A Methodology for Identification and Prediction of Economic Regimes in Electronic Marketplaces

Wolfgang Ketter

Dept of Computer Science and Engineering

University of Minnesota

Machine Learning Group - April 14th, 2005

Thanks to: John Collins, Maria Gini, Alok Gupta and Paul Schrater.

Overview

- Agents, Markets, and Auctions.
- TAC SCM - Game Overview
- Problem
- Economic Regime Identification
- Economic Regime Prediction
- Conclusion
- Future Work

Agents, Markets, and Auctions

Our long term goal is to enable programs ("agents") to do transactions on electronic markets on behalf of a user.

Why electronic markets and auctions?

- Electronic markets have the potential for reduced costs and increased access to world-wide markets.
- Auctions are a general and proven way to negotiate among rational entities.

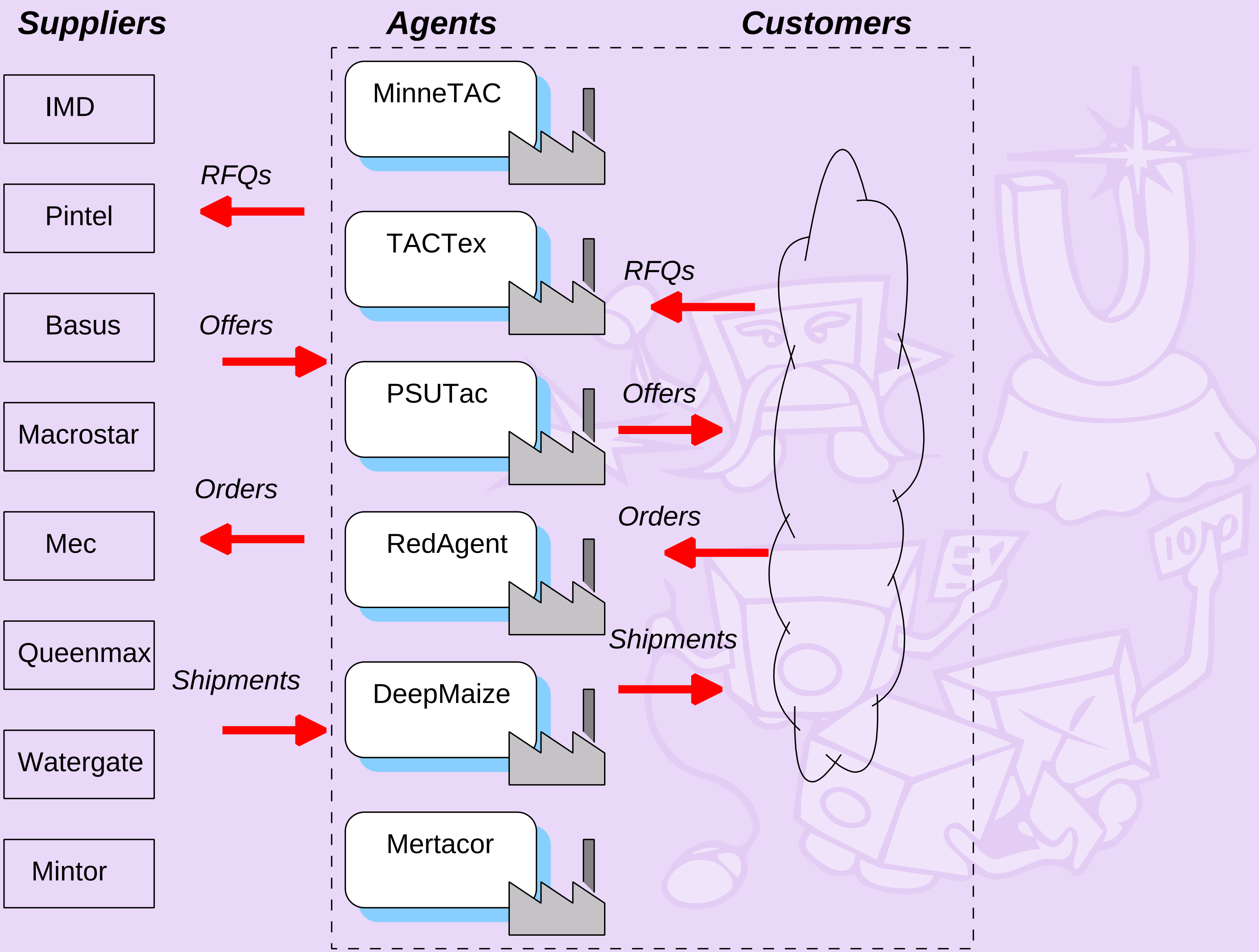
TAC SCM - Game Overview (1)

- Six autonomous agents:
  - Compete to acquire parts from suppliers.
  - Assemble parts into computers.
  - Compete for customer orders.
- The simulation takes place over 220 virtual days, each lasting fifteen seconds of real time.
- At the end (game/tournament), the agent with the most money in its bank account is the winner.

TAC SCM - Game Overview (2)

- Customers request computers of 16 different types to be delivered on a certain due date.
- Supply and demand varies randomly through the game and across three market segments (low, medium, and high computer price).
- Nominal price of a computer is the sum of the nominal cost of each part.

TAC SCM - Scenario



TAC SCM - Assembly Agents

These are the competitors in the game!

In this work we concentrate on bidding for customer orders, which includes:

- Which customer request for quotes (RFQs) to bid on. An RFQ includes: RFQid, computer type, due date, penalty, and reserve price.
- Product pricing.

Available Information in the Customer Market

Every day:

1. RFQs
2. Lowest and highest price paid per computer type from the previous day.

Every 20 days (market report):

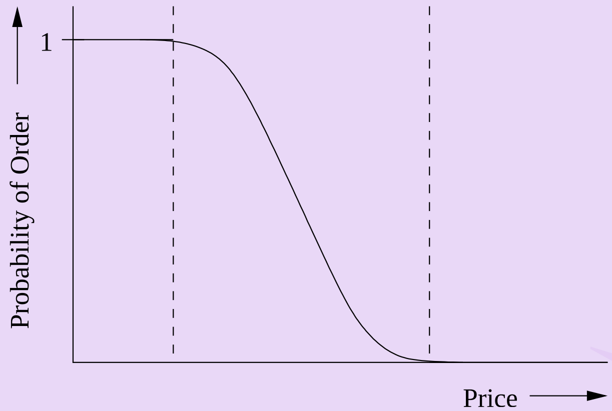
1. Total quantity per computer type ordered since the last market report.
2. Average order price per computer type ordered since the last market report.

Problem

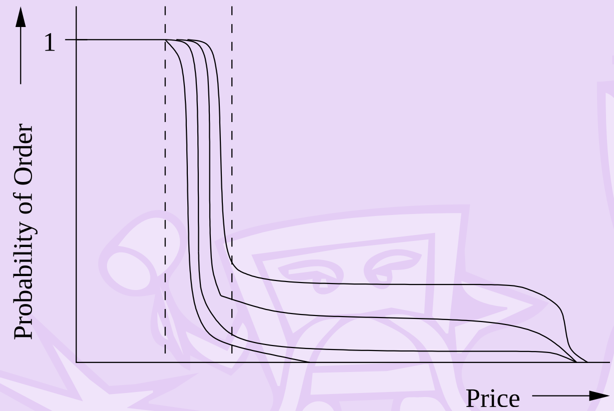
The TAC SCM game environment:

- Highly dynamic, since the demand and supply are changing constantly.
- Highly uncertain, because the agents change their behavior and adapt to new situations all the time.
- Very strategic, agents try to manipulate the market.
- Limited visibility, agents do not know their opponent's bids.

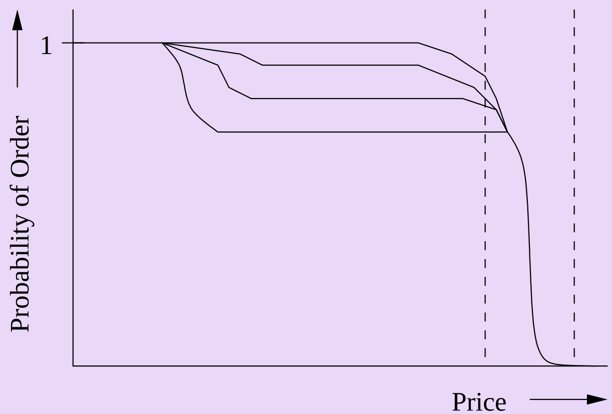
Economic Regimes



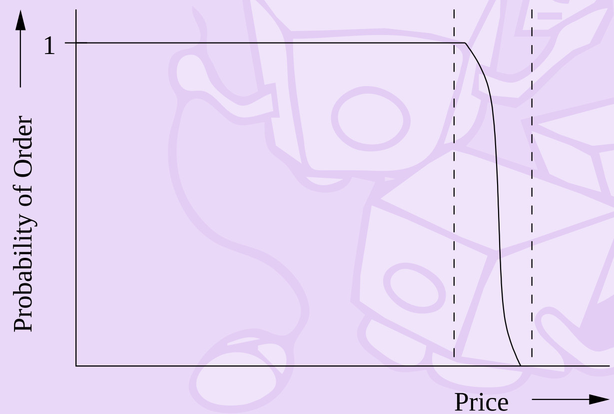
Balanced



Over-supply

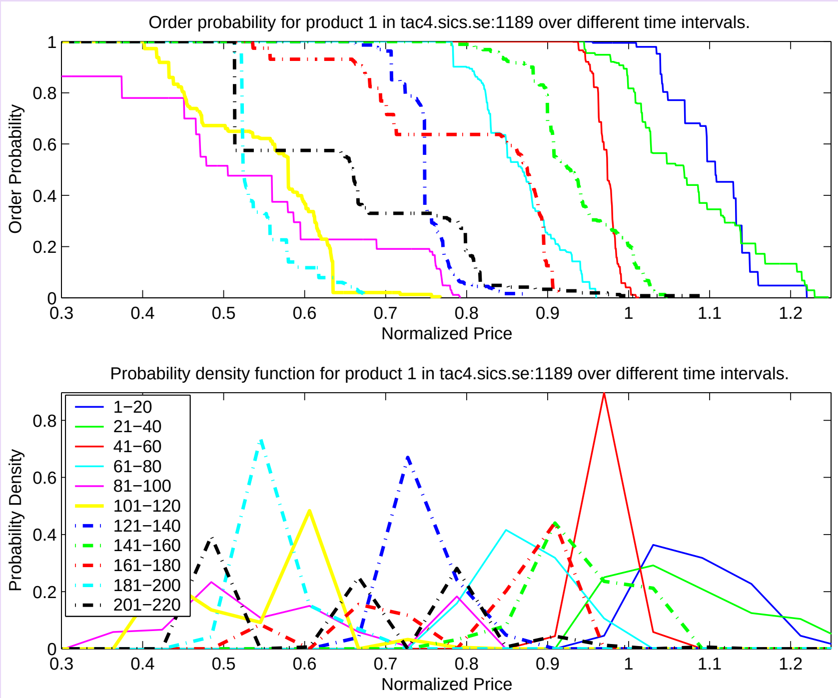


Scarcity



Extreme-scarcity

Probability of Order Example



1. Reverse cumulative density function represents the probability of order.
2. Computed from game data every 20 days.

Proposed Approach

1. Off-line identification of regimes from past game data.
2. Online identification of regimes from data available in the current game.
3. Prediction of regime transitions.

Off-line Regime Identification (1)

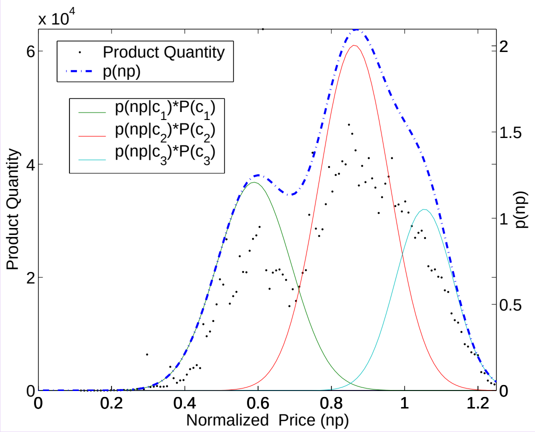
We use a Gaussian mixture model (GMM):

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

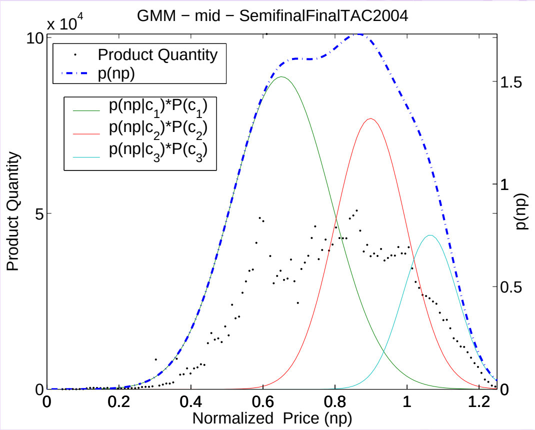
where $p(\mathrm{np}|c_i) = N[\mu_i, \sigma_i](\mathrm{np})$ is the i-th component of the normalized price density from the GMM, and $P(c_i)$ is the prior probability of the i-th component.

Off-line Regime Identification (2)

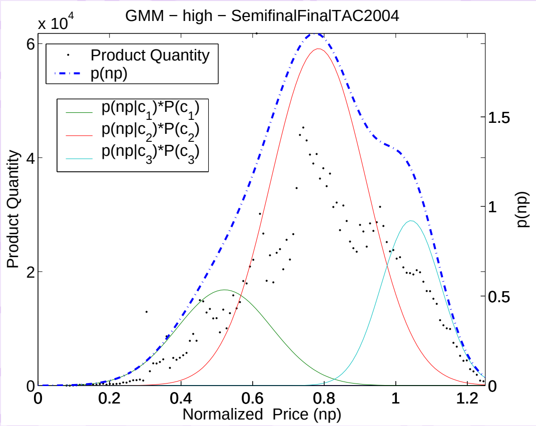
We apply the EM-Algorithm to determine the components (μ_i,σ_i, and $P(c_i)$) of the GMM, where $\forall i = 1, \cdots, N$. We assume $N = 3$.



Low Market



Medium Market



High Market

Off-line Regime Identification (3)

Using Bayes' rule we determine the posterior probability:

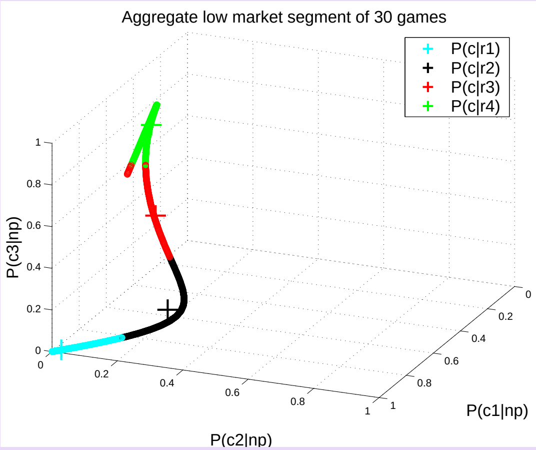
$$P(c_i|\mathrm{np}) = \frac{p(\mathrm{np}|c_i)\, P(c_i)}{\sum_{i=1}^{N} p(\mathrm{np}|c_i)\, P(c_i)} \quad \forall i = 1, \cdots, N$$

Off-line Regime Identification (4)

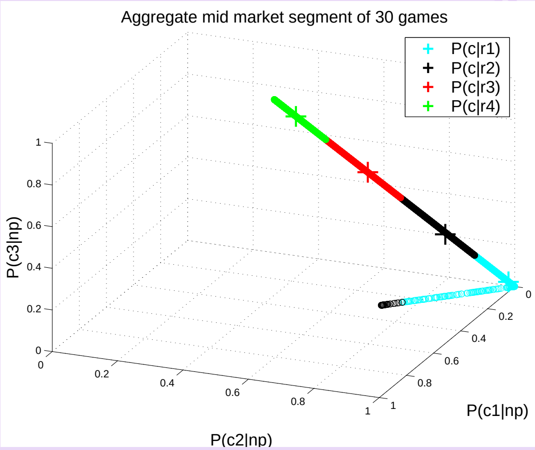
We defi ne the N-dimensional vector

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

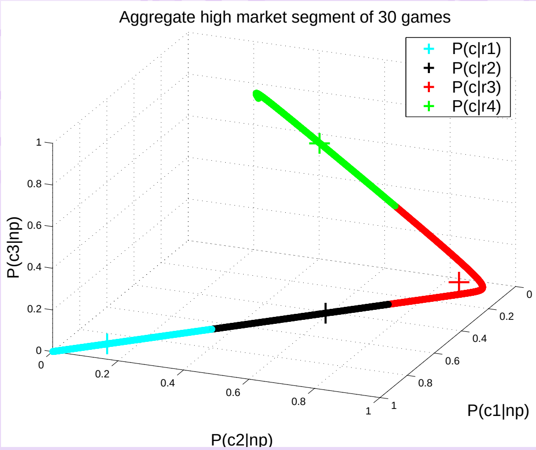
1. Compute $\vec{\eta}(\mathrm{np}_j)$ which is $\vec{\eta}$ evaluated at the np_j price.
2. Cluster these collections of vectors using k-means.
3. The center of each cluster corresponds to a regime R_k.



Low



Medium



High

Off-line Regime Identification (5)

Marginalizing over the components c_i we obtain:

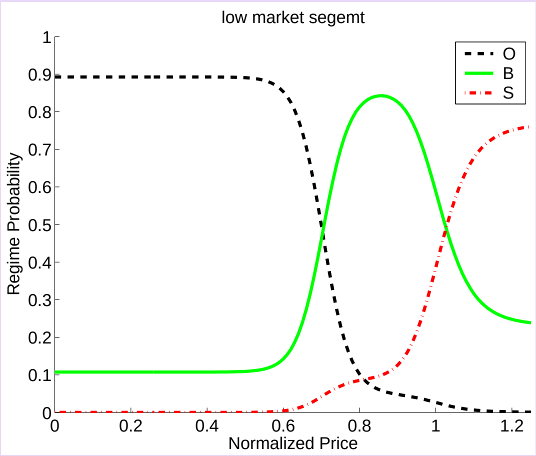
$$P(\mathrm{np}|R_k) = \sum_{i=1}^{N} p(\mathrm{np}|c_i)\, P(c_i|R_k)$$

Using Bayes' rule we determine the posterior probability:

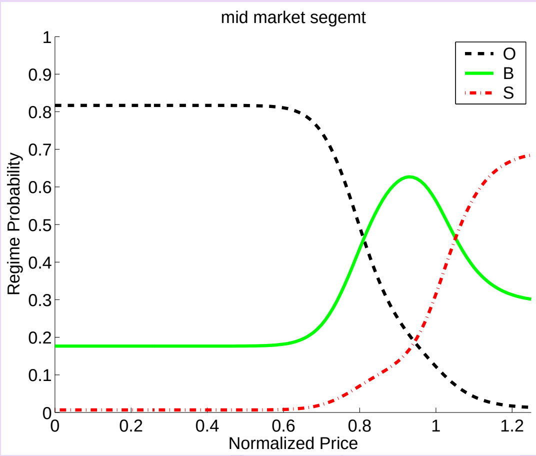
$$P(R_k|\mathrm{np}) = \frac{P(\mathrm{np}|R_k)\, P(R_k)}{\sum_{k=1}^{R} P(\mathrm{np}|R_k)\, P(R_k)} \quad \forall k = 1, \cdots, M$$

The prior probabilities $P(R_k)$ are determined by a counting process over a collection of entire games.

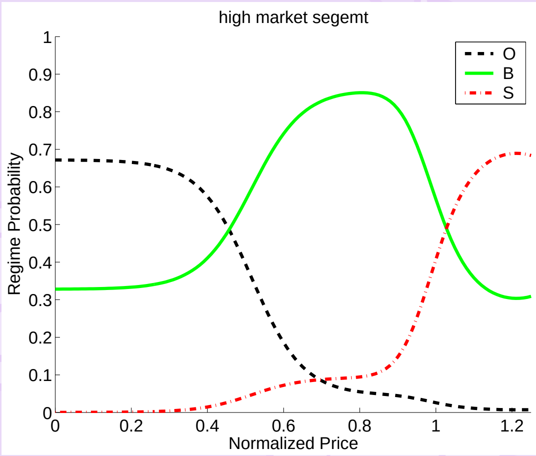
Regime Probability over Price for the 3 Market Segments



Low Market



Medium Market



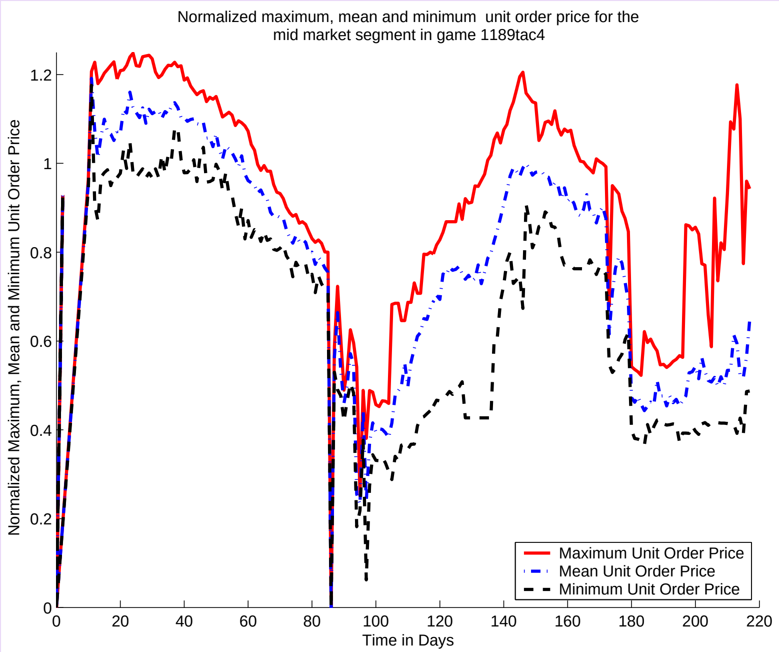
High Market

Regime probability over price - calculated off-line from 30 games.

Proposed Approach

1. Off-line identification of regimes from past game data.
2. Online identification of regimes from data available in the current game.
3. Prediction of regime transitions.

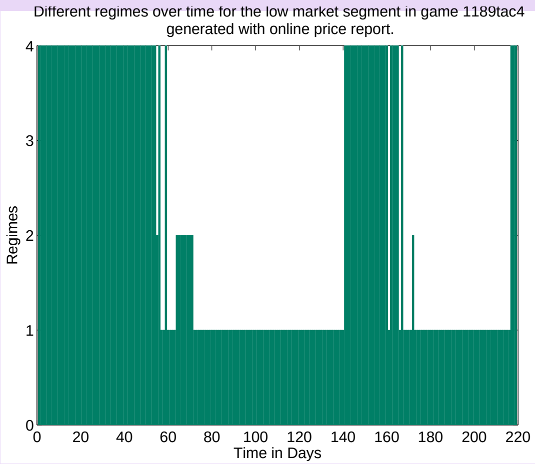
Online Identification of the Current Regime



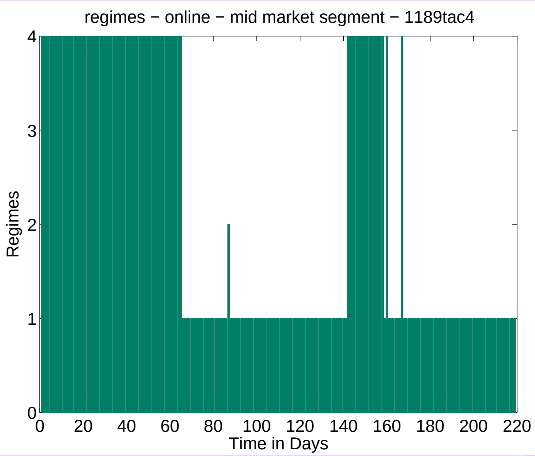
Daily price report: Minimum and maximum order prices.

1. Online every day we estimate the current regime by calculating the mean normalized price $\overline{\mathrm{np}}_{day}$ based on the daily price report.
2. Select regime which has the highest probability, i.e. $\mathrm{argmax}_{1 \leq k \leq M} \vec{P}(R_k|\overline{\mathrm{np}}_{day})$

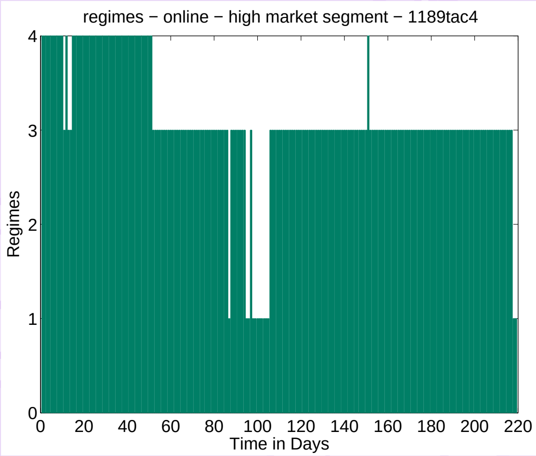
Regimes over time – online for the 3 Market Segments



Low Market



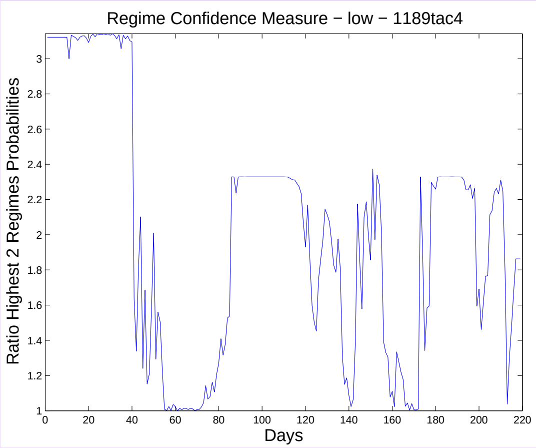
Medium Market



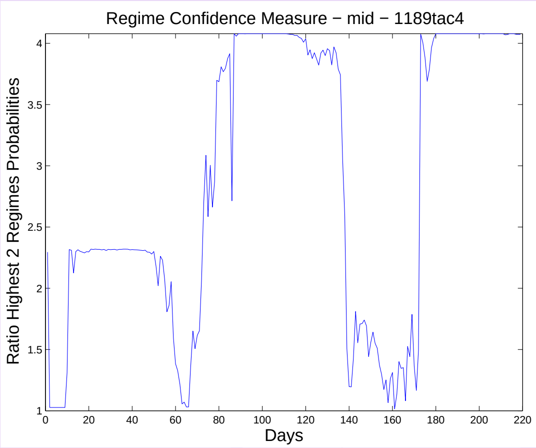
High Market

Game 1189@tac4 - Final TAC SCM’04: Regimes over time.

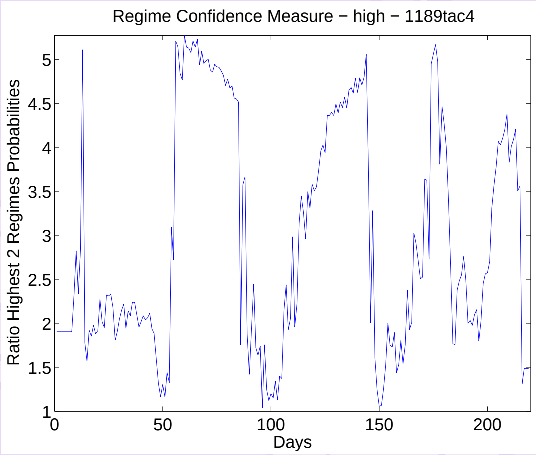
Regime Confidence Measure for the 3 Market segments



Low Market



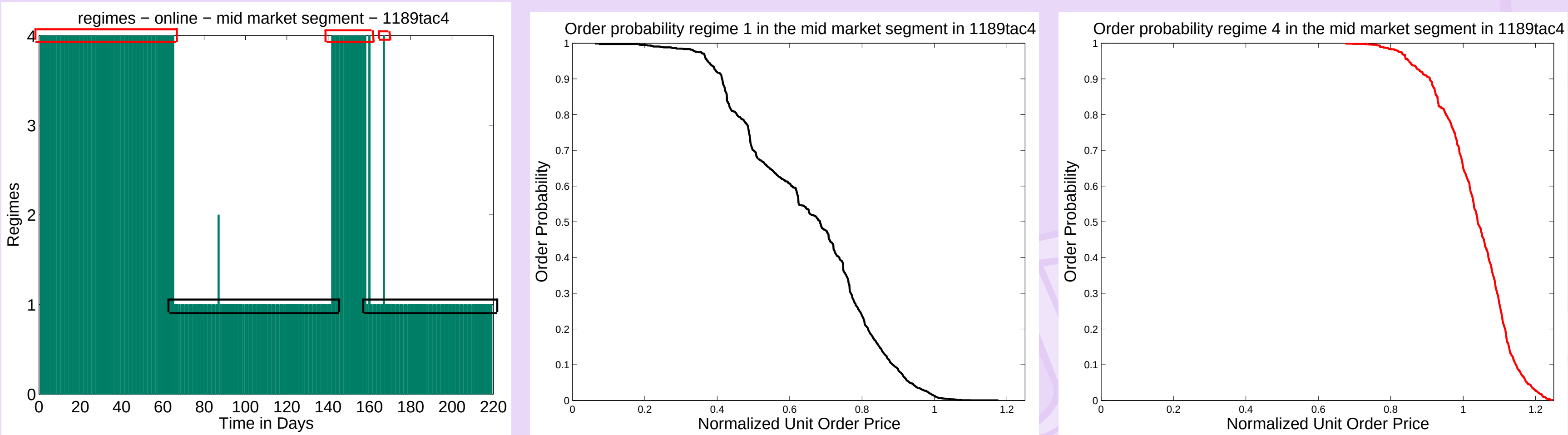
Medium Market



High Market

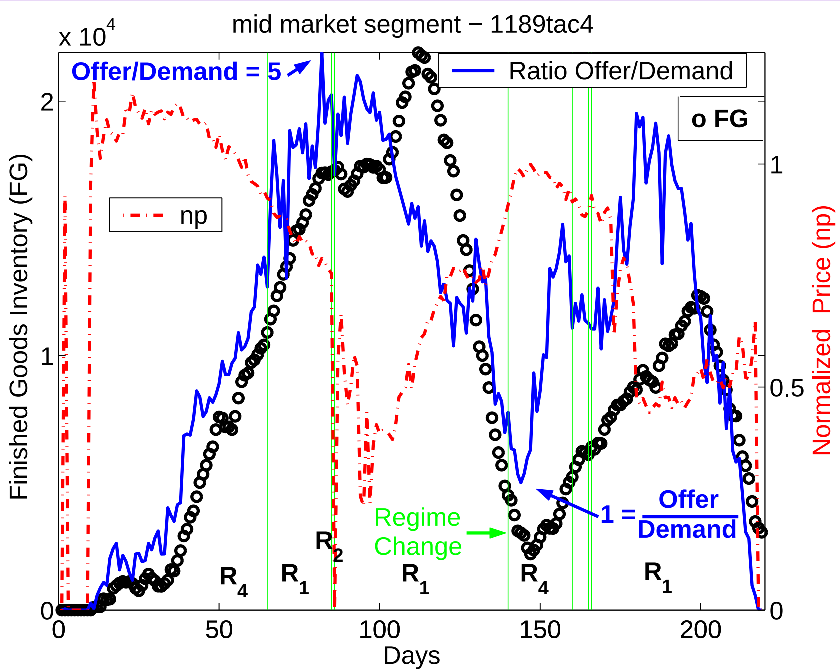
Game 1189@tac4: Ratio of highest and second highest regime probabilities.

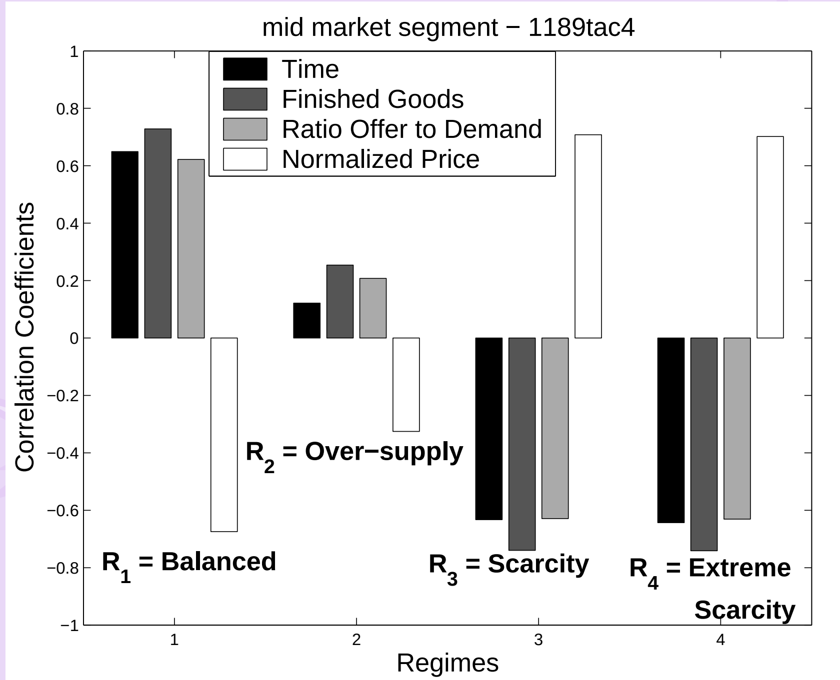
Regime 1 and 4 Order Probability Medium Market Segment



Game 1189@tac4: Regimes over time (left), order probability balanced regime R_1 (middle) and order probability scarcity regime R_4 (right).

Regime Market Parameters for the Medium Market Segment





Game 1189@tac4: Ratio $\frac{Offer}{Demand}$, and finished goods inventory (left) and correlation coefficients between regimes and market parameters (right).

Proposed Approach

1. Off-line identification of regimes from past game data.
2. Online identification of regimes from data available in the current game.
3. Prediction of regime transitions.

Prediction of Regimes

Agent behavior should depend on the current market regime and expectation of future regimes.

Distinguish between:

1. Short term prediction.
2. Long term prediction.

Short Term Prediction

- Uses the ratio of the probability of the highest vs the second highest regime to estimate the confidence in the regime identification.
  - High ratio: High confidence in regime identification.
  - Ratio close to 1: Mixture of almost equally likely regimes. This can be an indication of an upcoming regime change.
- Collect these ratios over an interval of days in the immediate past and we use interpolation to predict a regime shift.

Long Term Prediction

- Market exhibits time dependent behavior – long stable regimes and rapid switches to new regimes.
- Model the regime estimation process as a discrete, semi-Markov process:
  - It's a continuous time and a discrete space stochastic process.
  - Does not have the memoryless property of Markov processes.

Markov Transition Matrix

- Construct a Markov transition matrix, $\mathbf{T_{predict}}$, to model regime behavior:

$$\mathbf{T_{predict}}(r_{t+1}|r_t) \;=\; (1-\omega(.))\,\mathbf{T_{steady}}(r_{t+1}|r_t) + \omega(.)\,\mathbf{T_{change}}(r_{t+1}|r_t)$$

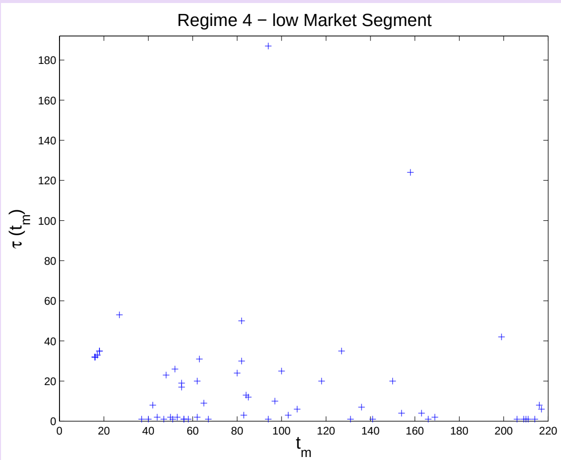
- $\mathbf{T_{steady}}$ – $M \times M$ identity matrix, where M is the number of regimes.
- $\mathbf{T_{change}}$ – posterior probability of transitioning to a regime given the current regime.
- $\omega(.)$ – probability of a regime change

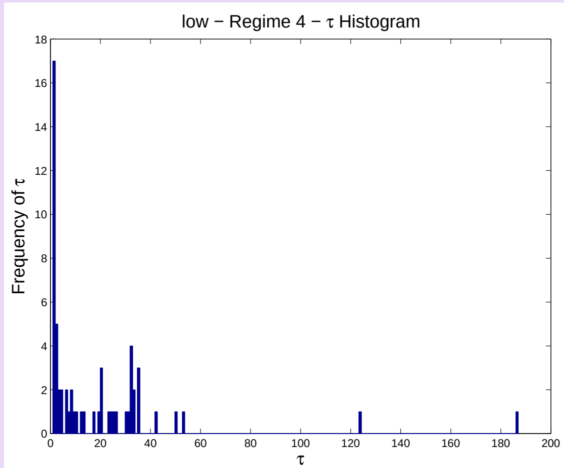
Off-line Regime Transition Probability (1)

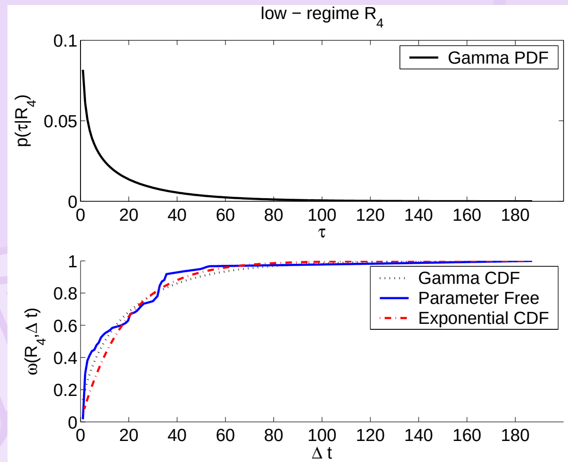
To compute $\omega(.)$ we model the distribution F_{ij} of how long (τ_i) the market stays in a regime as a random variable.

- Estimate the distribution F_{ij} of τ_i off-line.
- To model F_{ij} we hypothesized: $p(\tau_i|t, R_i)$
- Result: Hypothesis is not correct! F_{ij} – does not depend on time!
- Result: F_{ij} is dependent on R_i: $p(\tau_i|R_i)$

Off-line Regime Transition Probability (2)







Model $p(\tau_i|R_i)$ with the gamma density function:

$$g(t;\alpha,\lambda) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t} & \text{for } t \geq 0 \\ 0 & \text{for } t < 0 \end{cases},$$

where $\Gamma(x)$ is the gamma function.

Online Regime Transition Probability

Δt – time since the last regime transition at t_0:

$$\Delta t = t - t_0$$

Probability of a regime transition:

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

where $p(\Delta t | r = R_i) = g(\Delta t; \alpha_i, \lambda_i)$, and the parameters α_i, λ_i are fitted separately for each regime.

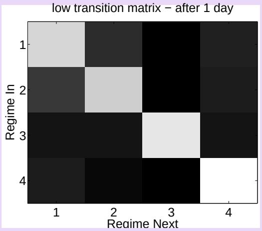
Online Prediction of Regimes (1)

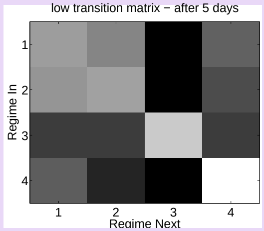
Determine the current transition matrix $\mathbf{T}_{\mathbf{predict}}$:

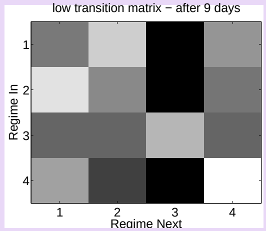
$$\mathbf{T}_{\mathbf{predict}}(r_{t+1}|r_t, \Delta t) = (\mathbf{I} - \mathbf{A})\,\mathbf{T}_{\mathbf{steady}}(r_{t+1}|r_t) + \mathbf{A}\,\mathbf{T}_{\mathbf{change}}(r_{t+1}|r_t)$$

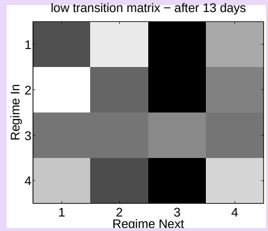
- $\mathbf{A}$ is a diagonal matrix with the probabilities $\omega(r = R_i, \Delta t)$ on the diagonal.
- $\mathbf{I}$ is an $M \times M$ identity matrix.

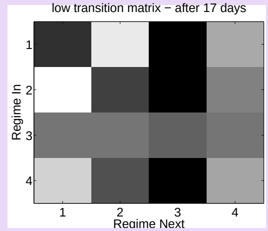
Online Prediction of Regimes (2) for the Low Market Segment

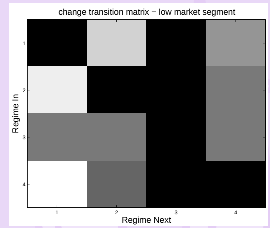












- Probabilities are shown as shades of gray, white = 1 and black = 0.
- Rows: Active regime - Columns: Next regime.
- $\mathbf{T}_{\mathbf{predict}}$ is evaluated for 1, 5, 9, 13 and 17 days from left to right.
- The rightmost matrix represents $\mathbf{T}_{\mathbf{change}}$.

Example Online Prediction of Current Regime (1)

Day = 1: Best prediction is the prior $\vec{P}(r_1)$.

Day = 2 and $\Delta t = 1$:

Step 1: Measurement $\vec{P}(\mathrm{np}_1|r_1)$ and correction:

$$\vec{P}(r_1|\mathrm{np}_1) = \frac{\vec{P}(\mathrm{np}_1|r_1)\,\vec{P}(r_1)}{\sum_{r_1}\vec{P}(\mathrm{np}_1|r_1)\,\vec{P}(r_1)}$$

Step 2: Prediction:

$$\vec{P}(r_2|\mathrm{np}_1) = \sum_{r_1}\mathbf{T_{predict}}(r_2|r_1,\Delta t = 1)\,\vec{P}(r_1|\mathrm{np}_1)$$

Example Online Prediction of Current Regime (2)

Day = 3 and $\Delta t = 2$:

Step 1: Measurement $\vec{P}(\mathrm{np}_2|r_2)$ and correction:

$$\vec{P}(r_2|\mathrm{np}_2, \mathrm{np}_1) = \frac{\vec{P}(\mathrm{np}_2|r_2)\,\vec{P}(r_2|\mathrm{np}_1)}{\sum_{r_2} \vec{P}(\mathrm{np}_2|r_2)\,\vec{P}(r_2|\mathrm{np}_1)}$$

Step 2: Prediction:

$$\vec{P}(r_3|\mathrm{np}_2, \mathrm{np}_1) = \sum_{r_2} \mathbf{T}_{\mathbf{predict}}(r_3|r_2, \Delta t = 2)\,\vec{P}(r_2|\mathrm{np}_2, \mathrm{np}_1)$$

Online Prediction Future Regimes

$$
\begin{aligned}
&\vec{P}(r_{t+k} | \mathrm{np}_{t-1}) \\
&\quad = \sum_{r_{t+k-1}} \cdots \sum_{r_{t-1}} \vec{P}(r_{t-1} | \mathrm{np}_{t-1}) \\
&\qquad\qquad \cdot \prod_{j=1}^{k} \mathbf{T}_{\mathbf{predict}}(r_{t+j} | r_{t+j-1}, \Delta t + j - 1)
\end{aligned}
$$

Future Work

- Use $\mathbf{T}_{\mathbf{predict}}$ to make long term predictions.
- Develop sales strategies that take advantage of regime forecasting.
- Integrate regime forecasting in the decision making process of the agent.

Conclusions

- Off-line identification of economic regimes from past game data.
- Online identification of economic regimes from data available in the current game.
- Prediction of economic regime transitions.

Contacts

email: ketter@cs.umn.edu
URL: www.cs.umn.edu/∼ketter