 학사
 2013 Osaka University 지구종합공학과 선박해양공학 석사
 2013~현재 국방과학연구소 연구원

관심분야 : 수치해석, 모델링&시뮬레이션, 유체역학, 전산유체역학(CFD)



김 정 훈 (kimjh@add.re.kr)

1998 서울대학교 해양학과 학사
 2001 서울대학교 지구환경과학부 석사
 2001~현재 국방과학연구소 선임연구원

관심분야 : 해상/수중무기체계 전투실현, 수상/수중 합성전장환경 모의



이 동 훈 (dhlee@add.re.kr)

1986 고려대학교 통계학과 학사
 1988 고려대학교 응용통계학 석사
 1990~현재 국방과학연구소 책임연구원

관심분야 : 모델링&시뮬레이션, 최적화