

시계열자료의 효율적 군집분석을 위한 구간특징화와 계층적 베이지안 기법의 융합

정영애*, 전진호**

선문대학교 IT학부^{*}, 가톨릭관동대학교 경영학과^{**}

A Fusion of the Period Characterized and Hierarchical Bayesian Techniques for Efficient Cluster Analysis of Time Series Data

Young-Ae Jung^{*}, Jin-Ho Jeon^{**}

Dept. of Information Technology Education, Sun-Moon University^{*}

Dept. of Business Administration, Catholic Kwan-Dong University^{**}

요 약 주가지표처럼 동적이며 시간흐름을 따르는 시계열자료들을 이해하는 효과적인 방법은 주어진 시계열자료들에 대하여 모델을 결정함으로써 이해하는 것이 좋다. 주어진 자료들에 대한 모델 결정과정은 수집되어진 대용량 시계열자료 전체를 한 번에 다 살펴보는 것보다 자료를 특성의 중요한 몇 개의 하위그룹으로 군집화하여 각 군집별 모델 결정을 통해 자료 전체를 이해하는 것이 효율적이다. 본 연구에서는 주어진 시계열자료들에 대하여 하위그룹으로의 효율적 군집화 과정 그리고 각 군집별 모델결정의 두 과정 중 첫 번째 과정인 하위집단으로 군집화 과정에 자료의 구간특징화 기법과 휴리스틱 베이지안기법의 융합을 이용하여 시간 및 계산비용을 감소시킬 수 있는 기법을 제안하였으며 실제적인 주가지표를 이용한 실험을 통해 제안하는 기법의 유효성을 확인하였다.

주제어 : 구간특징, 베이지안, 시계열자료, 융합, 군집

Abstract An effective way to understand the dynamic and time series that follows the passage of time, as valuation is to establish a model to analyze the phenomena of the system. Model of the decision process is efficient clustering information of the total mass of the time series data of the relevant population been collected in a particular number of sub-groups than to look at all a time to an understand of the overall data through each community-specific model determination. In this study, a sub-grouping of the group and the first of the two process model of each cluster by determining, in the following in sub-population characterized by a fusion with heuristic Bayesian clustering techniques proposed a process which can reduce calculation time and cost was confirmed by experiments using actual effectiveness valuation.

Key Words : Period Characterized, Bayesian, Time Series Data, Fusion, Clustering

Received 16 April 2015, Revised 20 May 2015

Accepted 20 July 2015

Corresponding Author: Jin-Ho Jeon
(Catholic Kwan-Dong University)

Email: jhgy@cku.ac.kr

ISSN: 1738-1916

© The Society of Digital Policy & Management. All rights reserved. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0>), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

1. 서론

경제시스템, 기업시스템, 의료시스템 및 공학 등의 많은 시스템에서 발생하는 다양한 형태의 자료 중 시간의 흐름에 따라 일정기간 동안 의미 있는 동적인 운동양태를 갖는 자료형태가 있다. 이렇게 시간의 흐름에 의해 표현되는 자료를 수집하여 기록한 자료를 시계열자료라 한다. 즉, 시계열자료(Time Series Data)란 동적으로 변화되는 값들의 연속으로 구성된 데이터시퀀스(Data Sequence)들의 집합이다.

우리나라의 경제시스템은 대외거래 규모의 확대 및 경제정책의 제도개선 등에 따라 지속적인 발전을 이루어 왔으며 개인뿐만 아니라 기업 및 다양한 사회구성 조직들에게 미치는 영향력이 커졌다. 이에 따라 경제시스템을 이해, 분석하고 정확한 의사결정의 필요성이 점차 커지고 있다. 이러한 경제시스템에서 경기변동을 예측하는 지표로 사용하는 것을 경기지표라고 한다. 이러한 경기지표들 중 현대 경제시스템을 대표하여 나타내는 주가지표는 국내외의 다양한 경제변수들에 의해 영향을 주고받는 상관관계가 밀접해짐에 따라 경제시스템의 전체를 이해하는데 좋은 지표라 할 수 있다. 이러한 주가지표처럼 동적이며 시간의 흐름에 따른 자료들을 이해하는 방법은 수집되어진 시계열자료를 분석하여 모델을 결정하는 것이다. 시계열자료와 같은 대용량의 자료들의 모델결정 과정은 수집되어진 자료의 전체 정보를 한 번에 다 살펴보는 것보다는 특정의 하위그룹으로 군집화하여 각 군집별로 모델결정을 통해 전체자료의 의미를 이해하는 것이 효과적이라 할 수 있다.

본 연구에서는 집단전체의 자료를 특정의 하위집단으로 군집화 과정에 있어서 시계열자료에 대하여 구간특징화와 계층적 베이지안 군집 방법을 이용하여 시간과 계산비용을 감소시킬 수 있는 효율적인 계층적 군집과정을 제안하는 것이 목적이며 다양한 경제지표 중 실제적인 주가지표를 통하여 실험하여 그 유효성을 확인하였다.

2. 관련 연구

군집화기법은 이질적이며 이해하기 어려운 전체 자료 집합에 대하여 동일 군집내에서는 유사도가 크게 각 군

집사이에서는 비유사도가 최대가 되도록 함으로서 전체 자료에 대하여 윤곽 파악을 쉽게 하는 것이 목적이다. 군집 기법들에 대한 연구는 계층적 기법들로부터 최적화기법들까지 다양한 영역에서 발전되어 왔다.

시계열자료들에 대한 군집기법 연구는 일반적으로 세 영역으로 구분되어진다. 첫 번째 영역으로 자료들 사이에 유사도 또는 거리측정을 이용하는 방식으로 상관측정(Correlation Measure), 헤밍거리(Hemming Distance)[1], 스트링편집거리(String Edit Distance)[2] 및 동적시간정합(Dynamic Time Warping)[3] 등이 있다. 두 번째 영역으로는 자료들의 특징을 추출하여 이용하는 방식으로 푸리에 기술자(Fourier Descriptor), 웨이블릿 분석(Wavlet Analysis)[4], MDL[5] 등이 있다. 세 번째 영역으로 주어진 자료들에 대하여 설명하는 모델결정의 방법으로 회귀 모델, 신경망, 마야코프체인[6,7,8] 등이 있다. 회귀모델은 주로 단기간의 자료들의 분석에 이용되므로 시계열자료에 대한 유용한 해석이 쉽지 않으며 신경망은 모델의 구조가 알려져 있으나 해석을 지원하지 않아 군집 모델링에는 적합하지 않다[7]는 단점을 가지고 있다. 마야코프체인 모델은 모델의 단순성 때문에 하나의 이산값을 시계열특징으로 묘사되는 시계열자료의 표현 모델링에 유용하다[6,9].

군집분석에서 꾸준히 제기되는 문제는 효율적인 군집의 수를 어떻게 선택하는가에 대한 문제이다. 보통의 일반적인 군집기법은 군집의 수를 미리 정하고 군집화 과정을 한다. 또 다른 기법으로서는 주어진 임계 기준값을 통하여 점진적 과정을 거쳐 가장 큰 임계값을 갖는 것을 유효한 군집 수로 결정하는 휴리스틱 기법이 있다. 본 연구에서는 시계열자료의 군집과정에 자료의 구간특징화와 마야코프체인 모델을 이용한 휴리스틱 기반의 베이지안 군집 과정을 제안한다.

3. 구간특징화를 이용한 베이지안 군집

본 연구에서 제안하는 군집기법은 주어진 자료를 구간화하여 각 구간을 특징하는 대표값으로 차원을 줄인 후 자료들에 대하여 베이지안 군집기법을 이용한다. 이러한 제안하는 기법을 통하여 기존의 군집과정보다 시간 및 계산비용의 절감 효율성을 얻을 수 있다.

3.1 마야코프체인 모델

마야코프체인 모델은 시간에 따른 상태변화를 파악하는 통계적 모델이다. 시간 t 에서 상태변화를 나타내는 상태전이확률 $A(t)$ 는 식(1)와 같다[6].

$$A(t) = \{a_{ij}(t)\}_{i,j=1}^N \quad (1)$$

여기서 $a_{ij}(t)$ 는 시간 $t-1$ 에서 상태가 i 일 때 시간 t 에서 상태가 j 로 전이될 확률을 나타낸다. 또한 상태 i 에서 다른 상태로 전이될 모든 확률의 합은 $\sum_{j=1}^N a_{ij}(t) = 1$ 이다. 마야코프체인 모델의 상태변화는 바로 이전 상태에 의존적인 특징을 갖는다. 즉, 시간 t 에서의 상태는 오직 시간 $t-1$ 에서의 상태에서만 영향을 받는다. 좀 더 확장하면 현재 시간 t 에서의 상태는 t 이전의 n 개의 상태에서만 영향을 받는다고 할 수 있으며 식(2)로 표현이 가능하다[6,10,11].

$$A(t) = \{a_{i_1 \dots i_n j}(t)\}_{i_1, \dots, i_n, j=1}^N \quad (2)$$

식(2)를 통해서 현재 $(n-1)$ 번째의 상태와 전이확률을 안다면 n 번째의 상태는 전이확률을 통해서 알 수 있다[6].

3.2 시계열자료의 구간특징화

각 시계열자료 개체 x_i 는 K 개의 특징들과 각 특징의 길이는 L 로서 표현된다. 즉, $K \times L$ 의 행렬로 표현된다. 이러한 시계열자료 개체들의 표현과 해석에서 쉬운 방법은 각 시간에서의 특징값들이 독립적이라고 가정하는 것이다. 이는 하나의 자료개체 x_i 는 L 개의 정적데이터 개체를 생산한다. 예를 들어, N 개의 자료개체가 있을 경우, $N \times L$ 의 자료개체 집합이 만들어진다. 이러한 시계열자료들의 계산비용 절감을 위해 구간특징화를 이용한다. 구간특징화 대표적 기법으로 구간상수화 기법을 살펴본다. 구간상수화(Piecewise Constant) 기법은 시계열자료 개체를 동일한 길이의 분할로 나눈 후 각 분할들에 대한 평균값을 특징으로 하는 기법이다. n 길이의 시퀀스 c 를 벡터 $\bar{c} = \bar{c}_1, \dots, \bar{c}_w$ 로 표현한다. \bar{c} 의 i 번째 요소

는 다음의 식(3)에 의하여 계산되어지며 이 값들은 차원이 축소된 표현이 된다[12]. 이러한 기법은 원래 시계열자료 개체들에 대하여 대응하는 변형된 시계열자료를 통해 최소계산을 보증하지만 시계열자료 개체가 내포하는 핵심적 특징행위는 구간특징화를 통해 얻는 것이 상대적으로 부족하다는 제약점이 있다[13].

$$\bar{C}_i = \frac{w}{n} \sum_{j=\frac{n}{w}(i-1)+1}^{\frac{n}{w}i} C_j \quad (3)$$

3.3 우도(Likelihood) 기준의 베이저안 군집

베이저안 군집은 모델기반 군집문제를 베이저안 모델 선택의 문제형태로 바꾼 것이다. 모델기반 군집화에서 자료는 확률분포의 혼합(Mixture)에 의해 생성되어지는 것을 가정하며 혼합모델 M 은 K 개의 모델들에 의해 표현되고 독립 이산변수 C 로 표현된다. C 의 각 값인 i 는 λ_i 에 의해 모델 되어지는 군집의 수를 표현한다.

자료 $X = (x_1, \dots, x_n)$ 이 주어지면, k 번째 λ_k 에 속하는 객체 x_i 집합확률을 $f(x_i|\theta_k, \lambda_k)$ 으로 표현한다. 혼합모델이 주어졌을 때 자료의 우도는 다음 식(4)으로 표현된다. 다음 식(4)에서 P_k 는 모델 λ_k 의 사전확률이다.

$$\begin{aligned} P(X|\theta, M) &= P(X|\theta_1, \dots, \theta_k, \lambda_1, \dots, \lambda_k) \quad (4) \\ &= \prod_{i=1}^N P(X_i|\theta_1, \dots, \theta_k, \lambda_1, \dots, \lambda_k) \\ &= \prod_{i=1}^N \sum_{k=1}^K P_k \cdot f(x_i|\theta_k, \lambda_k) \end{aligned}$$

베이저안 군집은 서로 다른 군집들을 갖는 분할들이 주어졌을 때, 가장 큰 사후확률을 갖는 가장 좋은 모델 M 을 선택하는 것이다. 베이즈 이론으로부터 모델의 사후확률은 $P(M|X) = \frac{P(M)P(X|M)}{P(X)}$ 에 의해 주어진다. $P(X)$ 는 데이터, $P(M)$ 은 모델의 사전확률을, $P(X|M)$ 은 데이터의 한계우도이다. 모델 비교 목적을 위해 서로 다른 모델들 사이에서 자료의 사전확률이 변하지 않고 유지되고 모든 모델들이 사전에 같다고 가정하면, 모든 모델들에 대하여 같은 사전확률을 할당할 수

있다. 즉, 사후확률은 자료의 한계우도에 비례한다. 그러므로 베이지안 선택 목적은 가장 큰 한계우도를 갖는 모델을 선택하는 것이다.

주어진 관측값들이 모델의 모든 매개변수에 대응되어지면 한계우도의 계산은 매우 간단하다. 하지만 주어진 자료가 불완전한 경우에는 정확한 해를 얻는 것은 매우 복잡하므로 일반적으로 근사기법들을 적용한다. 이러한 근사기법은 정확도는 낮지만 계산복잡도를 줄임으로서 계산비용 절감의 효율성을 얻을 수 있다[14]. 다음 식 (5),(6)과 같이 표현된다.

$$\log P(X|M) \approx \log P(X|\hat{\theta}, M) + \log P(\hat{\theta}|M) + \frac{d}{2} \log(2\pi) - \frac{1}{2} \log|A| \quad (5)$$

$$\log P(M | X) \approx \log P(X | M, \hat{\theta}) - \frac{d}{2} \log N \quad (6)$$

식(5),(6)에서 d 는 모델 매개변수 수, N 은 자료 수, $\hat{\theta}$ 는 모델에 대한 한계우도의 매개변수이다. 식(5)에서 첫 번째 항은 자료의 모델에 대한 우도항이며 두 번째 항은 매개변수의 수를 작게 하여 모델을 일반화시키는 항이다. 즉, 두 항의 조화를 통해 군집이 결정되며 하나의 군집으로부터 시작하여 하나씩 증가해 가며 가장 높은 값을 나타내는 군집 수가 유효한 군집 수를 나타내는 것이다 [15].

4. 실험

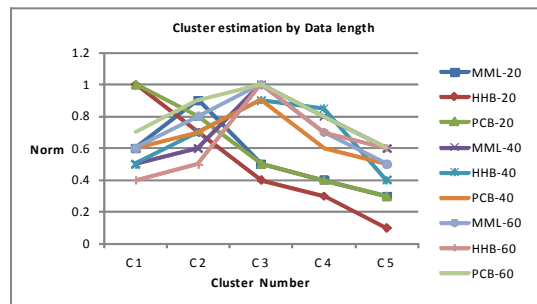
본 연구에서 제안하는 구간특징화를 이용한 계층적 베이지안 군집기법이 효율적인 군집을 추정하는지 실험을 통하여 살펴본다. 실험을 위한 자료길이는 매해 1월 2일을 기준으로 20일, 40일, 60일로 2012년, 2013년, 2014년의 코스피(KOSPI) 종합주가지수와 업종별에서 전기전자, 제조업, 유통업의 총 네 가지의 주가 자료를 선정하였으며 수집되어진 자료를 통해 마야코프모델을 생성한 후 생성된 모델로부터 각 모델별로 실험데이터 자료를 5개체, 10개체, 15개체를 생성하였다.

실험방법으로서 각 모델로부터 임의로 생성된 실험데

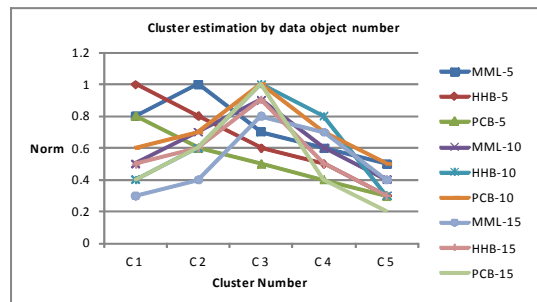
이터에 대한 군집과정은 첫 번째, 마야코프모델에 대하여 가장 큰 우도를 주는 모델에 할당하는 기법(Markov Model Likelihood : MML), 두 번째는 가장 큰 사후확률에 모델 복잡도에 따른 패널티값을 같이 고려하는 계층적 휴리스틱 베이지안 군집과정 기법 (Hierarchical Heuristic Bayesian : HHB), 세 번째는 두 번째 기법에 세그먼트(구간)를 3일로 설정한 후 구간에 대하여 구간특징화를 적용한 기법(Piecewise Character Bayesian : PCB)으로 세 방식에 대한 실험을 통해 제안한 구간특징화를 통한 베이지안 군집기법에 대한 유효성 및 효율성을 살펴본다.

4.1 세 업종에 대한 군집 추정

코스피, 전기전자, 유통업 세 가지 지수에 대한 군집추정을 실험하였다. 먼저 시계열자료의 길이에 따라 세 업종의 군집화가 정확하게 이루어지는지 확인하기 위하여 주어진 자료로부터 생성된 임의의 실험자료 길이를 20일, 40일, 60일로 하였으며 각 모델별로 실험자료 데이터는 5개체, 10개체, 15개체로 하였다.



[Fig. 1] Cluster estimation by data length



[Fig. 2] Cluster estimation by data object Number

[Fig. 1]는 마야코프모델 군집, 계층적 베이지안 군집, 구간특징화를 이용한 베이지안 군집기법 별로 시계열자료 길이를 20일, 40일, 60일 경우에 군집 수 추정결과를 보여주고 있다. [Fig. 2]는 각 군집기법별로 시계열자료 자료데이터 수를 5개체, 10개체, 15개체일 경우의 군집 수 추정결과를 보여주고 있다. [Fig. 1],[Fig. 2]에서 C1,C2,C3,C4,C5는 실험에 사용된 주어진 지수의 각 업종 별로 하나의 군집으로 보였다.

[Fig. 1],[Fig. 2]를 통해 시계열자료의 길이와 객체의 수가 상대적으로 짧거나 적은 경우는 부정확한 추정을 하는 것을 확인할 수 있다. 위의 그림을 표로 요약하면 <Table 1>,<Table 2>와 같다. <Table 1>,<Table 2>에서 보는 것처럼 자료의 길이가 40일, 60일처럼 충분한 길이의 자료와 자료 데이터가 10개체, 15개체처럼 충분한 자료의 객체에서는 세 기법 모두 정확하게 군집을 추정하는 것을 확인할 수 있다.

<Table 1> Cluster estimation by data lengths

	MML	HHB	PCB
Length 20	2	1	1
Length 40	3	3	3
Length 60	3	3	3

<Table 2> Cluster estimation by data object number

	MML	HHB	PCB
Object 5	2	1	1
Object 10	3	3	3
Object 15	3	3	3

하지만 자료의 길이가 상대적으로 짧은 20일과 자료 데이터가 상대적으로 적은 5개체에서는 세 기법 모두 정확하게 추정하지 못하는 것을 확인할 수 있다. 이유는 시계열자료의 길이가 짧을수록 또는 자료의 객체가 적을수록 자료 내에 내포된 규칙성을 정확하게 추출하여 반영하기에는 다소 어려움이 있다고 생각된다.

또한 세 모델들에서 생성된 실험 자료들을 통해 계층적 베이지안과 구간특징화 베이지안 군집이 아닌 마야코프모델 우도를 고려한 군집과정에서도 3군집이 아닌 2군집으로 잘못 추정된 이유는 실험에 사용된 세 군집 중 코스피(KOSPI)와 전기전자 모델에서 유사한 우도를 보여준 수치가 많았다. 이는 우리나라 대표기업인 삼성전자의 주가가 전기전자 업종별과 전체 코스피(KOSPI)지수

에 많은 영향을 주기 때문에 코스피(KOSPI)와 전기전자 업종 지수가 유사한 운동패턴을 보이는 것으로 생각된다.

4.2 네 업종에 대한 군집 추정

코스피(KOSPI), 전기전자, 제조업, 유통업 네 가지 지수에 대한 군집추정의 실험결과도 세 업종에 대한 군집의 실험결과와 같은 결과를 보여주고 있으며 <Table 3>,<Table 4>와 같다.

<Table 3> Cluster estimation by data lengths

	MML	HHB	PCB
Length 20	3	2	2
Length 40	4	4	4
Length 60	4	4	4

<Table 4> Cluster estimation by data object number

	MML	HHB	PCB
Object 5	3	2	2
Object 10	4	4	4
Object 15	4	4	4

<Table 3>,<Table 4>의 결과를 보면 앞서의 실험에서와 같이 충분한 시계열자료의 길이와 객체가 주어지면 자료 내에 내포된 규칙성을 잘 반영할 수 있음을 보여주고 있고 반대로 짧은 길이와 적은 객체가 주어진다면 정확하지 않은 군집추정 결과를 보여주고 있다. 즉, 시계열자료의 길이가 짧고 자료의 객체가 적다면 모든 기법들이 올바른 결과를 내기 어렵다. 그렇기 때문에 자료의 길이와 객체가 충분히 길고 많게 주어진 환경이라면 구간특징화를 통한 베이지안 군집과정이 올바른 군집추정 결과를 보여주기 때문에 시간 및 계산비용의 절감을 위하여 효율적으로 사용될 수 있다.

5. 결론

본 논문에서는 시계열자료의 군집과정에서 시간 및 계산비용의 절감을 위하여 구간특징화를 이용한 베이지안 군집과정을 제안하였다.

실제의 주가지수들을 이용하여 모델을 결정하고 모델로부터 실험데이터를 생성하여 실험을 하였으며 실험결

과, 짧은 길이와 적은 양의 자료 객체에서는 기존 기법과 제시한 군집과정이 시계열자료의 특징으로 인해 정확한 군집과정이 어려웠다. 하지만 충분한 자료의 길이와 충분한 양의 자료 객체가 주어진다면 기존의 기법들과 제안하는 구간특징화를 이용한 계층적 베이지안 군집과정도 같은 결과를 보여주는 것을 확인하였다. 특히 본 연구에서 제시된 기법을 이용한다면 점차 대용량으로 발생되는 시계열자료들의 군집과정에 시간 및 계산비용을 절감시킬 수 있으므로 효율적인 적용이 가능할 것이다.

향 후 더 다양한 부분의 일반화된 시계열자료에 적용 가능한 연구를 통하여 일반적이고 유연한 분석에 적용할 수 있는 부분으로 연구영역의 확대가 필요할 것이다.

REFERENCES

- [1] A. K. Jain and D. C. Dube, Algorithms for Clustering Data, Prentice Hall, 1988.
- [2] T. Okuda, E. Tanara and T. Kasai, "A Method for the Correction of Garbled Words based on the Levenshtein Metric", IEEE Transaction on Computers C25, 2, pp.172-177, 1976(2).
- [3] T. Oates, "Identifying Distinctive Subsequence in Multivariate Time Series by Clustering", Proceedings of the Sixteenth International Conference on Machine Learning, 1999.
- [4] Y. Huhtala, J. Karkkinen, H. Toivonen and N. R, "Mining for Similarity in aligned Time Series using Wavlets", Proceedings of SPIE on Data Mining and Knowledge Discover: Theory, Tools and Technology, 1999.
- [5] S. ManGanaris, "Learning to Classify Sensor Data", IJCAI'95 Workshop on Machine Learning in Engineering, 1995.
- [6] Y. Cho and G. Lee., "Modeling and Prediction of Time Seires Data based on Markov Model", The Korea Society of Computer and Information", Vol. 16, no. 2, pp. 225-233, 2011.
- [7] L. Rabiner., "A Tutorial on Hidden Markov Models and selected applications in speech recognition," Proc. of IEEE77, pp.257-286, 1989.
- [8] M. Siddiqi, J. Gordon and W. Moore., "Fast State Discovery for HMM Moel Selection and Learning," In Proc. Int'l Conference on Artificial Intelligence and Statistics, 2007.
- [9] J. Jeon., "A Study on Determining Prediction Models using Model-based Clustering of Time Series Data", Dankook Univ Ph. D, 2007.
- [10] J. Jeon and m. Kim., "A Study of Economic Indicator Prediction Model using Dimensions Decrease Techniques and HMM", The Journal of Digital Policy & Management, Vol. 11, no 10, pp305-311. 2013.
- [11] Y. Cho and G. Lee., "Prediction on Clusters by using Information Crtterion and Multiple Seeds", The Institute of Webcasting, Internet and Telecommunication, Vol. 10, no. 6, pp 145-152, 2010.
- [12] Y. Byungki and F. Christos., "Fast Time Sequence Indexing for Arbitrary Lp norms", In The VLDB Journal, pp 385-394, 2000.
- [13] L. Jessica, K. Eamonn, L. Stefano and C. Bill., "A Symbolic Representation of Time Series, with Implication for Streaming Algorithms", 8th ACM SIGMOD Workshop on Research Issues in DMKD, 2003.
- [14] J. Jeon and m. Kim., "A Study of Criterion for Efficient Clustering Estimation of Temporal Data", The Institute of Webcasting, Internet and Telecommunication, Vol. 11, no. 5, pp 139-144, 2011.
- [15] J. Jeon and m. Kim., "A Study on Prediction the Movement Pattern of Time Series Data using Information Criterion and Effective Data Length", The Institute of Webcasting, Internet and Telecommunication, Vol. 13, no. 1, pp 101-107, 2013.

정 영 애(Jung, Young Ae)



- 1991년 3월 ~ 1995년 2월 : 호서대 학교 컴퓨터과학 이학사
- 1998년 3월 ~ 2000년 2월 : 호서대 학교 컴퓨터과학 이학석사
- 2000년 3월 ~ 2007년 2월 : 단국대 학교 컴퓨터과학 이학박사
- 2009년 3월 ~ 현재 : 전문대학교 IT교육학부 교수

· 관심분야 : 소프트웨어공학, 인공지능, 정보 리터러시
· E-Mail : mariachoi911@gmail.com

전 진 호(Jeon, Jin Ho)



- 1990년 3월 ~ 1994년 2월 : 관동대 학교 경영학과 경영학사
- 1996년 3월 ~ 1998년 8월 : 명지대 학교 경영정보학과 경영학석사
- 2000년 3월 ~ 2007년 2월 : 단국대 학교 컴퓨터과학 이학박사
- 2009년 9월 ~ 현재 : 가톨릭관동대 학교 경영학과 조교수

· 관심분야 : 지능형시스템, 경영정보시스템, IT전략 및 정책
· E-Mail : jhgy@cku.ac.kr