

퍼지 시계열 예측을 위한 개선된 Particle Swarm Optimization 기법

Advanced Particle Swarm Optimization Technique for Fuzzy Time Series Forecasting

박진일¹, 이대종², 전명근¹

¹ 충북대학교 전기전자컴퓨터공학부

E-mail: moralskr@yahoo.co.kr, mgchun@chungbuk.ac.kr

² 충북대학교 BK21 충북정보기술사업단

E-mail: djmidori@empal.com

요 약

퍼지 시계열 예측은 전체 퍼지 구간에 따른 퍼지 소속 함수의 개수와 범위에 따라서 예측성능에 많은 영향을 미치고 있으며, 이러한 문제점을 개선하고자 최근 많은 연구가 진행 중에 있다. 본 논문에서는 이러한 퍼지 시계열의 문제점을 개선하기 위한 방법으로 다수 객체들의 학습 및 군집 특성을 이용한 Particle Swarm Optimization 기법을 도입하였다. 제안된 방법에서는 군집의 최적 객체를 전체 최적해와 각각의 퍼지 소속 함수들에 대한 최적해로 구분하여 탐색하는 기법을 제안한다. 실제 시계열 데이터를 이용한 실험을 통하여 기존의 연구 결과들과 비교함으로써 제안된 방법의 우수한 성능을 가짐을 검증하였다.

Key Words : Fuzzy Time Series, Genetic Algorithm, Particle Swarm Optimization, TAIFEX

1. 서 론

Song과 Chissom(1994, 1996)은 퍼지 이론에 기반을 둔 퍼지 시계열(Fuzzy Time Series) 모델을 제안하였다. Chen(1996)은 퍼지 규칙의 수가 많은 경우 퍼지 연산에 따른 연산시간이 많이 필요하다는 퍼지 시계열의 단점을 개선하고자 높은 예측 성능을 가지면서 단순한 형태의 퍼지 수 연산을 이용한 방법과 high-order 퍼지 시계열에 기반을 둔 예측 방법을 제안하였다. Lee(2008)등은 예측 성능의 향상을 위하여 퍼지 전체 집합에서 각각의 구간의 길이를 최적화하기 위하여 유전자 알고리즘에 simulated annealing 기법을 사용하는 방법을 제안하였다[1][3].

PSO는 군집 기반 알고리즘으로 새 떼와 물고기 떼와 같은 생체군집의 사회적 행동양식을 바탕으로 Kennedy와 Eberhart에 의해 처음 소개되었다[2]. 유전자 알고리즘이 종종 국부 해 수렴문제와 탐색시간이 오래 걸리는 단점에 비하여 PSO는 간결하고 구현이 용이하며, 연산의 효율성 및 짧은 계산 시간 안에 최적의 해를 생성할 수 있어 유전자 알고리즘에 비하여 안정적인 수렴특성을 가지고 있다. 본 논문에서는 Chen이 제안한 high-order 퍼지 시계열

의 예측 성능을 향상시키기 위하여 PSO를 적용하는 방법으로 PSO의 최적 해를 전체 퍼지 집합과 구간별 퍼지 집합의 최적 해로 나누어 탐색하는 방법을 제안하고자 한다.

2. 퍼지 시계열 예측 기법

two-factor k 차 order 퍼지 시계열은 다음 식과 같이 main-factor X_i 와 second-factor Y_i 의 퍼지 논리 관계로 나타낼 수 있다.

$$((X_{ik}, Y_{ik}), \dots, (X_{i2}, Y_{i2}), (X_{i1}, Y_{i1})) \rightarrow X_i \quad (1)$$

각각의 퍼지 논리 관계들은 그룹의 형태로 다음과 같이 정의되어진다.

$$\begin{aligned} (X_{ik}, Y_{ik}), (X_{i(k-1)}, Y_{i(k-1)}), \dots, (X_{i1}, Y_{i1}) &\rightarrow X_{j1} \\ (X_{ik}, Y_{ik}), (X_{i(k-1)}, Y_{i(k-1)}), \dots, (X_{i1}, Y_{i1}) &\rightarrow X_{j2} \\ &\vdots \\ (X_{ik}, Y_{ik}), (X_{i(k-1)}, Y_{i(k-1)}), \dots, (X_{i1}, Y_{i1}) &\rightarrow X_{jp} \end{aligned} \quad (2)$$

최종 예측 값은 다음과 같이 계산되어진다.

$$\frac{n_{j1} \times t_{j1} + n_{j2} \times t_{j2} + \dots + n_{jp} \times t_{jp}}{n_{j1} + n_{j2} + \dots + n_{jp}} \quad (3)$$

3. 퍼지 시계열 예측을 위한 Particle Swarm Optimization 기법

PSO에서 각 객체의 적합도를 평가하기 위하여 MSE(Mean Square Error)를 사용하였다.

main-factor의 전체 집합 U 는 $u_1 = [U_{\min}, x_1], u_2 = [x_1, x_2], \dots, u_n = [x_{n-1}, U_{\max}]$ 의 n 개의 구간으로 정의하고, second-factor의 전체 집합 V 는 m 개의 구간 $v_1 = [y_1, V_{\max}], v_2 = [y_2, y_1], \dots, v_m = [V_{\min}, y_{m-1}]$ 으로 구성된다.

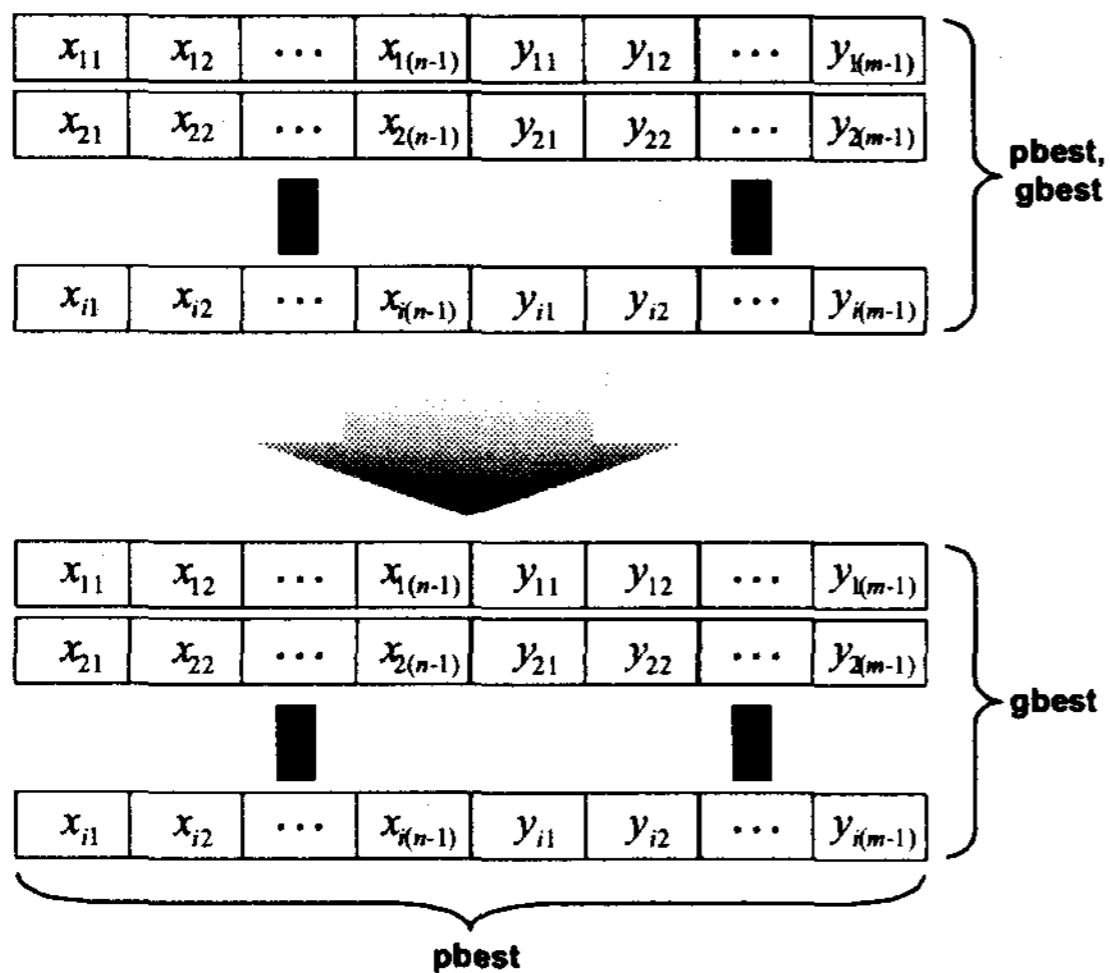


그림 1. 제안된 방법에서의 pbest와 gbest

퍼지 시계열에서의 예측 성능은 각각의 퍼지 규칙의 후건부(consequence)를 어떻게 설정하는가에 따라서 성능에 많은 영향을 받게 된다. 본 논문에서는 일반적인 PSO(Conventional PSO)에서 gbest와 pbest를 그림 1과 같이 전체 퍼지 집합(gbest)과 각각의 구간별 퍼지집합(pbest)으로 구분하여 탐색하는 PSO(Advanced PSO)방법을 제안한다.

4. 실험 및 결과

제안된 방법의 성능을 보이기 위하여, 1998년 8월 3일부터 9월 30일의 TAIFEX(Taiwan Futures Exchange)를 예측하는 문제에 적용하였으며, two-factor 퍼지 시계열을 위해 main-factor로 TAIFEX를 second-factor로 TAIEX(Taiwan stock exchange capitalization weighted stock index)를 사용하였다[3]. 각 퍼지 시계열의 구간 개수를 16개로 설정하였으며, 전체 객체의 수를 30개로 제한하여 실험을 하였다. 그림 2와 표 1은 기존의 방법들과의 성능 비교를 나타내었다. 표 1에서는 order의 변화에 따

른 각각의 성능을 비교하였으며, 제안된 방법이 기존의 방법들에 비하여 가장 우수한 성능을 보임을 알 수 있다.

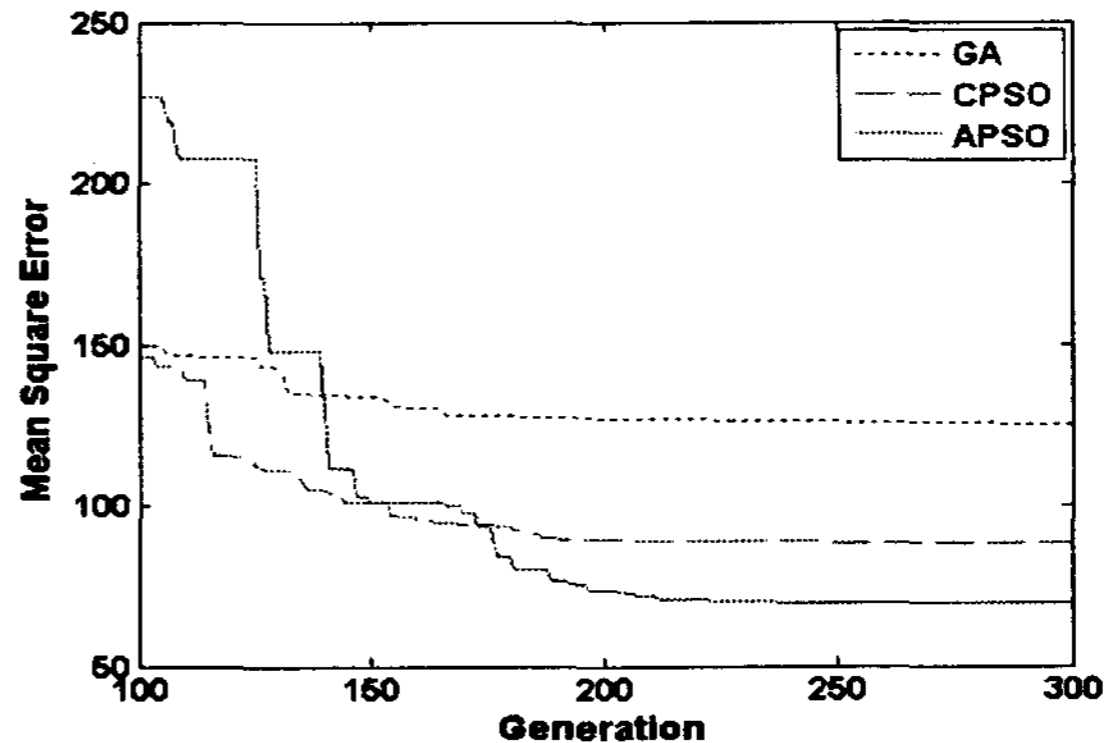


그림 2. 학습과정에서의 최적 객체의 적합도 변화 (Seventh order인 경우)

표 1. order의 변화에 따른 MSE의 비교

Algorithms	Fifth order	Sixth order	Seventh order	Eighth order
Lee et al. $\alpha=0.9$ [3]	143.31	147.14	105.02	124.45
CPSO	96.47	94.47	88.26	90.59
APSO	86.42	82.65	69.62	89.91

5. 결 론

본 논문에서는 퍼지 시계열에서 예측 성능을 향상시키는 방법으로 PSO를 이용하여 전체 퍼지 집합의 각 퍼지 구간의 길이를 최적화하는 기법을 제안하였다. 제안된 방법은 일반적인 PSO를 퍼지 시계열 예측에 적합한 형태로 개선함으로써, 실제 데이터를 통한 실험결과를 통하여 제안된 방법이 가장 우수한 성능을 가지고 있음을 확인 할 수 있었다.

참 고 문 헌

- [1] Lee, L. W., Wang, L. H., Chen, S. M., & Leu, Y. H. "Handling forecasting problems based on two-factors high-order fuzzy time series", *IEEE Trans. on Fuzzy Systems*, vol. 14, no. 3, pp. 468-477, 2006.
- [2] Kennedy, J., Eberhart, R. "Particle Swarm Optimization", *Proc. IEEE Int. Conf. Neural Networks*, vol. IV, pp. 1942-1948, 1995.
- [3] Lee, L. W., Wang, L. H., and Chen, S. M., "Temperature prediction and TAIFEX forecasting based on high-order fuzzy logical relationships and genetic simulated annealing techniques", *Expert Syst. with Applications*, vol. 34, pp. 328-336, 2008.