

경매 시스템에서 시계열 분석에 기반한 낙찰 예정가 추천 방법

고 민 정* · 이 용 규**

Reserve Price Recommendation Methods for Auction Systems Based on Time Series Analysis

Min Jung Ko* · Yong Kyu Lee**

Abstract

It is very important that sellers provide reasonable reserve prices for auction items in internet auction systems. Recently, an agent has been proposed to generate reserve prices automatically based on the case similarity of information retrieval theory and the moving average of time series analysis. However, one problem of the previous approaches is that the recent trend of auction prices is not well reflected on the generated reserve prices, because it simply provides the bid price of the most similar item or an average price of some similar items using the past auction data. In this paper, in order to overcome the problem, we propose a method that generates reserve prices based on the moving average, the exponential smoothing, and the least square of time series analysis. Through performance experiments, we show that the successful bid rate of the new method can be increased by preventing sellers from making unreasonable reserve prices compared with the previous methods.

Keywords : Time Series Analysis, Reserve Price, Agent, Internet Auction System

1. 서 론

전자상거래가 보편화되면서 인터넷 경매를 통한 물품 구매가 증가하고 있다. 이러한 인터넷 경매 시스템에서는 판매자가 경매 물품에 대하여 적절한 낙찰 예정가를 제시하는 것이 매우 중요하다[Peter F. Drucker, 1991 ; Eric V. Heck & Peter Vervest, 1998]. 기존의 경매 시스템에서는 경매 물품에 대해서 판매자가 낙찰을 원하는 최저 가격인 낙찰 예정가(Reserve Price) eBay Auction Guideline, 2002]를 추천하지 못하여, 판매자들의 주관에 따라서 정해야 하는 어려움이 있었다. 이로 인하여 경매 물품에 대한 초기 낙찰률이 낮아져서 판매자들이 재입찰에 참여 하거나 중간 조정자들이 오프라인으로 판매자와 구매자 사이의 거래를 유도해야 하는 문제점이 있었다[서울 자동차 경매장 매물 리스트, 2002 ; 파인드물 경매 자동차 리스트, 2003 ; 폰 사이트, 2004 ; 폰모아, 2004]. 또한 경매 물품이 지나치게 낮은 낙찰가에 낙찰되는 경우가 있어서 판매자가 손해를 보는 경우가 많았다.

이를 해결하기 위하여, 최근 인터넷 경매에서 정보검색의 사례 유사도와 시계열 분석의 이동평균에 의하여 낙찰 예정가를 자동으로 생성하는 방법이 제안되었다[고민정 외 3인, 2002 ; 김신우 외 3인, 2002 ; Lee, Y. K. 외 3인, 2002 ; 고민정, 이용규, 2004]. 그러나 전자는 과거 전체 기록을 기반으로 가장 유사한 사례의 낙찰가로 낙찰 예정가를 생성하여, 경매 물품의 낙찰 가격에 영향을 미치는 최근의 추세를 반영하지 못하는 경우가 많았다. 또한 후자는 이러한 문제는 해결하였으나, 경매 물품에 따라서 과거 기록의 중요성이 달라지는 경우에 객관적인 가중치를 부여하기가 어려워 잘못된 낙찰 예정가를 생성하는 경우가 발생한다. 그러므로 경매 물품의 특성에 적합한 낙찰 예정가를 추천하여

초기 낙찰률을 높이고, 이로 인하여 판매자의 재입찰과 오프라인 거래를 줄이며, 판매자의 입장에서 손해를 줄일 수 있는 방법이 필요하다.

본 논문은 이러한 문제점들을 해결하기 위하여 기존의 연구로서 판매자가 낙찰 예정가를 결정할 때 최근 자료에 더 높은 가중치를 부여한 이동 평균이 우수한 성능을 보인 것을 확장한다[고민정, 이용규, 2004]. 즉 지수 평활이나 최소자승과 같은 다양한 시계열 분석을 적용하여 경매 물품에 적합한 낙찰 예정가 추천 방법을 제안하고, 실험 평가 척도에 의하여 각 방법의 성능을 비교하고 분석한다. 또한 성능 실험을 통하여 시계열 분석이 인터넷 경매의 낙찰 예정가 추천에 적합하고, 경매 물품에 따라 다르게 제안하는 낙찰 예정가 추천 방법이 실제 낙찰가와 가까워서 경매의 초기 낙찰률을 높이고, 판매자가 재입찰이나 불합리한 가격에 물품이 낙찰되는 사례를 방지할 수 있음을 보인다.

본 논문의 구성은 다음과 같다. 먼저, 2장에서는 본 논문과 관련된 이론과 기존 인터넷 경매 시스템의 낙찰 예정가 추천 방법에 대하여 살펴본다. 그리고 3장은 본 시스템에서 낙찰 예정가를 자동으로 추천하는 방법을 시계열 분석에 따라서 설계하고, 4장에서는 성능 실험을 하고, 그 결과에 대하여 분석하고 비교한다. 또한, 5장에서는 본 논문의 결론과 향후 연구 과제를 정리한다.

2. 관련 연구

기존의 인터넷 경매 시스템에서 낙찰 예정가를 추천하기 위하여 제안된 방법의 특징을 분석하고, 이를 해결하기 위하여 적용된 시계열 분석을 살펴본다.

2.1 경매 물품의 낙찰 예정가 결정법

기존에 인터넷 경매 물품의 가격과 관련된 연

구로는 인터넷 경매에서 낙찰 예정가와 입찰가 생성에 사례 유사도를 이용하는 방법이 있었다[고민정 외 3인, 2002 ; 김신우 외 3인, 2002 ; Lee, Y. K. 외 3인, 2002]. 이는 벡터 공간 모델의 사례 유사도를 이용하여 과거 경매 기록을 대상으로 경매 물품과의 유사도를 분석하고, 가장 유사한 경매 물품의 낙찰 가격으로 낙찰 예정가를 자동으로 생성하였다[고민정 외 3인, 2002 ; 김신우 외 3인, 2002]. 여기서는 과거 전체 경매 기록과 시장 가격 데이터베이스의 가격 정보를 기준으로 각각 가장 유사한 사례 3건에 가중치를 부여하여 낙찰 예정가를 계산하고, 두 데이터베이스 사이의 가중치를 7 : 3으로 적용한 낙찰가를 낙찰 예정가로 생성하였다. 하나, 이는 낙찰 예정가를 생성하는 대상 정보가 저장된 과거 기록인지 시장 가격인지 여부를 고려하여 가중치를 부여하므로, 경매 물품의 최근 가격 추세를 반영하기 어려운 경우가 많았다.

이러한 문제를 해결하기 위하여 시계열 분석의 이동 평균을 이용하여 최근 몇 개의 가격 정보에 가중치를 두어 추세가 반영된 낙찰 예정가를 생성하는 방법이 제안되었다[고민정, 이용규, 2004]. 그러나 여기서도 최근의 가격 추세가 포함된 낙찰 예정가는 생성되지만, 경매 물품에 따라서 과거 실제 낙찰가의 가중치를 정하는데 주관의 개입되어 실제 낙찰 가능가와 차이가 큰 낙찰 예정가를 생성하는 경우가 발생하는 문제점이 있었다.

따라서 본 논문에서는 경매 물품에 따라서 다양한 시계열 분석을 적용하여 낙찰 예정가 생성하고, 이를 비교, 분석하여 판매자에게 손해를 줄이고, 낙찰률을 높일 수 있는 낙찰 예정가 추천 방법을 제시하고자 한다.

2.2 시계열 분석

기존의 인터넷 경매 시스템에서는 경매 물품

에 대하여 판매자가 낙찰 예정가를 제시하는데 도움을 주기 위하여 과거의 실제 낙찰 가격 정보를 제공한다[eBay Auction Guideline, 2002]. 이를 기초로 하여 판매자는 경매 물품에 적당한 낙찰 예정가를 제시하게 되고, 여기서 제공된 과거 기록이 미래의 낙찰 예정가를 결정하는 중요한 요인이 된다. 이러한 과거 기록은 시계열 자료의 특성을 가지는데, 여기서 시계열이란 어떤 현상에 대한 계량적 자료가 일정시간 마다 주어져 있는 것을 말하며, 이를 이용하여 미래를 예측하는 것이 시계열 분석(Time Series Analysis)이다[이순용, 1998].

여기에는 이동 평균(Moving Average)과 지수 평활(Exponential Smoothing), 최소 자승(Least Square) 등이 있다[Blake LeBaron, 1992 ; Robert H. Shumway & David S. Stoffer, 2002]. 이 중에 이동 평균은 과거 자료의 불규칙적인 변동을 고르게 하여 예측 오차를 줄이는 방법이며, 여기에는 단순 이동 평균(Simple Moving Average), 가중 이동 평균(Weighted Moving Average)이 있다[오길록, 1975 ; 조신섭 & 손영수, 1999]. 그리고, 지수 평활이란 현재부터 과거 모든 기간의 실적에 지수적으로 작은 가중치를 반영하는 방법이다[지철원, 1992 ; Blake LeBaron, 1992]. 여기에는 단순 지수 평활(Simple Exponential Smoothing), 이중 지수 평활(Double Exponential Smoothing)이 있다[Holt C. Carbone, 1957 ; Michael P. Clements 외 2인, 2004]. 또한 최소 자승은 실제치와 예측치 편차 자승의 총합이 최소가 되도록 평균선을 그려서 예측하는 방법이다.

이들은 수요나 판매 예측, 주식 시장 분석, 이익 분석 등에 많이 활용된다[Peter F. Drucker 외 5인, 1991 ; Charles P. Jones, 2001 ; Robert H. Shumway & David S. Stoffer, 2002]. 그러나, 이 방법들이 인터넷 경매 시스템에서 경매

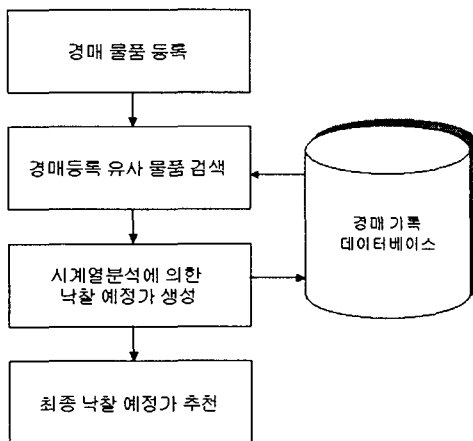
물품의 낙찰 예정가 추천에 이용된 경우는 찾아보기 어렵다. 본 논문에서는 이들을 이용하여 낙찰 예정가를 추천하는 방법을 제시하고, 성능 실험을 통하여 결과를 비교 분석한다.

3. 낙찰 예정가 추천 방법

시계열 분석을 인터넷 경매 시스템에 적용하여 판매자를 위한 낙찰 예정가를 자동으로 추천하는 방법을 설계한다.

3.1 낙찰 예정가 추천 절차

인터넷 경매시스템에서 경매 물품의 낙찰 예정가를 추천하는 과정은 다음 <그림 1>과 같다.

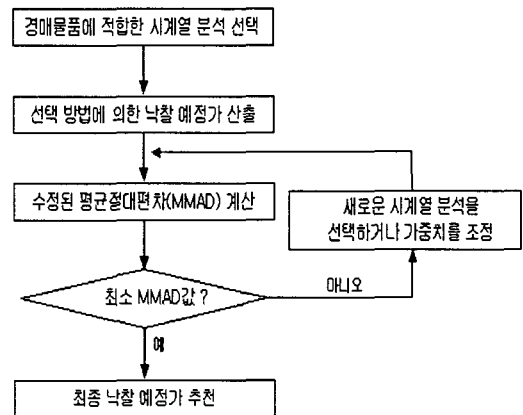


<그림 1> 낙찰 예정가 추천 절차

먼저, 경매 물품을 등록하고, 등록된 물품 중에서 가장 유사한 물품을 검색한다. 이것은 경매 기록 데이터베이스를 대상으로 물품의 가격 결정에 영향을 미치는 요소로 자동차의 경우는 연식과 주행거리, 핸드폰의 경우는 사용 기간과 모델명을 기준으로 이루어진다. 여기서 검색된 자료를 기초로 시계열 분석을 적용한 낙찰 예정가를 산출하게 되는데, 그 방법은 다음 <그

림 2>에서 설명한다. 그리고 마지막 단계에서는 제안된 가격으로 최종 낙찰 예정가를 결정한다.

시계열 분석에 의하여 낙찰 예정가를 추천하는 절차는 다음 <그림 2>와 같다. 우선, 경매 물품에 대하여 하나의 시계열 분석을 선택하고 이를 이용하여 초기 낙찰 예정가를 계산한다. 이렇게 생성된 낙찰 예정가는 실제 낙찰가를 기준으로 성능 평가 척도를 산출하고, 이 값이 가장 작은 것을 선택하기 위하여 가중치를 조정해가면서 성능 평가 척도인 수정된 평균 절대 편차(Modified Mean Absolute Deviation)를 비교한다. 이를 통해서 결정된 가중치를 적용하여 선택된 시계열 분석을 적용하여 최종 낙찰 예정가를 추천하게 된다.



<그림 2> 시계열 분석에 의한 낙찰 예정가 추천 흐름도

본 논문에서는 시계열 분석의 이동 평균, 지수 평활, 최소 자승을 이용하여 낙찰 예정가를 추천하는데, 그 방법과 성능 실험 척도는 다음 절들에서 설명한다.

3.2 시계열 분석에 의한 낙찰 예정 추천 방법

시계열 분석을 인터넷 경매 시스템에 적용하

여 낙찰 예정가를 자동으로 추천하는 방법을 제시한다.

3.2.1 이동 평균에 의한 낙찰 예정가 산출

이동 평균은 과거 일정기간의 실제 낙찰가를 평균해서 새로운 낙찰 예정가를 예측하는 방법이며, 여기에는 단순 이동 평균과 가중 이동 평균이 있다[이순용, 1998 ; Robert H. Shumway & David S. Stoffer, 2002]. 본 논문에서는 이들의 기본 공식에 과거 경매 물품의 실제 낙찰가를 적용하여 낙찰 예정가를 추천하는데, 그 방법은 다음과 같다.

과거 여러 기간의 가격 정보에 동일한 가중치를 부여하는 것이 단순 이동 평균이고[이순용, 1998], 이를 적용하여 낙찰 예정가를 추천하는 방법은 식 (1)과 같다.

$$V_t = \frac{1}{n} \sum_{i=1}^n (P_{t-i}) \quad (1)$$

V = 낙찰 예정가

P = 경매 물품의 실제 낙찰가

n = 이동 평균을 구하기 위한 기간

($n = 1, 2, \dots, 12$)

t = 시간을 나타내는 인덱스

이 방법은 현재부터 과거의 일정 기간 동안 검색된 낙찰가의 평균으로 예상 낙찰 예정가를 결정하는데, 여기서는 낙찰하고자 하는 시점에서 과거의 누적 개월 수를 기준으로 이동 평균의 기간을 결정하고, 이를 통해서 낙찰 예정가를 계산한다.

가중 이동 평균은 최근 가격 정보에 가장 높은 가중치를 부여한 평균이며[Leonard J. Tashman, 2000], 낙찰 예정가는 다음 식 (2)에 의해서 구해진다.

$$V_t = \sum_{i=1}^n (w_{t-i} \cdot P_{t-i}) \quad (2)$$

V = 낙찰 예정가

P = 경매 물품의 실제 낙찰가

n = 이동 평균을 구하기 위한 기간

($n = 1, 2, \dots, 12$)

t = 시간을 나타내는 인덱스

w = 가격에 대한 가중치

$$\left(\sum_{i=1}^n w_i = 1 \right)$$

본 논문에서는 현재부터 과거 순으로 높은 가중치를 부여하고, 개월 수를 기준으로 0과 1사이에서 0.1 단위로 가중치를 적용하여 낙찰 예정가를 추천한다. 예를 들면 이동 평균을 구하기 위한 기간이 3개월인 경우에는 최근으로부터 가중치를 3 : 2 : 1로 반영한 3개월의 평균으로 낙찰 예정가를 생성하며, 여기에 이용된 가중치의 합은 1이다.

3.2.2 지수 평활을 이용한 낙찰 예정가 산출

지수적으로 감소된 과거의 정보와 증가되는 현재의 가격 정보를 기초로 하여 미래를 예측하는 것이 지수 평활이며[Robert H. Shumway & David S. Stoffer, 2002], 여기에는 단순 지수 평활과 이중 지수 평활이 있다[이순용, 1998]. 다음 절들에서 이를 통하여 인터넷 경매 물품의 낙찰 예정가를 추천하는 방법을 제시한다.

단순 지수 평활[Robert H. Shumway & David S. Stoffer, 2002]은 경매 물품의 과거 낙찰가를 지수적으로 작게 반영하는 것이며, 본 논문에서는 다음 식 (3)과 같은 방법으로 낙찰 예정가를 계산한다.

$$V_t = \alpha \cdot P_{t-1} + (1 - \alpha) \cdot V_{t-1} \quad (3)$$

V = 낙찰 예정가

P = 경매 물품의 실제 낙찰가

t = 시간을 나타내는 인덱스
 α = 지수 평활 계수 ($0 < \alpha < 1$)

이는 가중 이동 평균에서 가중치 부여에 대한 주관성을 피하기 위하여 지수 평활 계수를 이용한다. 여기서는 지수 평활 계수를 0에서 1 사이에 0.1 간격으로 적용한다. 예를 들어 지수 평활 계수 0.2는 경매 물품 등록 시점의 최근 실제 낙찰가를 20%, 나머지 과거의 실제 낙찰가를 지수적으로 감소시켜 80%로 반영하여 낙찰 예정가를 추천하는 것을 의미한다.

정해진 기간 내에 상향이나 하향의 추세가 있는 자료에 지수 평활을 두 번 적용하는 방법이 이중 지수 평활이며[Michael P. Clements 외 2인, 2004], 이를 적용하여 낙찰 예정가를 다음 식 (4)과 같이 구한다.

$$V_t = \alpha \cdot P_t + (1 - \alpha) \cdot (V_{t-1} + b_{t-1}) \quad (4)$$

$$b_t = \beta \cdot (V_t - V_{t-1}) + (1 - \beta) \cdot b_{t-1} \quad (5)$$

V = 낙찰 예정가

P = 경매 물품의 실제 낙찰가

t = 시간을 나타내는 인덱스

b = 추세를 나타내는 요소

α = 1차 지수 평활 계수 ($0 < \alpha < 1$)

β = 2차 지수 평활 계수 ($0 < \beta < 1$)

여기서 식 (5)에 의하여 계산되는 b 는 경매 물품의 상향이나 하향의 추세를 반영하는 값이며, 이를 통해서 식 (4)에 경매 물품의 추세를 반영할 수 있다. 본 논문에서는 지수 평활 계수인 α , β 의 값을 0에서 1사이에 0.1 단위로 동일하게 적용하고, b 의 초기값으로 가장 최근의 실제 낙찰가를 적용한다.

3.2.3 최소 자승에 의한 낙찰 예정가 산출

다음 식 (6)은 예측치와 실제치의 편차 자승의 총합이 최소가 되도록 동적인 평균선을 이용

하는 최소 자승[이순용, 1998]을 기초로 여기에 경매 물품의 판매 시점과 실제 낙찰가를 적용하여 낙찰 예정가를 계산한 것이다.

$$V_t = X + Y \cdot D_t \quad (6)$$

$$Y = \left(\sum_{i=1}^n (D_i - \overline{D}) \cdot (P_i - \overline{P}) \right) / \left(\sum_{i=1}^n (D_i - \overline{D})^2 \right) \quad (7)$$

$$X = \overline{P} - Y \cdot \overline{D} \quad (8)$$

V = 낙찰 예정가

P = 경매 물품의 실제 낙찰가

t = 시간을 나타내는 인덱스

D = 경매 물품의 판매 월

\overline{P} = 경매 물품의 실제 낙찰가의 평균

\overline{D} = 경매 물품의 판매 월의 평균

X, Y = 회귀선을 이루는 계수

여기서 식 (7)과 식 (8)에서 계산된 X, Y 를 적용한 평균선으로 낙찰 예정가를 생성하고, 이를 통해서 판매자가 경매 물품을 등록하는 시점을 기초로 한 예상 낙찰 예정가를 식 (6)에 의하여 생성한다.

4. 성능 실험

시계열 분석을 이용한 낙찰 예정가 추천 방법들에 대한 실험 환경을 소개하고, 실험 결과를 비교, 분석한다.

4.1 성능 실험 환경

성능 실험 대상 자료의 구축 방법과 실험 시나리오를 설명하고, 각 방법의 성능 평가 기준을 설명한다.

4.1.1 성능 실험 대상

성능 실험을 위해서 실제 경매가 이루어지는

중고 자동차[서울 자동차 경매장 매물 리스트, 2002 ; 파인드몰 경매 자동차 리스트, 2003 ; 글로비스(주) 중고차 시세, 2004]와 핸드폰[중폰 사이트, 2004 ; 폰모아, 2004] 사이트에 제시된 과거 낙찰가 기록을 이용하였는데, 이 중에서 동일한 자동차 모델 자료 2854 건과 최근 2년간의 핸드폰 모델 자료 1512 건을 경매 기록 데이터베이스로 구축하였다. 그리고 각각의 경매 물품에 대하여 과거 실제 낙찰 가격을 1년간 월별로 저장하여 활용하였다.

또한 성능 실험을 위하여 인터넷 경매 사이트[엔카 매물 리스트, 2004 ; 폰모아 사이트, 2004]에서 임의로 중고 자동차와 중고 핸드폰을 각각 413건과 103건을 수집하여 경매 기록 데이터베이스를 대상으로 낙찰 예정가를 추천하는 성능 실험을 한다. 이들은 경매 물품의 낙찰된 시점에 저장된 실제 낙찰가와 예측 오차를 생성하여 성능이 평가된다.

4.1.2 성능 실험 시나리오

본 연구는 실험 자료를 대상으로 시계열 분석 중에서 가장 성능이 우수한 모델을 선택하기 위하여 이동 평균, 지수 평활, 최소 자승을 이용하여 성능 실험을 실시한다. 여기서 이동 평균의 단순 이동 평균과 가중 이동 평균을 적용하고, 지수 평활의 단순 지수 평활과 이중 지수 평활을 활용하며, 최소 자승을 이용한다. 각 방법 내에서 예측 오차가 가장 작은 가중치를 찾아내고, 이들을 비교하여 성능이 우수한 것을 최종 낙찰 예정가로 추천한다.

또한 기존에 낙찰 예정가를 제안 받지 못한 경우에는 사용자가 경매 물품에 대하여 낙찰을 희망하는 낙찰 희망가를 이용하였고, 이를 본 논문에서는 No Agent로 표기한다. 또한 이전의 연구였던 사례 유사도에 기초하여 과거 전체 기록에서 가장 유사한 경매 물품 3건의 낙찰가에

가중치를 부여하여 낙찰 예정가를 생성하여 본 논문에서 제시한 새로운 방법으로 추천된 낙찰 예정가와 성능 실험을 실시한다. 여기서는 평가 척도를 기준으로 오차나 낙찰률 등을 통하여 성능 실험이 이루어지는데, 이들에 대한 실험 결과와 분석은 다음 절들에서 제시한다.

4.1.3 성능 평가 척도

시계열 분석을 이용한 낙찰 예정가 추천의 성능을 평가하기 위하여 수요 예측의 오차 측정에 많이 사용되는[이순용, 1998 ; Robert H. Shumway & David S. Stoffer, 2002] MAD(Mean Absolute Deviation)를 이용하는데, 이는 추천된 낙찰 예정가와 실제 낙찰가 차의 절대 값을 합하여 예측 기간 수로 나눈 것이다. 여기서는 MAD를 다음 식 (9)와 같이 각각의 실제 낙찰가로 나누어 수정된 MAD(Modified Mean Absolute Deviation)로 조정하여 적용한다. 왜냐하면 실험 자료인 중고 자동차와 핸드폰의 기본 가격이 서로 다르므로 실제 낙찰가를 기준으로 한 편차를 계산하여 비교하기 위해서이다.

$$E = \frac{1}{n} \sum_{i=1}^n (|b_i - r_i|) / b_i \quad (9)$$

n = 물품의 수

b_i = 실제 낙찰가

r_i = 낙찰 예정가

(고객의 최저 낙찰 희망가)

본 논문에서는 식 (9)에서 물품 수로 나누지 않은 값을 예측 오차(기존의 방법에서는 오차)라고 정의하며, E 의 값이 작을수록 예측 오차의 크기가 작으며, 이것을 이용하여 낙찰 예정가와 실제 낙찰가와 유사성을 분석한다.

4.2 실험 결과

실험 자료를 대상으로 실험한 결과를 분석하

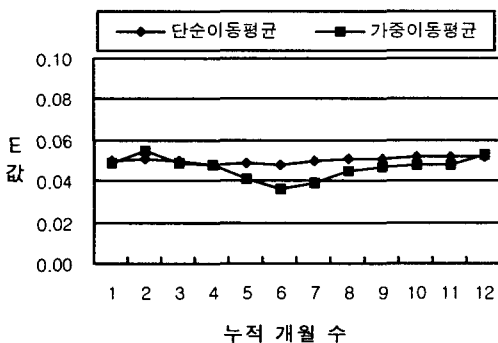
고, 이를 통하여 낙찰 예정가를 자동으로 추천하는 방법의 장점과 특징을 비교한다.

4.2.1 시계열 분석의 성능

실제 낙찰가와 가장 유사한 낙찰 예정가를 추천하는 방법을 선택하기 위하여 시계열 분석의 이동 평균, 지수 평활, 최소 자승을 이용한 성능 실험을 실시하고, 이들의 E 값을 비교하여 가장 성능이 우수한 방법을 선택한다.

(1) 이동 평균의 성능

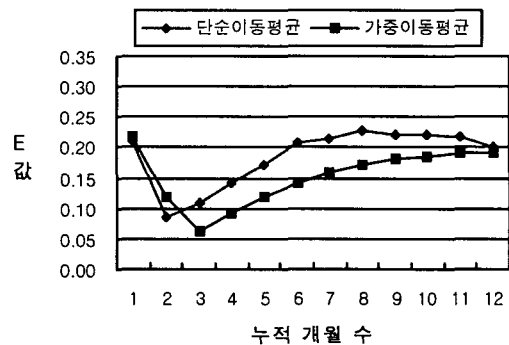
이동 평균에 의하여 낙찰 예정가를 추천하는 성능 실험을 위하여 단순 이동 평균은 식 (1)을 이용하고, 가중 이동 평균은 식 (2)을 적용하여 누적 개월 수에 따른 E 값을 생성하였고, 실험 결과는 다음 <그림 3>과 <그림 4>이다.



<그림 3> 자동차의 E 값 비교

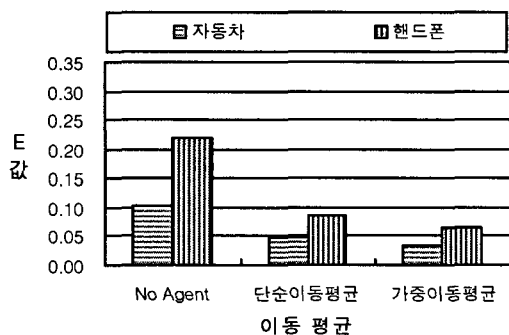
<그림 3>은 x축의 누적 개월 수에 따른 자동차의 E 값을 나타내는 그래프이다. 예를 들어 x축이 4인 경우는 단순이동 평균은 이전 4개월의 평균으로 낙찰 예정가를 추천하고, 가중 이동 평균은 현재를 기준으로 4:3:2:1의 가중치를 반영한 4개월의 평균으로 낙찰 예정가를 추천한 것이다. 결과는 단순 이동 평균의 4개월과 가중 이동 평균의 6개월의 E 값이 다른 가중치에 비해 비교적 작음을 알 수 있다. 이는 과거 1

년 동안의 실제 낙찰가를 동일하게 적용하는 단순 이동 평균 보다 최근을 기준으로 더 높은 가중치를 부여하는 가중 이동 평균의 성능이 우수함을 의미한다.



<그림 4> 핸드폰의 E 값 비교

<그림 4>는 핸드폰에서 누적 개월 수에 따른 E 값이고, 결과는 단순 이동 평균의 경우는 2개월, 가중 이동 평균의 경우는 3개월에 대한 E 의 값이 가장 작음을 보여 준다. 이 경우도 최근의 실제 낙찰가에 높은 비중을 반영하는 가중 이동 평균의 성능이 우수함을 알 수 있다.



<그림 5> 이동 평균에 의한 E 값 비교

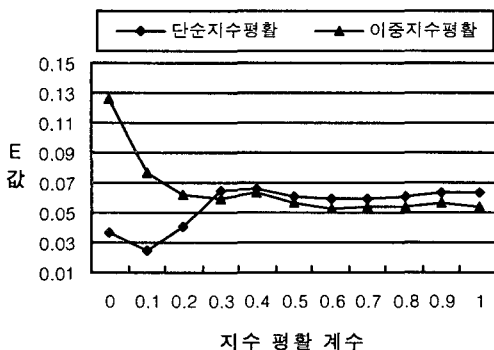
<그림 5>는 실험 자료인 자동차 413건과 핸드폰 103건을 대상으로 낙찰 예정가를 추천 받지 못한 기존의 방법과 이동 평균에 의한 결과를 나타내는 그래프이다. 여기서 이동 평균을 이용하는 경우는 각 방법 내에서 우수한 성능을

보인 누적 개월 수를 적용하여 실험하였다. 즉 자동차와 핸드폰에 대하여 4 개월과 2 개월의 단순 이동 평균과 6 개월과 3 개월의 가중 이동 평균을 적용하여 실험한 E 값이다.

결과는 낙찰 예정가를 추천 받지 못한 이전의 방법보다 이동 평균에 의한 E 값이 작아서 우수함을 보여준다. 더욱이 이동 평균 방법의 단순 이동 평균 경우에 E 값은 0.048과 0.086이고, 가중 이동 평균은 0.035와 0.064이며 두 경우 모두 가중 이동 평균의 성능이 더 우수함을 나타낸다. 이들은 모든 자료에 동일한 가중치를 부여하는 단순 이동 평균보다 최근 자료에 높은 가중치를 부여한 가중 이동 평균이 실제 낙찰가와 더 가깝고, 경매 물품의 낙찰 시점을 중심으로 한 최근 몇 개월의 자료가 과거 1년 동안의 모든 실제 낙찰가를 반영하는 것보다 성능이 좋아짐을 의미한다.

(2) 지수 평활의 성능

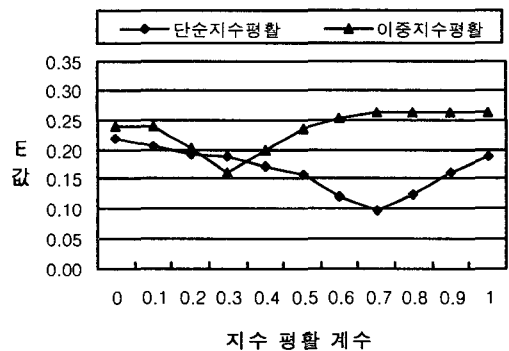
지수 평활에 의하여 낙찰 예정가를 추천하는 방법으로 단순 지수 평활은 식 (3), 이중 지수 평활은 식 (4)를 적용하여 낙찰 예정가를 추천하였고, 결과는 다음 <그림 6>과 <그림 7>이다.



<그림 6> 자동차의 E 값 비교

<그림 6>은 자동차에서 지수 평활 계수에 따

른 E 값을 나타내며, x 축 0.3의 의미는 낙찰하려는 바로 이전 달의 실제 낙찰가와 나머지 과거 1년간의 실제 낙찰가의 비율을 3 : 7로 반영하는 것을 의미한다. 결과는 단순 지수 평활은 지수 평활 계수가 0.1인 경우와 이중 지수 평활은 0.6인 E 값이 작아서 우수함을 보여준다. 여기서 바로 이전 달과 나머지 1년간의 실제 낙찰가 비율을 1 : 9로 반영한 단순 지수 평활 방법이 6 : 4로 반영한 이중 지수 평활보다 성능이 우수 것은 중고 자동차의 경우에 계절에 따라서 실제 낙찰가의 변동이 크지 않아서 상향이나 하향을 반영하는 이중 지수 평활 방법의 성능이 좋지 않음을 의미한다.

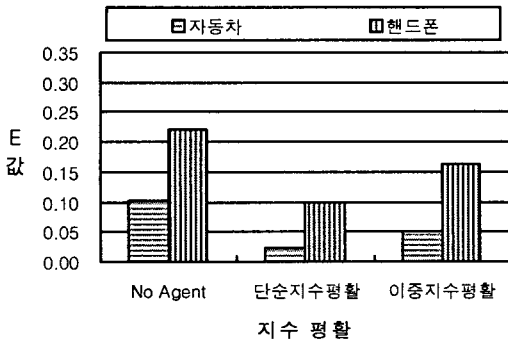


<그림 7> 핸드폰의 E 값 비교

<그림 7>은 핸드폰의 경우로서 지수 평활은 0.7, 이중 지수 평활은 0.3에 의한 E 의 값이 작음을 보여준다. 이것은 낙찰 시점 바로 이전 달의 실제 낙찰가와 과거 1년간의 실제 낙찰가를 단순 지수 평활은 7 : 3, 이중 지수 평활이 3 : 7로 반영하여 낙찰 예정가를 생성하는 것의 성능이 우수함을 의미한다. 여기서도 계절에 따른 변화 요인을 고려하지 않으면서 실제 낙찰가에 영향을 많이 받는 낙찰 예정가의 성능이 우수함을 알 수 있다.

<그림 8>은 실험 자료인 중고 자동차와 핸드폰을 대상으로 각 기법에서 가장 성능이 우수한

지수 평활 계수 이용하여 E 값을 기존의 낙찰 예정가를 추천 받지 못한 방법과 비교한 것이다.



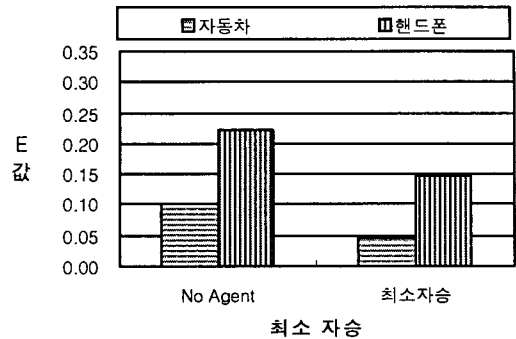
〈그림 8〉 지수 평활에 의한 E 값 비교

결과는 지수 평활에 의한 E 값이 기존의 방법보다 작아서 우수함을 보여준다. 또한 지수 평활을 이용한 경우에도 자동차와 핸드폰에서 단순 지수 평활의 E 값은 0.025과 0.098, 이중 지수 평활의 E 는 0.053와 0.163로 단순 지수 평활의 성능이 우수함을 알 수 있다. 이들은 낙찰 시점 이전 달 실제 낙찰가와 나머지 과거 1년간의 모든 가격 정보의 비율을 자동차의 경우는 90%와 10%, 핸드폰의 경우는 70%와 30%를 반영하여 새로운 낙찰 예정가를 생성하는 것의 성능이 우수함을 나타낸다. 또한 경매 물품의 낙찰 예정가를 생성하는 경우에 과거 1년간의 실제 낙찰가를 지수적으로 작게 반영하는 것이 실제 낙찰가와 유사함을 나타내고, 실험 자료가 년도나 계절에 따라서 실제 낙찰가의 변화가 적기 때문에 이를 반영하는 이중 지수 평활의 성능이 단순 지수 평활보다 비교적 좋지 못한 것이고, 이것은 실험 자료에 따라서 달라질 수 있다.

(3) 최소 자승의 성능

다음 <그림 9>는 판매자의 주관에 따라서 낙찰 예정가를 결정하는 기존의 방법과 경매 물품의 경매 시기와 실제 낙찰가를 기준으로 식 (6)

의하여 낙찰 예정가를 추천한 결과이다.

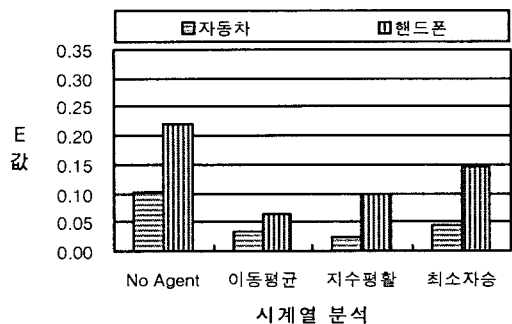


〈그림 9〉 최소 자승에 의한 E 값 비교

결과는 최소 자승에 의한 E 값이 기존의 방법보다 우수하며, 이것은 경매 물품의 경매 시기와 실제 낙찰가를 기준으로 한 최소 자승이 실제 낙찰가와 가까운 가격을 생성하는 방법임을 의미한다.

(4) 시계열 분석의 성능 비교

<그림 10>은 실험 자료에 대하여 기존에 낙찰 예정가를 추천하지 못하는 방법과 시계열 분석에 의해서 낙찰 예정가를 결정하는 E 값의 결과를 비교한 것이고, 시계열 분석은 각 방법 내에서 가장 우수한 성능을 보인 가중치를 적용한다.



〈그림 10〉 시계열 분석에 의한 E 값 비교

결과는 시계열 분석에 의한 E 값이 기존의

방법보다 작아서 우수함은 보여준다. 더욱이 자동차와 핸드폰에서의 E 값이 이동 평균의 경우 0.035와 0.064, 지수 평활의 경우는 0.025과 0.098, 최소 자승은 0.046와 0.148로 기존 시스템의 0.103와 0.221 보다 작음을 나타낸다. 또한 실험 결과 중에서도 자동차의 경우는 지수 평활, 핸드폰의 경우는 이동 평균에 의한 낙찰 예정가의 E 값이 가장 작아서 우수함을 보여준다. 이들은 자동차는 낙찰 시점 바로 이전달의 실제 낙찰가와 과거 1년간의 모든 가격 정보를 지수적으로 감소시켜서 반영하고, 핸드폰은 낙찰 시점 이전 3개월의 가중 평균으로 생성된 낙찰 예정가가 실제 낙찰가와 가장 가까움을 의미한다. 또한 실험 대상인 중고 자동차의 낙찰가 예정가는 과거의 모든 자료를 기반으로 한 추천이 적합하며, 핸드폰의 경우는 최근 일부 가격만을 적용하여 가격을 생성하는 것이 적합함을 보여준다. 본 논문에서는 이들을 적용하여 새로운 낙찰 예정가를 추천한다.

4.2.2 새로운 방법의 성능

시계열 분석을 이용하는 새로운 방법과 기존의 방법들을 실제 낙찰가를 기준으로 오차와 낙찰률을 비교하는 성능 실험을 실시한다.

(1) 실제 낙찰가와 오차 비교

성능 실험을 위하여 자동차 413건과 핸드폰 103건을 대상으로 낙찰 예정가를 제시하지 못하는 기존 방법과 제시하는 방법 중에서도 이전의 연구인 사례 유사도와 시계열 분석에 의한 새로운 방법으로 실제 낙찰가와 오차를 계산한다. 이것은 추천된 낙찰 예정가와 실제 낙찰가의 차를 각각의 실제 낙찰가로 나누어 생성하고, 낙찰 예정가를 추천 받지 못하는 기존 방법의 경우에는 낙찰 예정가를 대신하여 고객 최저 낙찰 희망가를 적용한다. 여기서 계산된 값을

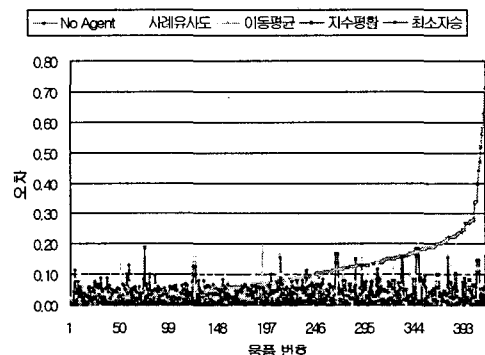
기존의 방법에서는 오차라고 사용하고, 추천된 방법에서 생성된 값은 예측 오차라고 정의한다.

〈표 1〉 오차와 예측 오차 비교

		No Agent	Agent			
			사례 유사도	이동 평균	지수 평활	최소 자승
자동차	최소 오차	0.000	0.000	0.000	0.000	0.000
	최대 오차	0.714	0.265	0.192	0.189	0.184
	E 값	0.103	0.059	0.035	0.025	0.046
핸드폰	최소 오차	0.038	0.014	0.000	0.001	0.001
	최대 오차	0.578	0.420	0.277	0.236	0.330
	E 값	0.221	0.186	0.064	0.098	0.148

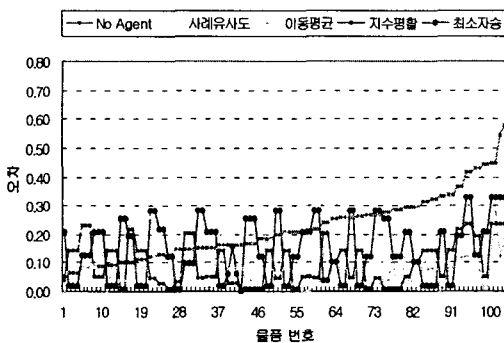
〈표 1〉은 새로운 방법의 예측 오차가 기존의 오차 보다 작아서 경매 물품에 대하여 추천된 낙찰 예정가의 성능이 우수함을 나타낸다. 특히 이전의 연구인 사례 유사도에 의한 방법보다 시계열 분석에 의한 최소와 최대 예측 오차의 범위와 E 값이 더 작아서 성능이 우수함을 보여준다.

〈그림 11〉과 〈그림 12〉는 실험 대상인 자동차 413건과 핸드폰 103건의 오차와 예측 오차를 보여주며, 이것은 기존 방법으로 계산된 오차의 오름차순으로 정렬된 그래프이다.



〈그림 11〉 자동차의 오차와 예측 오차 비교

<그림 11>은 사례 유사도나 새로운 방법으로 생성한 낙찰 예정가의 예측 오차 분포를 나타내며, 기존 방법의 오차는 0.000에서 0.0714로 나타나고 새로운 방법들의 예측 오차는 0.000에서 0.0192로 분포되어 있음을 볼 수 있다. 이는 새로운 방법에 의한 예측 오차가 지나치게 낮거나 높지 않은 범위에 있음을 의미한다.



<그림 12> 핸드폰의 오차와 예측오차 비교

<그림 12>는 핸드폰 103건에서 오차와 예측 오차를 분포를 나타내는데, 여기에서도 기존의 방법보다 사례 유사도의 경우나 새로운 방법에 의한 예측 오차의 범위가 작음을 알 수 있다. 이는 새로운 방법에 의한 낙찰 예정가 생성 방법이 실제 낙찰가와 차이점을 기존의 방법보다 줄여서 판매자가 경매 물품을 등록할 때 경매 물품에 대하여 지나치게 높은 낙찰 예정가를 결정하여 경매 물품이 유찰되는 경우와 지나치게 낮은 가격에 낙찰되는 경우를 방지할 수 있음을 의미한다.

(2) 성공한 경우의 오차 비교

<표 2>는 낙찰에 성공한 경우에 실제 낙찰가를 기준으로 한 오차와 예측 오차를 비교한 것이다. 즉 낙찰에 성공한 경매 물품만을 대상으로 실제 낙찰가와 낙찰 예정가(기존 방법은 고객 최저 낙찰 희망가)의 차를 실제 낙찰가로 나눈 최소, 최대 오차를 나타낸다. 여기서는 추천

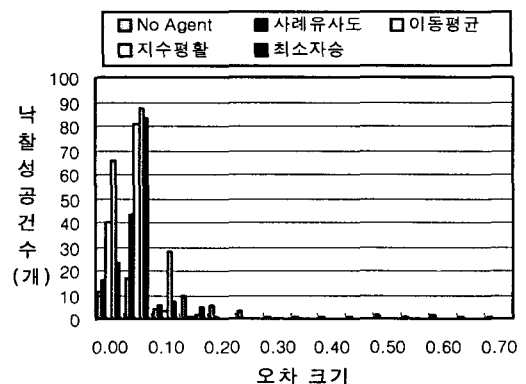
된 낙찰 예정가가 실제 낙찰가와 동일하거나 낮은 경우를 낙찰 성공으로 간주한다.

<표 2> 낙찰에 성공한 경우의 오차와 예측 오차 비교

		No Agent	Agent			
			사례 유사도	이동 평균	지수 평활	최소자승
자동차	최소 오차	0.00	0.00	0.00	0.00	0.00
	최대 오차	0.67	0.25	0.14	0.07	0.14
핸드폰	최소 오차	0.04	0.06	0.00	0.01	0.00
	최대 오차	0.04	0.33	0.08	0.23	0.28

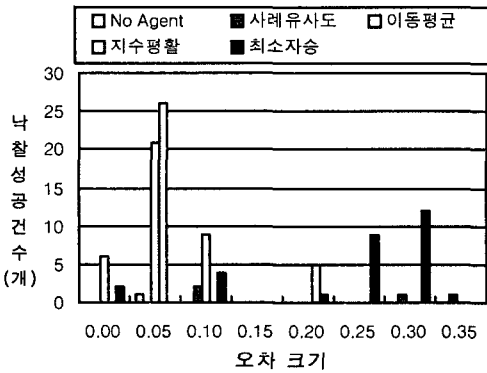
낙찰에 성공한 경우도 <표 2>와 같이 새로운 방법에 의한 예측 오차의 범위가 기존의 방법이나 사례 유사도를 이용하는 경우보다 작음을 알 수 있다. 여기서 핸드폰의 경우에 기존의 방법이 우수한 것은 낙찰 건수가 1건인 경우만 있어서 나타난 결과이고, 실험 자료가 많아질 경우는 달라질 것이 예상된다.

<그림 13>과 <그림 14>는 오차의 크기에 따른 낙찰의 성공 건수를 나타내는 그래프이다. 여기서 x 축은 오차의 크기를 나타내며, 예를 들어 0.05는 오차의 크기가 0.01에서 0.05 사이에 있는 낙찰 성공 건수를 보여준다.



<그림 13> 자동차의 오차 크기에 따른 낙찰 성공 건수 비교

자동차의 경우에 <그림 13>과 같이 새로운 방법의 낙찰 성공 건수가 0.00에서 0.15에 주로 나타나는데, 기존의 방법은 0.00에서 0.70까지 넓게 분포되어 있는 것을 보여준다. 이는 낙찰에 성공한 경우에 기존의 방법은 판매자가 지나치게 낮은 가격에 경매 물품을 낙찰하여 손해를 입었음을 의미한다.



<그림 14> 핸드폰의 오차 크기에 따른 낙찰 성공 건수 비교

<그림 14>는 핸드폰의 경우에는 새로운 방법에서도 자동차보다는 오차의 범위가 넓게 낙찰에 성공한 것을 보여준다. 그러나 이는 핸드폰의 실험 자료가 103건으로 자동차의 403건에 비해서 적기 때문이고, 실험 자료가 증가할 경우 오차의 범위가 줄어들 것으로 예상된다.

(3) 전체 낙찰률 비교

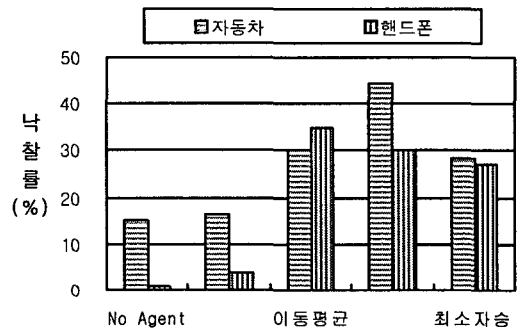
<표 3>와 <그림 15>는 각 기법들에 대하여 실험 자료인 중고 자동차와 중고 핸드폰 자료 각각 413건과 103건 내의 낙찰률을 분석한 것으로서, 기존 방법은 재입찰을 고려하지 않은 초기 입찰의 경우만을 적용한다.

<표 3> 전체 낙찰 건수에 대한 실험 결과 데이터

	No Agent	Agent			
		사례 유사도	이동 평균	지수 평활	최소자승
자동차	15%	16%	30%	44%	29%
핸드폰	1%	4%	35%	30%	27%

결과는 <표 3>과 같이 새로운 방법에 의한 예측 낙찰률이 기존의 낙찰률의 15%와 1%보다 최대 44%와 35%로 높아짐을 알 수 있고, 이는 추천된 낙찰 예정가가 실제 낙찰률을 높일 수 있음을 의미한다.

<그림 15>는 기존의 방법을 이용한 실제 낙찰률과 새로운 방법을 통하여 생성된 낙찰 예정가의 예측 낙찰률을 나타내는 그래프이다.



<그림 15> 실제 낙찰률과 예측 낙찰률 비교

결과는 기존의 방법이나 사례 유사도에 의한 낙찰 예정가 생성보다 새로운 방법에 의한 낙찰률이 모두 높음을 볼 수 있다. 더욱이 자동차의 경우는 지수 평활에 의한 방법의 성능이 우수하고, 핸드폰은 이동 평균에 의한 예측 낙찰률이 높음을 알 수 있다. 이는 새로운 방법을 이용하면 실제 낙찰가와 가까운 낙찰 예정가를 생성하여 기존의 인터넷 경매 시스템에서 초기 낙찰률이 낮아서 재입찰을 하는 경우를 줄임으로서 조정가에 의한 오프라인 거래를 감소시킬 수 있음을 의미한다.

5. 결론 및 향후 연구

본 논문에서는 인터넷 경매 시스템에서 판매자가 경매 물품을 등록하는 시점을 반영하여 시계열 분석에 기초한 낙찰 예정가를 추천하는 방법을 제안하였고, 성능 실험을 통하여 이전의

방법과 비교, 분석하였다.

여기서 제시한 낙찰 예정가 자동 추천 방법을 이용하여 중고 자동차와 핸드폰을 대상으로 성능 실험을 하였는데, 중고 자동차는 경매 물품 등록 시점의 실제 낙찰가가 새로운 낙찰가 생성에 중요한 영향을 미치면서 과거의 모든 가격 정보가 고려되는 단순 지수 평환에 의한 낙찰 예정가 추천이 우수한 성능을 보였다. 반면에 중고 핸드폰은 과거의 모든 자료보다는 가중치가 부여된 경매 물품 등록 시점 이전 3개월의 실제 낙찰가를 반영한 가중 이동 평균에 의한 낙찰 예정가의 성능이 우수하였다. 이를 적용한 실험 결과는 새로운 시스템의 낙찰률이 자동차와 핸드폰에서 44%와 35%로 기존 방법의 15%와 1%, 이전의 연구인 사례 유사도 방법의 16%와 4% 보다 높아짐을 알 수 있었다. 또한 낙찰에 성공한 경우에 실제 낙찰가를 기준으로 한 최대 예측 오차가 자동차의 경우에 0.14로 기존 방법의 0.67, 사례유사도 방법의 0.25 보다 작아서 판매자가 손해를 보는 경우를 줄일 수 있음을 보였다.

이들은 새로운 방법이 경매 물품에 따라서 시계열 분석에 기초하여 가장 성능이 우수한 낙찰 예정가를 추천하여 실제 낙찰가와와의 차이를 줄여서 낙찰률을 높일 수 있고, 판매자가 경매 물품의 가격을 지나치게 높게 제시하여 유찰되거나 낮아서 손해를 보는 경우를 방지할 수 있음을 의미한다. 그리고 이를 통해서 판매자가 경매 물품의 초기 낙찰률이 낮아서 재입찰을 하거나 조정자가 오프라인으로 거래에 참여 하는 경우를 줄임으로서 시스템의 효율성을 높일 수 있다. 또한 성능 분석을 통하여 시계열 분석이 인터넷 경매 시스템의 낙찰 예정가 추천에 적합함을 보였고, 이를 이용하여 새로운 형태의 낙찰 예정가 추천 방법을 제시한 것이 본 논문의 의의라고 할 수 있다.

향후에는 다양한 경매 물품을 가진 대규모 데

이터베이스를 대상으로 기존의 사례 유사도에 최근의 추세를 반영하는 방법과 ARIMA등과 같은 고급 시계열 분석법들을 적용하여 낙찰 예정가를 생성하고, 이들 사이의 성능 실험을 통하여 본 연구의 실용성 검증이 필요하다. 또한 새로운 시스템 기술을 이용하여 인터넷 경매 시스템을 구현하는 것이 요구된다.

참 고 문 헌

- [1] 고민정, 김신우, 박성은, 이용규, "인터넷 경매 시스템에서의 낙찰 예정가 자동생성 에이전트", *정보처리학회논문지*, 제9-D권 제5호, 2002년 10월, pp. 955-962.
- [2] 고민정, 이용규, "이동 평균을 이용한 인터넷 경매 시스템의 낙찰 예정가 자동생성", *한국전자거래학회지*, 제9권 제2호, 2004년 5월, pp. 17-31.
- [3] 김신우, 고민정, 박성은, 이용규, "공동 구매 시스템에서의 낙찰 예정가 및 입찰가 자동생성", *한국전자거래학회지*, 제7권 제2호, 2002년 8월, pp. 55-68.
- [4] 이순용, 생산 관리론, 법문사, 1998.
- [5] 오길록, *이단계 이동평균법에 의한 단기 판매 예측*, 한국과학기술원 산업공학과 석사학위논문, 1975.
- [6] 조신섭, 손영수, *시계열 분석*, 울곡 출판사, 1999.
- [7] 지철원, *시계열 예측 의사결정 지원시스템: ARMA 모형 수립을 위한 신경망 접근방법*, 한국과학기술원 경영과학과 박사학위논문, 1991.
- [8] Holt C. Carbone, "Forecasting Trends and Seasonals by Exponentially Weighted Moving Averages", *Carnegie Institute of Technology*, Vol. 2, No. 52, 1957.
- [9] Michael P. Clements, Philip H. Franses and Norman R. Swanson, "Forecasting Economic and Financial Time-series with

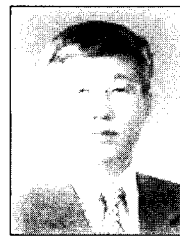
- Non-linear Models," *Int'l Journal of Forecasting*, Vol. 20, No. 2, April 2004, pp. 169-183.
- [10] Peter F. Drucker, et al., *Accurate Business Forecasting*, Harvard Business School Press, 1991.
- [11] Eric V. Heck and Peter Vervest, "How Should CIO's Deal with Web-Based Auctions?" *Communications of the ACM*, Vol. 41, No. 7, July 1998, pp. 99-100.
- [12] Charles P. Jones, *Investments : Analysis and Management*, 8th Edition, Wiley, 2001.
- [13] Blake LeBaron, *Do Moving Average Trade Rule Imply Nonlinearities in Foreign Exchange Markets?*, Technical Report 9222, Social Systems Research Institute, Univ. of Wisconsin, Madison, 1992.
- [14] Yong K. Lee, Shin W. Kim, Min J. Ko and Sung E. Park, "Pricing Agents for a Group Buying System", *EurAsia ICT 2002*, Lecture Notes in Computer Science(LNCS), Vol. 2510, 2002, pp. 693-700.
- [15] Robert H. Shumway and David S. Stoffer, *Time Series Analysis and Its Application*, Springer, 2002.
- [16] Leonard J. Tashman, "Out-of-sample Tests of Forecasting Accuracy : an Analysis and Review", *Int'l Journal of Forecasting*, Vol. 16, No. 4, October 2000, pp. 437-450.
- [17] eBay Auction Guideline, <http://webhel-lp.ebay.com/cgi-bin/eHNC>, 2003.
- [18] 서울 자동차 경매장 매물 리스트, http://www.saaa.co.kr/sell/main_sell.html/sell/sise/retail_sise.asp, 2002.
- [19] 파인드물 경매 자동차 리스트, <http://auto.findall.co.kr/BuyAuto/Type/BuyCarList.asp>, 2003.
- [20] 엔카 매물 리스트, <http://www.encar.co.kr/proc/23/23000.jsp>, 2004.
- [21] 글로비스(주) 중고차 시세, <http://www.hkaa.co.kr/Sise/SerCond01.aspx>, 2004.
- [22] 중론 사이트, <http://used4989.com/home2>, 2004.
- [23] 폰모아 사이트, <http://www.phonemoa.co.kr>, 2004.

□ 저자소개



고민정

경기대학교 전자계산학과에서 학사, 이화여자대학교 교육대학원 컴퓨터교육학과에서 석사를 마치고, 현재 동국대학교 컴퓨터공학과에서 박사과정 중에 있으며 동국대학교 전자계산원 교수로 재직 중이다. 주요 관심분야는 XML 및 웹, 전자상거래 시스템, 전자상거래 보안, 데이터베이스이다.



이용규

동국대학교 전자계산학과에서 학사, 한국과학기술원 전자학과에서 석사를 마치고 Syracuse University에서 전자학 박사학위를 취득하였다. 정보통신부 행정직 국가공무원, 한국국방연구원 선임연구원, 한국통신 선임연구원, 콜로라도대학교 컴퓨터학과 방문교수를 거쳐 현재 동국대학교 컴퓨터멀티미디어공학과 교수로 재직 중이며 주요 관심 분야는 XML 및 웹, 전자상거래 시스템, 스토리지 시스템, 데이터베이스이다.