Automatic stock decision support system based on box theory and SVM algorithm

Qinghua Wen, Zehong Yang(楊澤紅),
Yixu Song(宋亦旭), Peifa Jia(賈培發)

State Key Laboratory on Intelligent Technology and Systems, Tsinghua National Laboratory for Information Science and Technology, Department of Computer Science and Technology, Tsinghua University, Bejing 100084, China

Expert Systems with Applications 37 (2010), pp. 1015-1022

Presenter: Hung-Hsi Chen
Date: Dec. 22, 2009

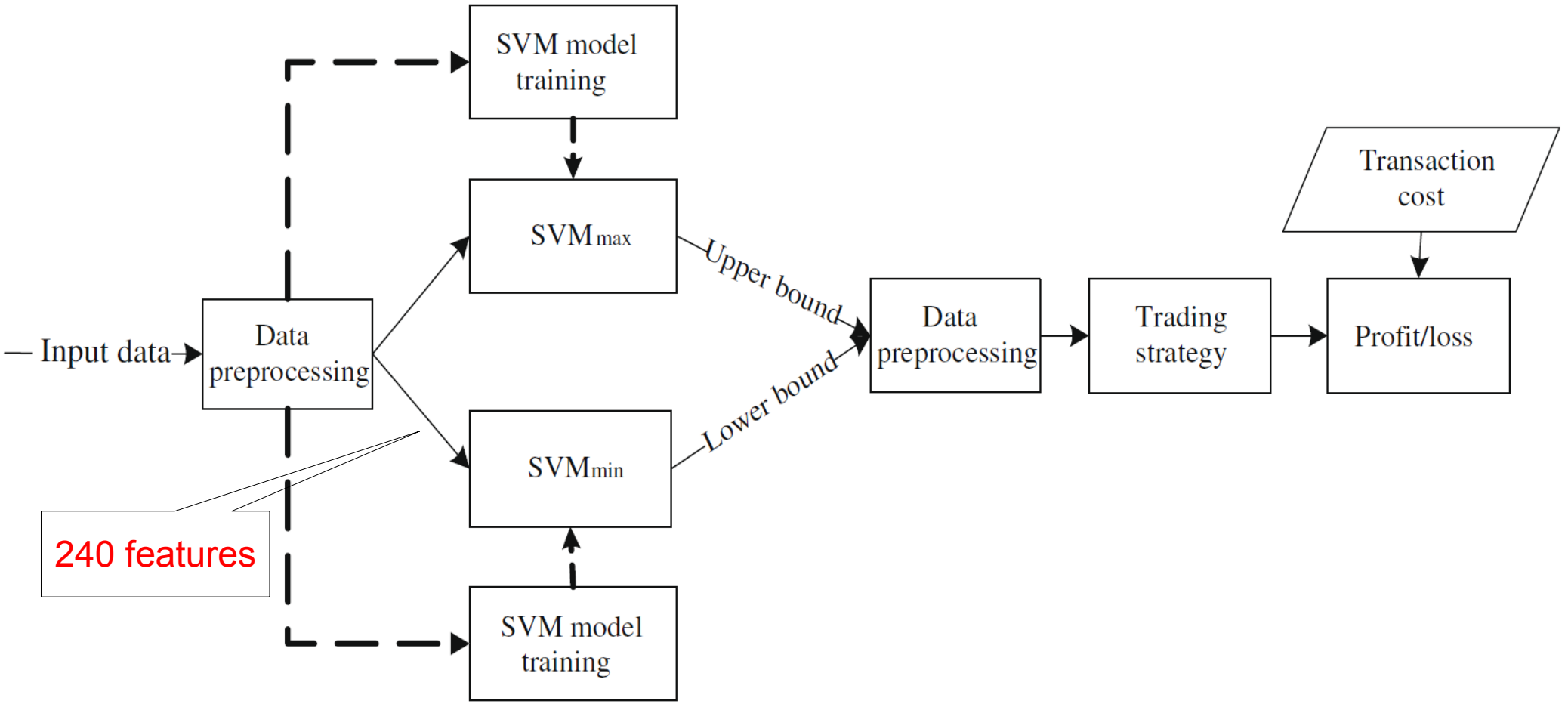
Abstract(1/2)

The stock market is considered as a high complex and dynamic system with noisy, non-stationary and chaotic data series. So it is widely acknowledged that stock price series modeling and forecasting is a challenging work. A significant amount of work has been done in this field, and in them, soft computing techniques have showed good performance. Generally most of these works can be divided into two categories. One is to predict the future trend or price; another is to construct decision support system which can give certain buy/sell signals. In this paper, we propose a new intelligent trading system based on oscillation box prediction by combining stock box theory and support vector machine algorithm.

Abstract(2/2)

The box theory believes a successful stock buying/selling generally occurs when the price effectively breaks out the original oscillation box into another new box. In the system, two SVM estimators are first utilized to make forecasts of the upper bound and lower bound of the price oscillation box. Then a trading strategy based on the two bound forecasts is constructed to make trading decisions. In the experiment, we test the system on different stock movement patterns, i.e. bull, bear and fluctuant market, and investigate the training of the system and the choice of the time span of the price box. The experiments on 442 S&P500 components show a promising performance is achieved and the system dramatically outperforms buy and-hold strategy.

Architecture



Support Vector Regression(SVR)

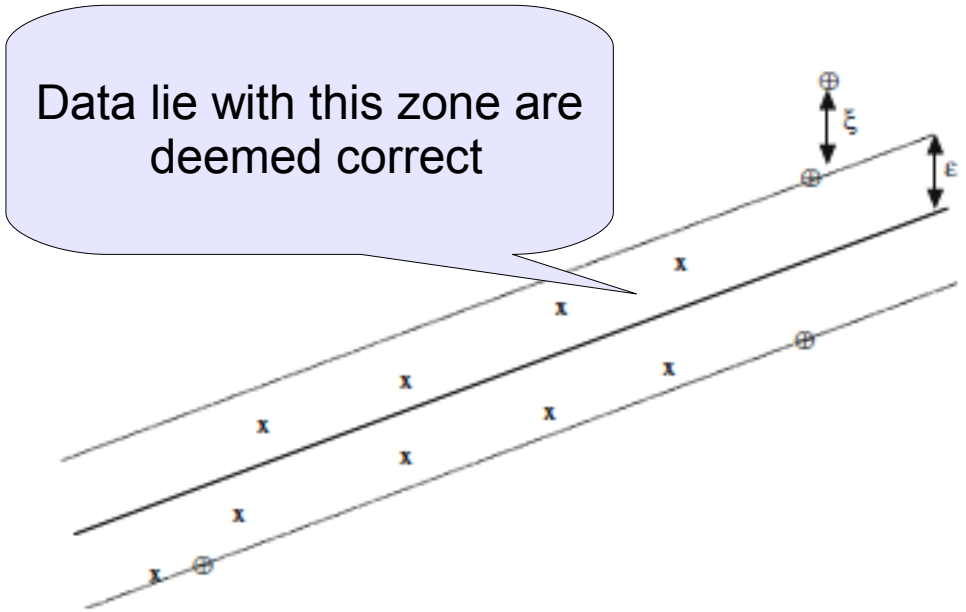
SVR is closely related to SVM classifiers in terms of theory and implementation.

The output is a real number(SVM: output is a type of class)

$$f(x) = w^T x + b \text{ with } w \in X, \quad b \in R$$

$$\min \quad \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{l}(\xi_i + \xi_i^*)$$

$$\text{s.t.} \quad \begin{cases} y_i - w^T x_i - b \leqslant \varepsilon + \xi_i \\ w^T x_i + b - y_i \leqslant \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geqslant 0 \end{cases}$$



Box theory

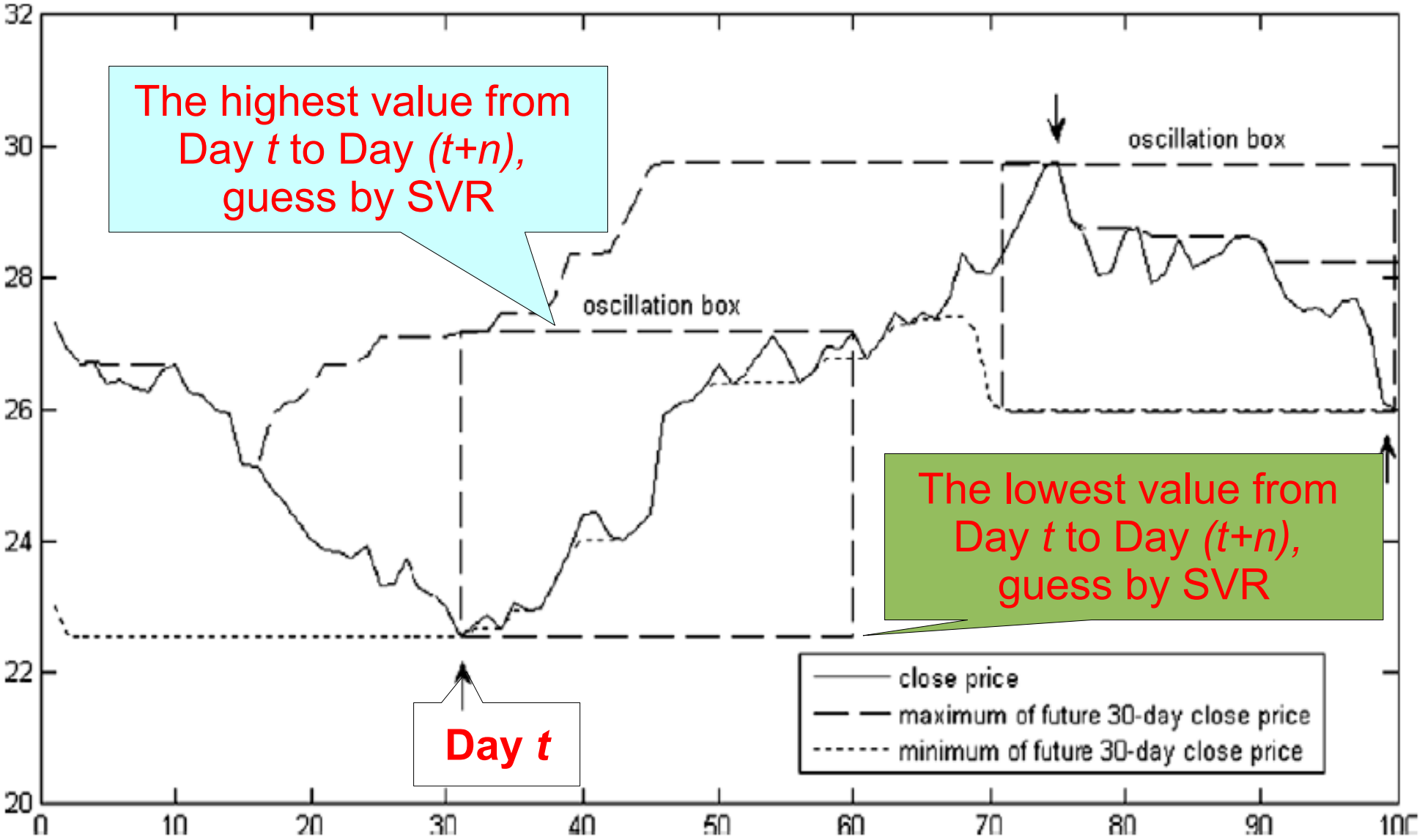


Fig. 1. '↑' means *buy*, '↓' means *sell*, or else do nothing. If the current close price is near to the bottom of the oscillation box and the bottom of the oscillation box is in uptrend, then buy. If the current close price is near to the top of the oscillation box and the top of the box is in downtrend, then sell. The time span n is set to 30.

Features

- 240 features
  1. Close Price(k)
  2. Moving Average(k)
  3. Relative Strength Index, RSI(k)
  4. Rate Of Change, ROC(k)
  5. FastK(k)
  6. SlowK(k)
  7. SlowD(k)
  8. $\tilde{H}$ (k-30), $\tilde{L}$ (k-30)
     - k=n, n-1, ...,(n-30)

All of the data is scaling to [-1, 1] by $x_{scaled} = -1 + 2\frac{x - min(x)}{max(x) - min(x)}$

Trading strategy

If next trade=buy

If $\frac{|C_i - \tilde{L}_i|}{C_i} \leqslant \sigma$ and $\tilde{L}_i$ is in uptrend

If $sellprice - C_i \geqslant \varphi$

then Buy, $buyprice = C_i$

next trade=sell

else If $\frac{|C_i - \tilde{H}_i|}{C_i} \leqslant \sigma$ and $\tilde{H}_i$ is in downtrend

If $C_i - buyprice \geqslant \phi$

then Sell, $sellprice = C_i$

next trade=buy

If $\frac{buyprice - C_i}{buyprice} \geqslant \theta$

then Sell, $sellprice = C_i$

next trade=buy

Parameter ：

σ ： transaction rate

Θ ： stop-loss rate

Φ, φ ： filter the false signal

收盤價非常靠近預測的最低價 (未來 n 日) 且預測的最低價之值在上升，即股市有反轉上升趨勢的可能性。

Trading strategy

If next trade=buy

If $\frac{|C_i - \tilde{L}_i|}{C_i} \leqslant \sigma$ and $\tilde{L}_i$ is in uptrend

If $sellprice - C_i \geqslant \varphi$

then Buy, $buyprice = C_i$

next trade=sell

else If $\frac{|C_i - \tilde{H}_i|}{C_i} \leqslant \sigma$ and $\tilde{H}_i$ is in downtrend

If $C_i - buyprice \geqslant \phi$

then Sell, $sellprice = C_i$

next trade=buy

If $\frac{buyprice - C_i}{buyprice} \geqslant \theta$

then Sell, $sellprice = C_i$

next trade=buy

Parameter ：

σ ： transaction rate

Θ ： stop-loss rate

Φ, φ ： filter the false signal

之前賣出的價錢與現在的收盤價已經有一段差距，避免震盪時的假訊號

Trading strategy

If next trade=buy

If $\frac{|C_i - \tilde{L}_i|}{C_i} \leqslant \sigma$ and $\tilde{L}_i$ is in uptrend

If $sellprice - C_i \geqslant \varphi$

then Buy, $buyprice = C_i$

next trade=sell

else If $\frac{|C_i - \tilde{H}_i|}{C_i} \leqslant \sigma$ and $\tilde{H}_i$ is in downtrend

If $C_i - buyprice \geqslant \phi$

then Sell, $sellprice = C_i$

next trade=buy

If $\frac{buyprice - C_i}{buyprice} \geqslant \theta$

then Sell, $sellprice = C_i$

next trade=buy

Parameter ：

σ ： transaction rate

Θ ： stop-loss rate

Φ, φ ： filter the false signal

之前買出的價錢與現在的收盤價已經有一段差距，避免震盪時的假訊號

Performance evaluation

$$MSE = \frac{1}{N}\sum_{i=1}^{N}(y_i - y_i^*)^2$$

y_i : actual output, y^*_i : the estimate

the smaller MSE, the better predictive performance

$$SCC = \frac{\left[\sum_{i=1}^{N}(y_i - \bar{y}_i)(y_i - \bar{y}_i^*)\right]^2}{\sum_{i=1}^{N}(y_i - \bar{y}_i)^2 \sum_{i=1}^{n}(y_i^* - \bar{y}_i^*)^2}$$

Squared correlation coefficient

the larger SCC, the better predictive performance

$$rate\ of\ profit = (Y - Y_0)/Y_0 \times 100\%$$

Experiment

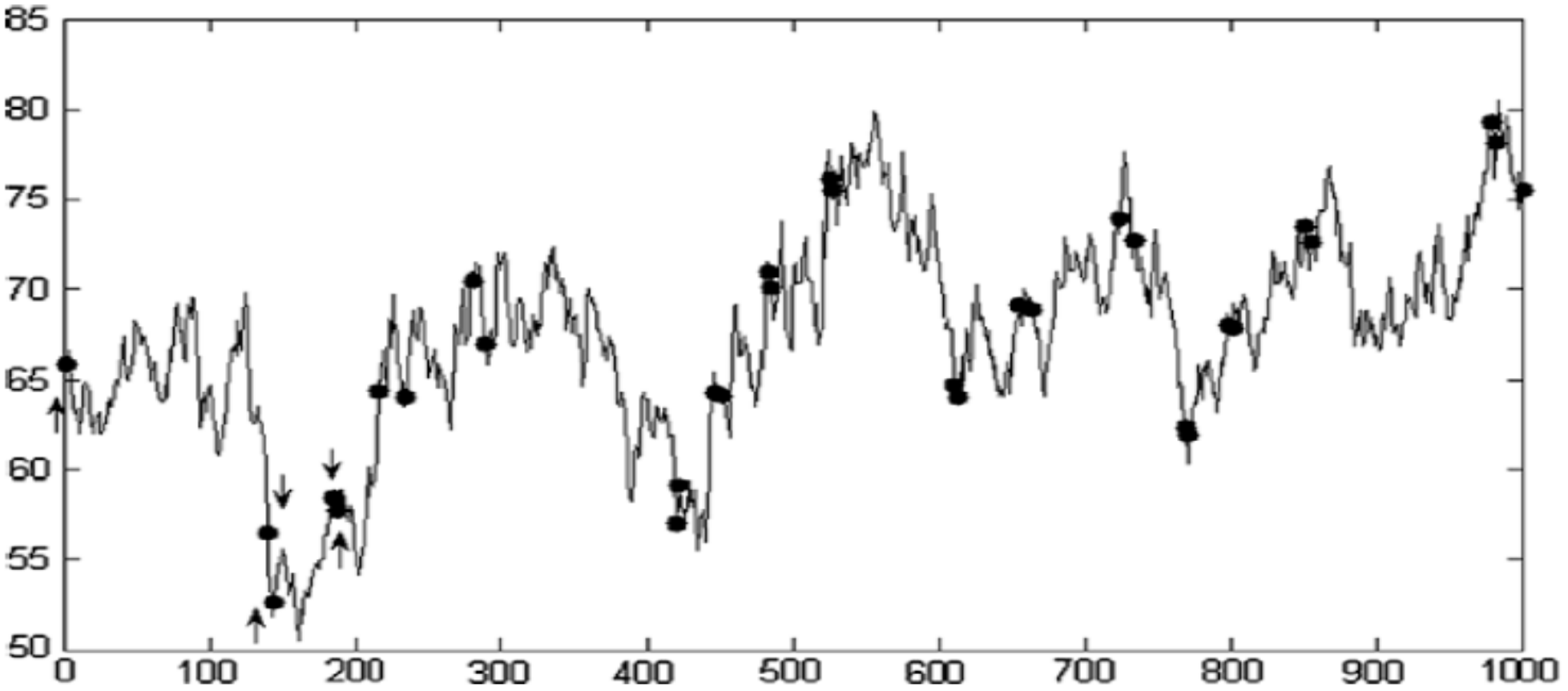


Fig. 3. Trading log('↑': buy, '↓': sell) of ALCOA INC for 4 years.

Parameter ：

Date: 3/27/1990~3/7/1994

Transaction rate σ : 0.05

Stop-loss rate Θ : 14%

Φ, φ : 0.5

Result ：

Our system: profit 121.63%

Market: profit 15.3%

MSE, SCC(SVMmax) : 0.00661, 0.7109

MSE, SCC(SVMmin): 0.00842, 0.6677

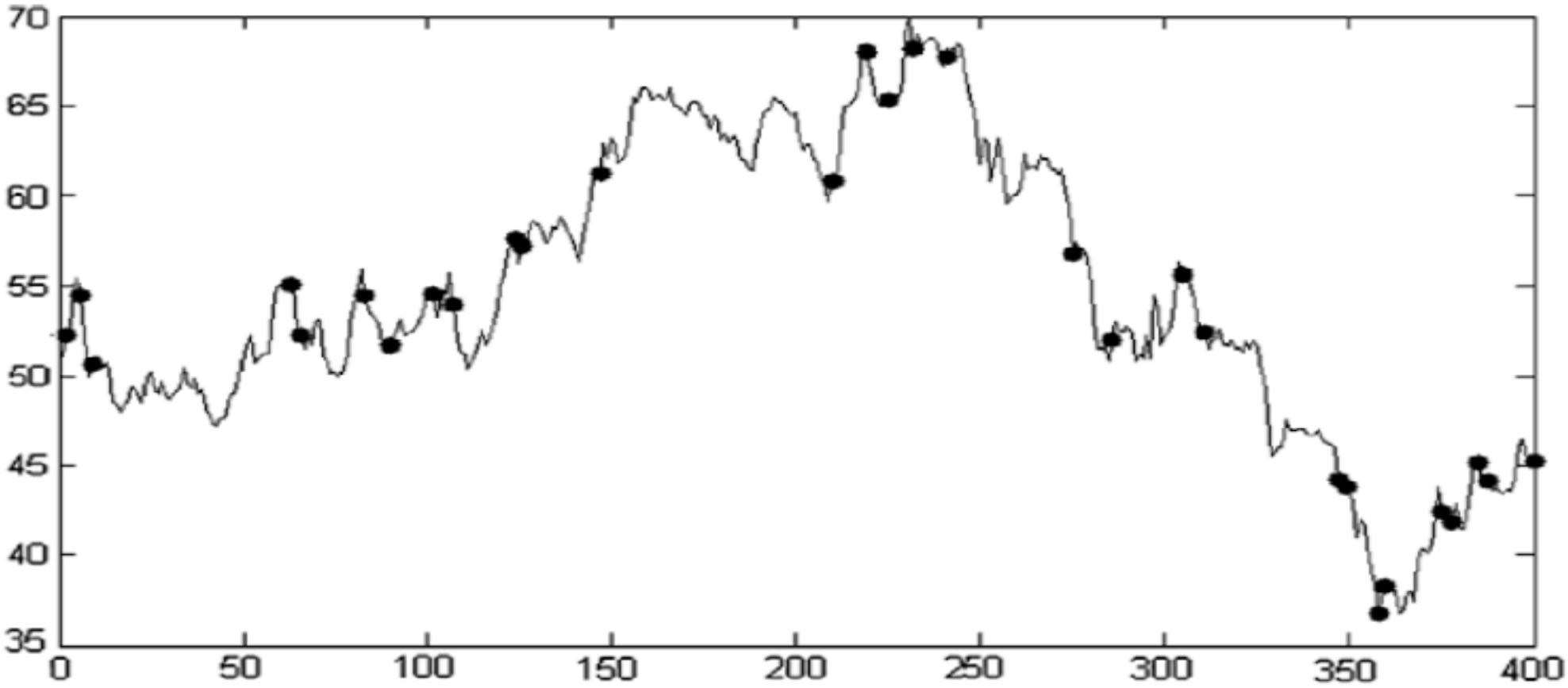


Fig. 4. Trading log of ALCOA INC for almost 2 years.

Parameter ：

Date: 6/13/1970~1/11/1972

Transaction rate σ : 0.12

Stop-loss rate Θ : 15%

Φ, φ : 0.2

Result ：

Our system: profit 21.04%

Market: loss 11.3%

MSE, SCC(SVMmax) : 0.00890, 0.8370

MSE, SCC(SVMmin): 0.00842, 0.6677

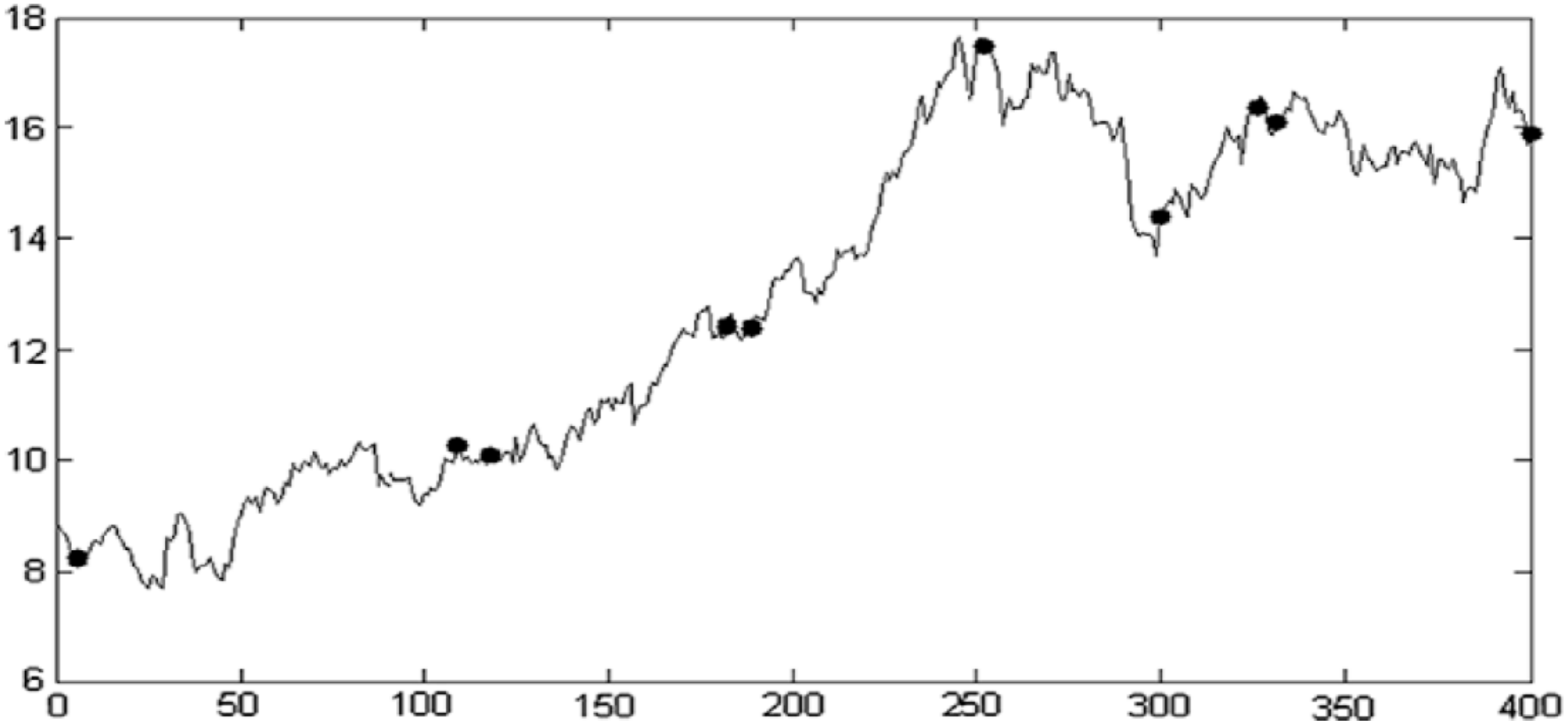


Fig. 5. Trading log of AES CP INC for 400 days.

Parameter ：

Date: 3/17/2004~10/17/2005
Transaction rate σ : 0.09
Stop-loss rate Θ : 15%
Φ, φ : 0.2

Result ：

Our system: profit 139.5%
Market: profit 80.6%
MSE, SCC(SVMmax) : 0.00662, 0.8806
MSE, SCC(SVMmin): 0.00776, 0.90364

Fig. 6. Trading log of TRIBUNE INC for 400 days.

Parameter ：

Date: 3/17/2004~10/17/2005
Transaction rate σ : 0.01
Stop-loss rate Θ : 15%
Φ, φ : 0.2

Result ：

Our system: loss 7.6%
Market: loss 35.2%
MSE, SCC(SVMmax) : 0.00157, 0.9921
MSE, SCC(SVMmin): 0.00218, 0.8336

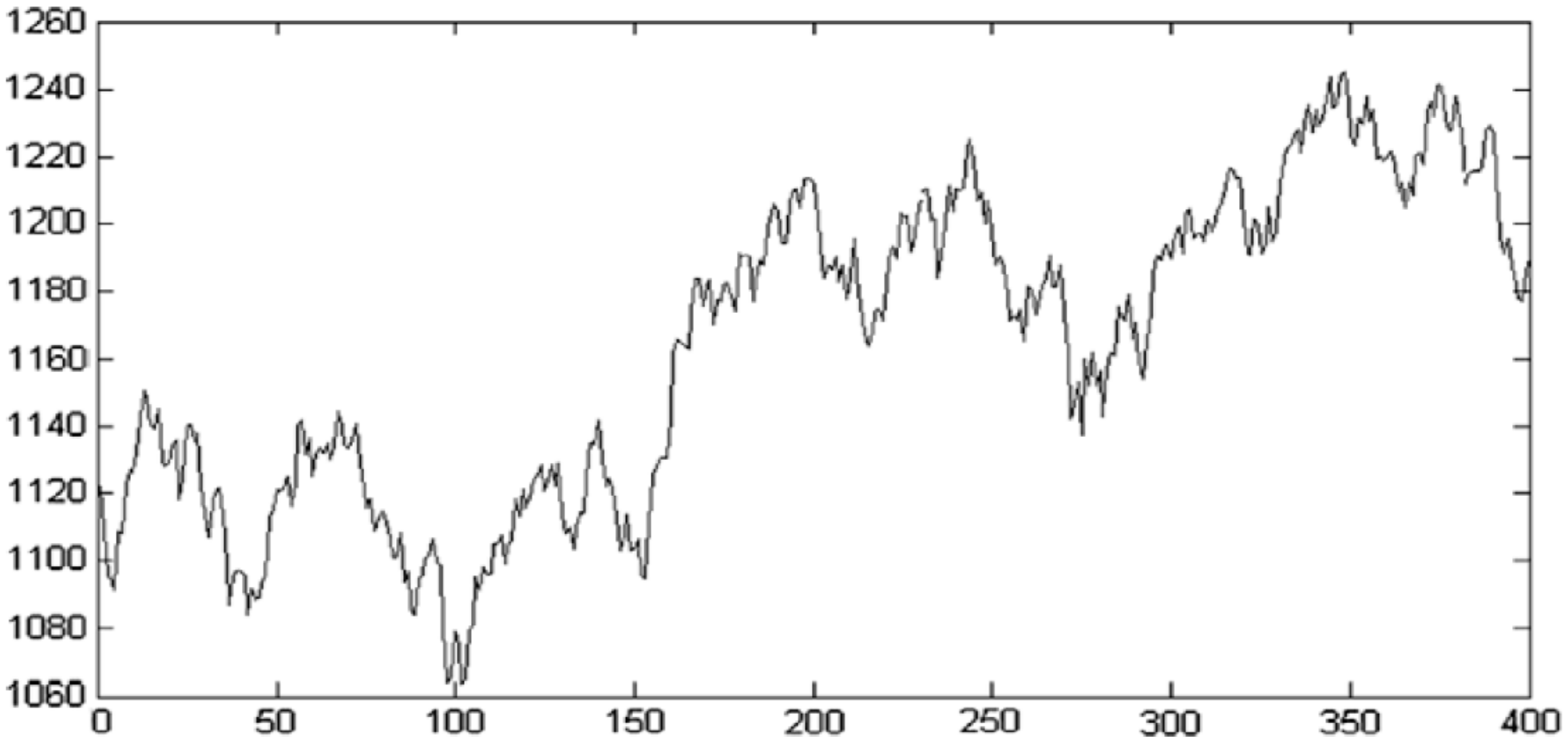


Fig. 7. S&P 500 index from Mar 17, 2004 to Oct 17, 2005.

SVM RBF kernel ：

SVMmax : c:550, g: 0.00002125

SVMmin: c:350, g: 0.00002125

parameter:

Transaction rate σ : 0.01

Stop-loss rate Θ : 15%

Φ, φ : 0.2

Result ：

Market profit: 6.57%

Choose 442 stocks for experiment

Average (our system)profit: 25.94%

Average Buy-and-hold profit: 5.77%

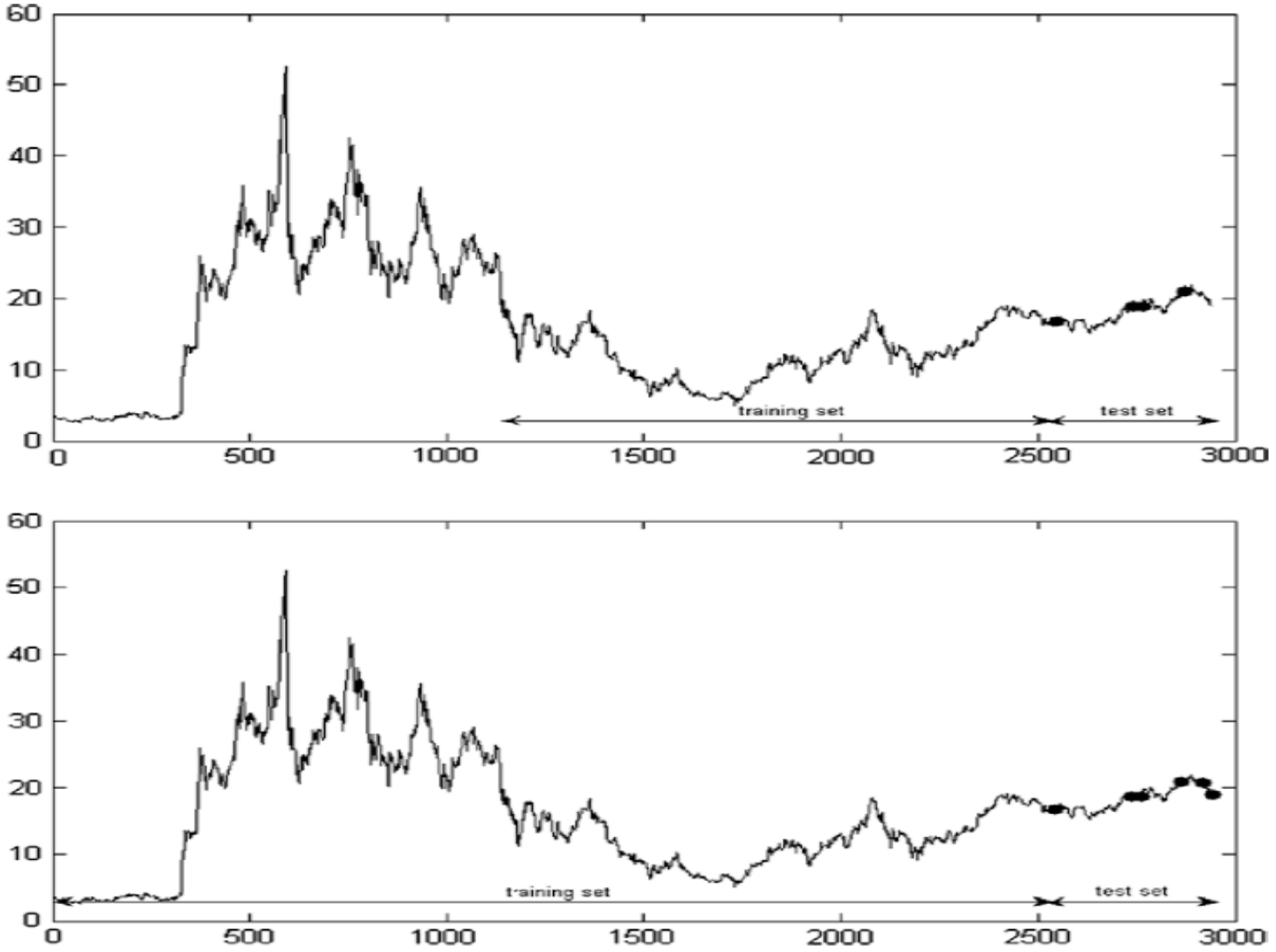


Fig. 8. Trading log of AES CP INC for 400 days from Mar 17, 2004 to Oct 17, 2005.with transaction rate σ set to 0.25. Top: There is 1400 data training points, the profit is 23.92%. Bottom: the training set extend to 2492 points, the profit is 14.02%.

The proportion of training set and testing sets is 3:1

Table 1
Average performance of trading 422 stocks for 400 days.

Market pattern	Number of stock	Less than buy-and-hold	Less (%)	Number of loss	Loss (%)	Average profit (%)	Average profit of buy-and-hold (%)
Bull	224	57	25.44	0	0	37.69	29.47
Bear	198	1	0.51	29	14.65%	12.65	−21.04
Total	422	58	13.12	29	6.87%	**25.94**	**5.77**

The proportion of training set and testing sets is 5:1 for the 57 stocks which underperform than buy & hold and 29 stocks which loss in downtrend

Table 2
Enhanced average performance of trading 422 stocks for 400 days.

Market pattern	Number of stock	Less than buy-and-hold	Less (%)	Number of loss	Loss (%)	Average profit (%)	Average profit of buy-and-hold (%)
Bull	224	39	17.41%	0	0	39.11	29.47
Bear	198	0	0	14	7.07%	13.79	−21.04
Total	422	39	9.24%	14	3.32	**27.23%**	**5.77**

The proportion of training set and testing sets is 5:1 for all 422 stocks

Table 3
Average performance of trading 422 stocks for 400 days with extended training set.

Market pattern	Number of stock	Less than buy-and-hold	Less (%)	Number of loss	Loss (%)	Average profit (%)	Average profit of buy-and-hold (%)
Bull	224	51	22.77	0	0	36.29	29.47
Bear	198	2	1.01	26	13.13%	10.33	−21.04
Total	422	53	12.56	26	8.53%	**24.11**	**5.77**

50 stocks(the alphabetically first 50 stocks in S&P500), */* means to predict the upper/lower bound of the oscillation box * days ahead

Table 4
Average performance of trading 50 stocks by varying *n*.

	3/3	5/5	8/8	10/10	15/15	20/20	30/30	40/40	50/50
Average profit (%)	35.5	39.49	**40.79**	39.83	34.35	35.86	**37.88**	37.01	34.89
number of transaction	9.8	8.7	8.1	7.7	7.3	8.4	8.8	7.9	7.7
Average SCC of *SVMmax*	0.969	0.945	0.914	0.888	0.843	0.792	0.706	0.623	0.562
Average SCC of *SVMmin*	0.959	0.925	0.881	0.852	0.786	0.732	0.609	0.523	0.462

Trading on Microsoft

SVM RBF kernel ：

SVMmax : c:550, g: 0.00002125

SVMmin: c:350, g: 0.00002125

Parameter ：

n: 30

Transaction rate σ : vary

Stop-loss rate Θ : 10%

Φ, φ : 0.2

Result ：

MSE, SCC(SVMmax) : 0.00890, 0.8370

MSE, SCC(SVMmin): 0.00842, 0.6677

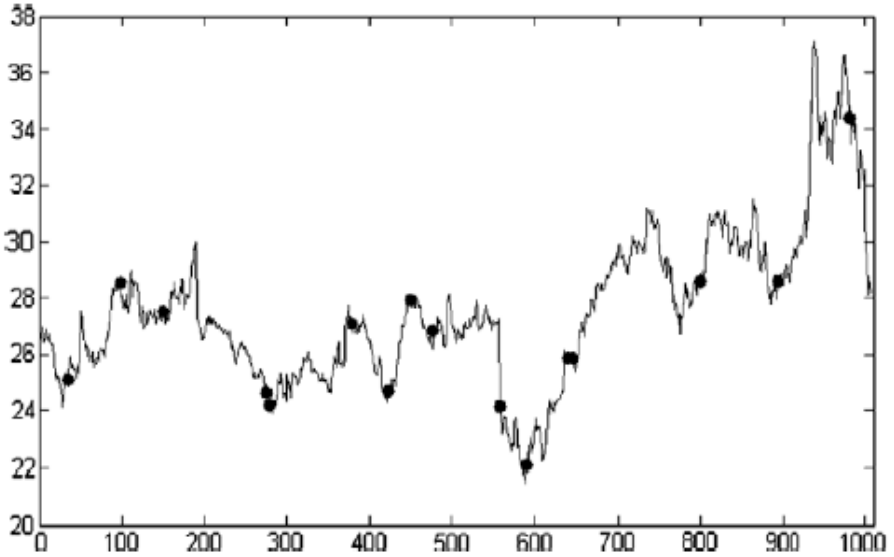


Fig. 9. Trading log of Microsoft for 4 years from Feb 12, 2004 to Feb 12, 2008 with trade ratio σ set to 0.025 and profit 74.00%. The profit is 5.16% by the buy-and-hold strategy.

Table 5
Performance of trading microsoft.

σ	0.005	0.01	0.015	0.02	0.025	0.03	0.035	0.04	Buy-and-hold
Number of transactions	4	5	7	8	8	8	6	6	1
Profit (%)	−2.77	19.03	53.37	73.94	**74.00**	72.17	29.37	15.41	5.16

Trading on IBM

SVM RBF kernel ：

SVMmax : c:1000, g: 0.00002125

SVMmin: c:2000, g: 0.00002125

Parameter ：

n: 30

Transaction rate σ : vary

Stop-loss rate Θ : 15%

Φ, φ : 0.2

Result ：

MSE, SCC(SVMmax) : 0.00016, 0.9111

MSE, SCC(SVMmin): 0.00020, 0.8706

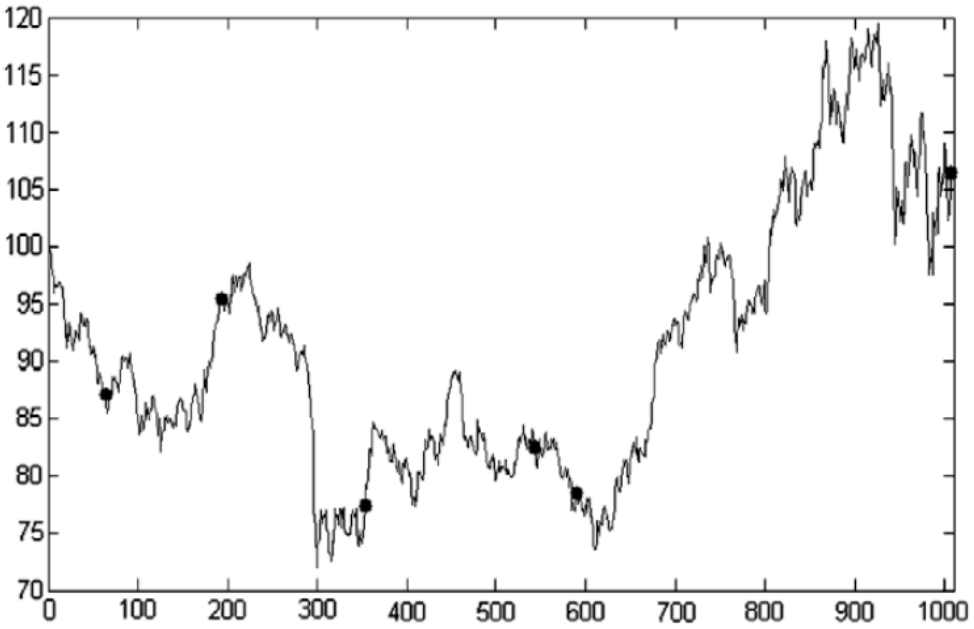


Fig. 10. Trading log of IBM for 4 years from Feb 12, 2004 to Feb 12, 2008 with trade ratio σ set to 0.015 and profit 53.31%. The profit is 7.28% by the buy-and-hold strategy.

Table 6
Performance of trading IBM.

σ	0.005	0.01	0.015	0.02	0.025	0.03	0.035	0.04	Buy-and-hold
Number of transactions	2	2	3	4	4	6	10	13	1
Profit (%)	10.10	47.38	**53.31**	20.83	24.33	5.97	15.33	−13.63	7.28

Thank you!!!