A Computational Approach to Predicting Economic Regimes in Automated Exchanges

Wolfgang Ketter, John Collins, Maria Gini, Alok Gupta⋆, and Paul Schrater
Department of Computer Science and Engineering
⋆Department of Information and Decision Sciences
University of Minnesota
{ketter,jcollins,gini,schrater}@cs.umn.edu, agupta@csom.umn.edu

Abstract

We present a computational approach to identify dominant market conditions, such as over-supply or scarcity, and to predict market changes in automated exchange environments. Intelligent agents can learn the characteristics of prevailing economic conditions, or *regimes*, from historical data. Agents can then use real-time observable information to identify the current market regime and forecast upcoming market changes. We show that different market regimes can be effectively identified using our methodology. We also present preliminary work on a method to predict regime transitions. We experimentally validate our approach with data from the Trading Agent Competition for Supply Chain Management.

1 Introduction

Electronic marketplaces are gaining popularity among producers seeking to streamline their supply chains and consumers looking for opportunities. Intelligent software agents can significantly facilitate human decision processes either by helping users to select strategies or by making autonomous choices that are consistent with a human decision maker's preferences.

The Trading Agent Competition for Supply Chain Management [6] (TAC SCM) is a market simulation in which six competing autonomous agents attempt to maximize profit by buying parts, assembling products (personal computers), and selling their products in daily auctions. The goal is to achieve the highest bank balance at the end of the game, which runs for one simulated year. An agent can produce 16 different types of computers, that are categorized into 3 different market segments. Demand in each market segment varies during the game.

An agent has to make many decisions such as how many parts to buy, when to have them delivered, what types of computers to build, when to sell them, and at what price. Availability of parts and demand for computers varies randomly through the game and across low, medium and high price market segments. The small number of agents in the game, their ability to adapt and change strategy during the game, and to manipulate the markets makes the game highly dynamic, uncertain, and challenging.

In this paper we examine how market conditions, such as oversupply or scarcity, can be detected and exploited by an autonomous agent. The long term objective of our work is to show how knowledge of current and anticipated market conditions can enable an agent to make better operational and strategic decisions, such as how many parts to order, how to schedule production, and how to price its products. While this type of prediction about the economic environment is commonly used at the macro economic level, such predictions are rarely done for a micro economic environment. After a brief discussion of related work, we describe our method to autonomously identify distinguishable economic conditions. The method uses historical data from previous games and real-time data available to the agent during the game. Finally, we present an approach to forecasting regime transitions.

2 Related Work

Many researchers have tried to characterize market conditions to inform strategic and operational decisions. For instance, [5] discusses how strategic windows of change alternate with long periods of stability in mature economic markets.

Sales strategies used in previous TAC SCM competitions have attempted to model the probability of receiving an order for a given offer price, either by estimating the probability by linear interpolation from the minimum and maximum daily price records [4], or by estimating the relationship between offer price and order probability with a reverse cumulative density function (CDF) and factors such as quantity and due date [2]. All these methods fail to take into account market conditions that are not directly observable. They are essentially regression models, and do not represent qualitative differences in market conditions. The method we present here, in contrast, is able to detect and forecast a broader range of market conditions for durable goods. Regression based approaches (including non-parametric variations) assume that the functional form which defines the relationship between dependent and independent variables has a constant structure. However, experience shows that these functional relationships often have a different structures for different regimes.

3 Economic Regime Identification

We believe that market conditions can be characterized by statistical patterns, and that such patterns can be learned from historical data. We call these distinguishable market conditions *regimes*.

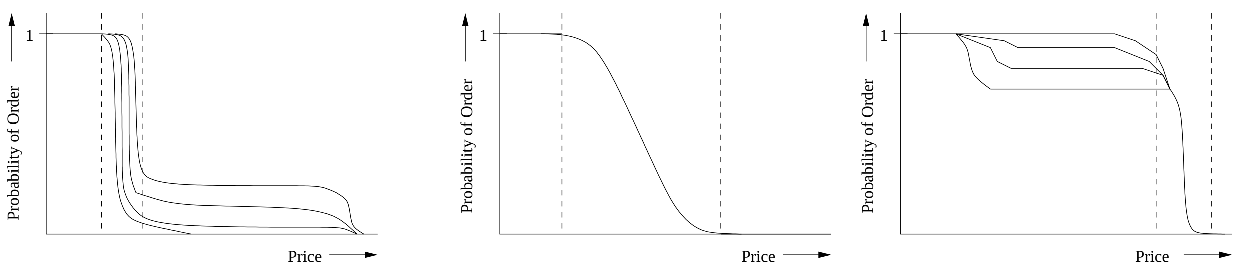

Figure 1: Typical order probability curves during over-supply (left), balanced (middle) and scarcity (right) regimes.

Figure 1 shows typical curves for the probability of receiving an order for a given offer price. The slope of the curve and its position changes over time. According to economic theory, high prices and a steep slope correspond to a situation of scarcity, where price elasticity is small, while a less steep slope corresponds to a balanced market where the range of prices is larger.

Clearly changes in market conditions should affect the strategy of the agent. When there is scarcity, prices are higher, so an agent can price aggressively. In balanced situations prices are lower and have more spread, so the agent has a range of options for maximizing expected profit. In over-supply situations prices are lower, and the agent should focus on controlling costs.

We identify and characterize regimes by analyzing data from previous games. The agent can then use these results along with data available in real-time during the game to identify regimes, forecast regime transitions, and adapt its procurement, production, and pricing strategies accordingly. For our experiments, we used data from a set of 26 games played during the semi-finals and finals of TAC SCM 2004.

We use a Gaussian mixture model (GMM) over past game data to identify and define regimes.

In the game, each computer type has a nominal price, which is the sum of the nominal costs of its parts. To allow comparisons between different computer types in a market segment, we normalize the

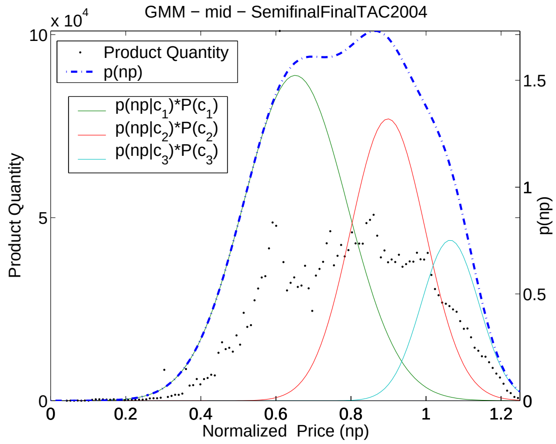

Figure 2: The Gaussian mixture model for the medium market segment.

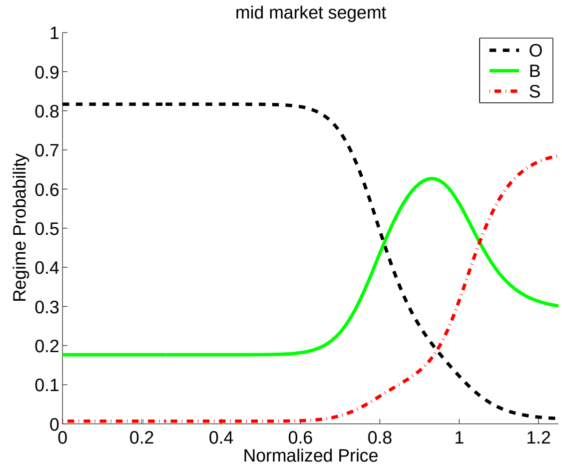

Figure 3: Regime probabilities over normalized price for the medium market segment.

prices by using the ratio of market price to nominal price. We call np the normalized price. We apply the EM-Algorithm [1] to determine the Gaussian components c_i of the GMM, $N[\mu_i, \sigma_i](\mathrm{np})$, and their prior probability, $P(c_i)$. The density of the normalized price[1] can be modeled as:

$$p(\mathrm{np}) = \sum_{i=1}^{N} p(\mathrm{np}|c_i)\, P(c_i) \tag{1}$$

where $p(\mathrm{np}|c_i)$ is the contribution of the i-th Gaussian to the normalized price density of the GMM. An example is shown in Figure 2.

For our experiments we chose $N = 3$, because we found experimentally that this provides a good balance between quality of approximation and simplicity of processing. Using Bayes' rule we determine the posterior probability:

$$P(c_i|\mathrm{np}) = \frac{p(\mathrm{np}|c_i)\, P(c_i)}{\sum_{i=1}^{N} p(\mathrm{np}|c_i)\, P(c_i)} \qquad \forall i = 1, \cdots, N \tag{2}$$

We then define the N-dimensional vector $\vec{\eta}(\mathrm{np})$, whose components are the posterior probabilities from the GMM,

$$\vec{\eta}(\mathrm{np}) = [P(c_1|\mathrm{np}), P(c_2|\mathrm{np}), \ldots, P(c_N|\mathrm{np})] \tag{3}$$

and for each observed normalized price np_j we compute $\vec{\eta}(\mathrm{np}_j)$ which is $\vec{\eta}$ evaluated at the np_j price. We cluster these collections of vectors using k-means. The center of each cluster corresponds to regime R_k for $k = 1, \cdots, M$, where M is the number of regimes.

We distinguish three regimes, namely over-supply (R_1), balanced (R_2), and scarcity (R_3). Regime R_1 represents a situation where there is a glut in the market, i.e. an over-supply situation, which depresses prices. Regime R_2 represents a balanced market situation, where most of the demand is satisfied. Regime R_3 represents a situation where there is scarcity of products in the market, which increases prices. We can rewrite $p(\mathrm{np}|c_i)$ in a form that shows the dependence of the normalized price np not on the Gaussian component c_i of the GMM, but on the regime R_k:

[1]We focus on market prices because basic economic theory tells us that price is a clear indication of balance between supply and demand.

$$P(\mathrm{np}|R_k) = \sum_{i=1}^{N} p(\mathrm{np}|c_i)\, P(c_i|R_k). \tag{4}$$

The probability of regime R_k dependent on the normalized price np can then be computed using Bayes' rule as:

$$P(R_k|\mathrm{np}) = \frac{P(\mathrm{np}|R_k)\, P(R_k)}{\sum_{k=1}^{M} P(\mathrm{np}|R_k)\, P(R_k)} \quad \forall k = 1, \cdots, M. \tag{5}$$

The prior probabilities $P(R_k)$ of the different regimes are determined by a counting process over multiple games. Figure 3 depicts the regime probabilities for the medium market segment. Each regime is clearly dominant over a range of normalized prices. To make things more intuitive, we label regime R_1 as O for over-supply, regime R_2 as B for balanced, and regime R_3 as S for scarcity. The relative dominance and range of the different regimes varies among the market segments, but we can see, as expected, that oversupply corresponds to lower prices, a balanced situation to prices closer to the average, and scarcity to high prices.

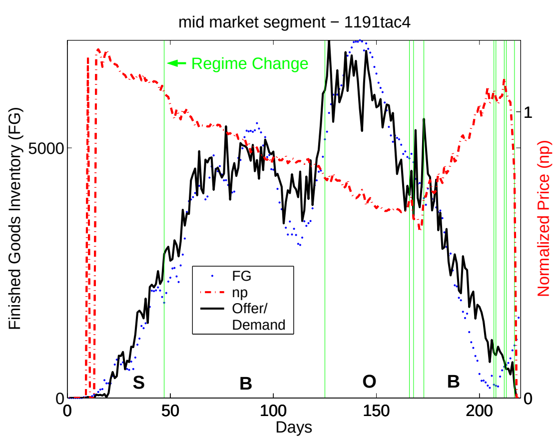

Figure 4: Relationships in the medium market between regimes and normalized prices.

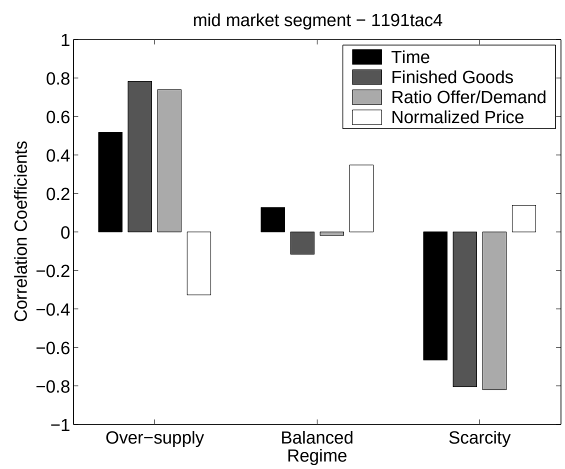

Figure 5: Correlation coefficients in the medium market.

Figure 4 shows the quantity of the finished goods inventory (FG), the ratio of offer to demand, and the normalized price over time in a game. The ratio of offer to demand represents the proportion of the market that is satisfied. While these factors are not directly visible to the agent during the game, they are clearly correlated with market regimes that are computed from observable factors, i.e. the mid-range price, which is the arithmetic mean of the minimum and maximum price which can be observed every day. Figure 4 shows a clear correlation between regimes and market parameters, substantiating our economic regime characterization. For example, the figure shows that when the offer to demand ratio is high (i.e. over-supply) prices are low and vice-versa. A correlation analysis of the market parameters is shown in Figure 5. The p-values for the correlation analysis are all less than 0.01. Regime R_1 (over-supply) correlates strongly and positively with time, ratio of offer to demand, and quantity of finished goods inventory, and negatively with normalized price. On the other hand, in Regime R_3 (scarcity) we observe a strong negative correlation with the market parameters shown in Figure 5.

3.1 Identification of current regime during the game

During the game, the agent can estimate the current regime by calculating the mean normalized price $\overline{\mathrm{np}}_{day}$ for a given day and by selecting the regime which has the highest probability, i.e. $\mathrm{argmax}_{1 \le k \le M}\, \vec{P}(R_k|\overline{\mathrm{np}}_{day})$. In TAC SCM, an agent has only limited information during the game. The best it can do is to estimate the mean of the normalized price of the computers sold the previous day using the daily price reports of minimum and maximum prices. This estimate can then be used to identify the corresponding regime online.

4 Regime Prediction

The behavior of an agent should depend on the current market regime as well as expectation of future regimes. We model the regime estimation process as a Markov process. We construct a Markov transition matrix, $\mathbf{T}_{\text{predict}}(r_{t+1}|r_t)$ off-line by a counting process over past games. This matrix represents the posterior probability of transitioning to a regime given the current regime. Equation 6 describes a recursive computation for predicting the posterior distribution of regimes at time $t+n$ days into the future, where $k = n+1$. In any given day we have access to yesterday's normalized prices (np).

$$\vec{P}(r_{t+k}|\mathrm{np}_{t-1}) = \sum_{r_{t+k-1}} \cdots \sum_{r_{t-1}} \vec{P}(r_{t-1}|\mathrm{np}_{t-1}) \cdot \prod_{j=1}^{k} \mathbf{T}_{\mathbf{predict}}(r_{t+j}|r_{t+j-1}) \tag{6}$$

Once the prediction of a regime change has high enough probability the agent should adjust its operational and strategic behavior in anticipation of the change.

We measure the accuracy of regime prediction using two separate values, (1) a count of how many times the regime predicted is the correct one, (2) a count of how many times the predicted time of the regime switch is correct. We assume the prediction is correct when the regime switch prediction is -2/+2 days from the one measured off-line.

We tested our approach on all 16 final games in the 2005 TAC SCM tournament. Starting with day 1 until day 199, we forecasted every day the regimes for the next 20 days and we forecast when the transition would occur. Using the Markov prediction, we identified the correct regime 3131 times out of 3184 trials (i.e. 98.34%) in the high market, 2722 times out of 3184 trials (i.e. 85.49%) in the medium market, and 2358 times out of 3184 trials (i.e. 74.06%) in the low market. The second measure of success, correctness of prediction of the time of regime change, in the high market ended up with 2821 out of 3184 trials (i.e. 88.60% success rate), in the medium market with 1976 out of 3184 trials (i.e. 62.06% success rate), and in the low market up with 1788 out of 3184 (i.e. 56.16% success rate).

We hypothesized that regime switches are not exponential (Markov), i.e. the future depends not only on the present state, but also on the length of time the process has spent in that state. This requires modeling the regime transition as a semi-Markov process [3].

To model this we modify the Markov transition matrix, $\mathbf{T}_{\text{predict}}$, to be a weighted sum of two matrices, the steady state matrix $\mathbf{T}_{\text{steady}}$ and the change matrix $\mathbf{T}_{\text{change}}$. $\mathbf{T}_{\text{steady}}$ is the $M \times M$ identity matrix, where M is the number of regimes. $\mathbf{T}_{\text{change}}$ is the Markov transition matrix, which is computed off-line as described earlier.

$$\mathbf{T}_{\text{predict}}(r_{t+1}|r_t) = (1 - \omega(.))\mathbf{T}_{\text{steady}} + \omega(.)\mathbf{T}_{\text{change}}(r_{t+1}|r_t) \tag{7}$$

where $\omega(.)$ represents the probability of a regime change, and r_t represents the current regime. To compute the value of $\omega(.)$, we need to introduce a few variables. We define Δt as the time since the last regime transition at t_0: $\Delta t = t - t_0$. We model the time τ_i spent in regime R_i before the transition to regime R_j occurs as a random variable with distribution F_{ij}. τ_i is estimated from historical data. We hypothesized that the probability density of τ_i is dependent on the current regime, R_i, i.e. $p(\tau_i|R_i)$. We computed the frequency of all values of τ_i in ascending order and fitted different distributions. The Gamma distribution, $g(t;\alpha,\lambda)$ is a reasonable fit to the data.

The probability of a regime transition $\omega(r,\Delta t)$ from the current regime, r, with respect to the time Δt that has elapsed since the last regime transition, t_0, is given by:

$$\omega(r = R_i, \Delta t) = \int_0^{\Delta t} p(\Delta t|r = R_i)\, \mathrm{d}\Delta t \tag{8}$$

where $p(\Delta t|r = R_i) = g(\Delta t; \alpha_i, \lambda_i)$. Equation 9 describes a recursive computation for predicting the posterior distribution of regimes at time $t+n$ days into the future, where $k = n+1$, for the semi-Markov process.

$$\vec{P}(r_{t+k}|\mathrm{np}_{t-1}) = \sum_{r_{t+k-1}} \cdots \sum_{r_{t-1}} \vec{P}(r_{t-1}|\mathrm{np}_{t-1}) \cdot \prod_{j=1}^{k} \mathbf{T_{predict}}(r_{t+j}|r_{t+j-1}, \Delta t + j - 1) \tag{9}$$

When we model the process as a semi-Markov process we obtain in the high market 2989 out of 3184 (i.e. 93.88% success rate), in the mid market 2395 out of 3184 trials (i.e. 75.22% success rate), and in the low market 2451 out of 3184 trials (i.e. 76.98% success rate).

5 Conclusions and Future Work

We presented an approach to characterizing and predicting economic market conditions. Our approach recognizes that market situations evolve over time and exhibit qualitative differences that can be used to guide strategic and tactical decisions. Our approach recognizes that prices are influenced by non-observable factors, such as the inventory positions of the other agents. Unlike price-following methods, our approach promises to enable an agent to anticipate and prepare for regime changes, for example by building up inventory in anticipation of better prices in the future or by selling in anticipation of an upcoming oversupply situation. Our next step is to integrate regime prediction with the ordering, inventory management and sales strategies of the agent.

References

[1] A. P. Dempster, N. M. Laird, and D. B. Rubin. Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society*, Series B, 39(1):1–38, 1977.

[2] W. Ketter, E. Kryzhnyaya, S. Damer, C. McMillen, A. Agovic, J. Collins, and M. Gini. MinneTAC sales strategies for supply chain TAC. In *Proc. of the Third Int'l Conf. on Autonomous Agents and Multi-Agent Systems*, pages 1372–1373, New York, July 2004.

[3] S. E. Levinson. Continuously variable duration hidden markov models for automatic speech recognition. *Comput. Speech Lang.*, 1(1):29–45, 1986.

[4] D. Pardoe and P. Stone. Bidding for Customer Orders in TAC SCM: A Learning Approach. In *Workshop on Trading Agent Design and Analysis*, pages 52–58, New York, July 2004.

[5] K. Pauwels and D. Hanssens. Windows of Change in Mature Markets. In *European Marketing Academy Conf.*, Braga, Portugal, May 2002.

[6] N. Sadeh, R. Arunachalam, J. Eriksson, N. Finne, and S. Janson. TAC-03: A supply-chain trading competition. *AI Magazine*, 24(1):92–94, 2003.