Stock Trading by Modelling Price Trend with Dynamic Bayesian Networks

Jangmin O[1], Jae Won Lee[2], Sung-Bae Park[1], and Byoung-Tak Zhang[1]

[1] School of Computer Science and Engineering, Seoul National University
San 56-1, Shillim-dong, Kwanak-gu, Seoul, Korea 151-744
{jmoh,sbpark,btzhang}@bi.snu.ac.kr
[2] School of Computer Science and Engineering, Sungshin Women's University,
Dongsun-dong, Sungbuk-gu, Seoul, Korea 136-742
jwlee@cs.sungshin.ac.kr

Abstract. We study a stock trading method based on dynamic bayesian networks to model the dynamics of the trend of stock prices. We design a three level hierarchical hidden Markov model (HHMM). There are five states describing the trend in first level. Second and third levels are abstract and concrete hidden Markov models to produce the observed patterns. To train the HHMM, we adapt a semi-supervised learning so that the trend states of first layer is manually labelled. The inferred probability distribution of first level are used as an indicator for the trading signal, which is more natural and reasonable than technical indicators. Experimental results on representative 20 companies of Korean stock market show that the proposed HHMM outperforms a technical indicator in trading performances.

1 Introduction

Stock market is a core of capitalism where people invest some of their asset in stocks and companies might raise their business funds from stock market. Since the number of investors is increasing everyday in this century, the intelligent decision support systems aiding them to trade are keenly needed. But attempts on modelling or predicting the stock market have not been successful in *consistently* beating the market. This is the famous Efficient Market Hypothesis (EMH) saying that the future prices are unpredictable since all the information available is already reflected on the history of past prices [4]. However, if we step back from *consistently*, we can find several empirical results saying that the market might be somewhat predictable [1].

Many of technical indicators such as moving averages have been developed by researchers in economic area [3]. There are some weak points in technical indicators. For example, if we use RSI, we must specify its parameters. The curves of RSI are heavily influenced by the parameters. Also, there are some vagueness in their interpretations, which might be varied according to the subjectiveness of the interpreters.

In this paper, we propose a trend predictor of stock prices that can produce the probability distribution of trend states under the dynamics of trend and

Z.R. Yang et al. (Eds.): IDEAL 2004, LNCS 3177, pp. 794–799, 2004.

price of a stock. To model the dynamics, we design a hierarchical hidden Markov model, a variant of dynamic bayesian networks (DBN). Given observed series of prices, a DBN can probabilistically inference hidden states from past to current. Also we can sample or predict the future from learned dynamics. To use an indicator of bid and ask signals, it is more natural to use our HHMM model than technical indicators.

The resulting trading performance is compared with the performances of a technical indicator through a simulated trading on Korean stock market.

2 HHMM for Mining Trend Patterns

The hierarchical hidden Markov model is an extension of the hidden Markov model (HMM) that is designed to model domains with hierarchical structure and dependencies at multiple length scales [2]. In HHMM, there are production states emitting single observations and abstraction states emitting sequences. An abstract state calls sub-HHMM which is responsible for emitting single observations or calls recursively other sub-HHMMs. HHMM can be converted to HMM, but its inference and learning are prohibited since the conversion leas to severely increased node sizes and multi fan-ins. Given a sequence of length T, the original inference algorithm for HHMM was $O(T^3)$ which is prohibitive when long sequences. Fortunately, there is efficient techniques to convert HHMM to DBN allowing $O(T)$ inference time as usual DBN [2, 6]. Figure 1 is an HHMM for mining trend patterns of stock prices which is designed in this paper.

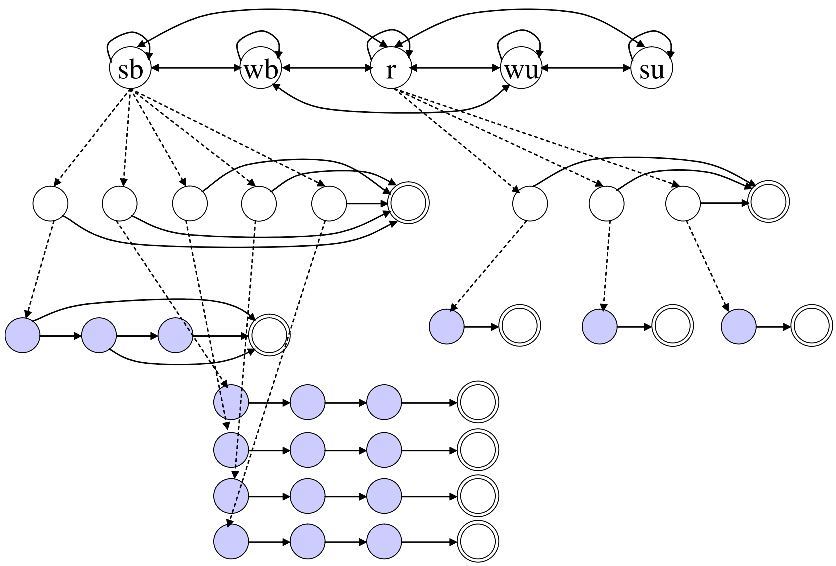

Fig. 1. State transition diagram for HHMM modelling the trend patterns of stock prices. Gray nodes are observation nodes, double circles are end states, and white circles are hidden states.

2.1 Fist Level

First level of HHMM is about the states of the trend. We divide the trend into five classes such as *strong bear* (`sb`), *weak bear* (`wb`), *random walk* (`r`), *weak bull* (`wu`), and *strong bull* (`su`) in Fig. 1. The sequence of price is the realization of the sub-HHMMs at the lower level called from this abstract states.

2.2 Second Level

In second level, there are five abstract states called from first level except the state `r`. These abstract states call their sub-HMMs of third level. In this paper, we choose observation at time t as relative closing price (RC) and is defined as

$$\mathrm{RC}(t) = \frac{\mathrm{close}(t) - \mathrm{close}(t-1)}{\mathrm{close}(t-1)}$$

where close(t) means closing price at day t.

As sub-HMM for the random walk, the simple mixture of Gaussian is introduced as shown the right-center in Fig. 1. The number of mixture components is set to three. And the hidden states of second level for random walk correspond to the mixture components. Each component has its own emission node at third level, which produces a Gaussian random variable.

After the k-means clustering on the pool of RC from the training data set, the gaussian emission node of each component is initialized with mean and covariance of clustered data. The transition probabilities from `r` to three abstract states of second level are initialized as the priors of clustered data.

2.3 Third Level

Bottom level HMMs are responsible for emission of outputs. We assume that there are five backbones which are Markov chains of length three. Each backbone is connected to one of abstract state in second level. To construct backbones, we use k-means clustering. From training data set, every RC sequence of length three is produced into a pool. Five clusters are produced from this pool by k-means clustering algorithm. Figure 2 shows the results of clustering and Fig. 3 shows the prior probabilities of the center of each cluster. Four clusters except '4' are used to initialize four Markov chains as shown bottom in Fig. 1.

While cluster '4' takes about 35% of sequences, the center of the cluster is nearly deviated from zero. Therefore to compromise the information loss from k-means clustering, we design a flexible HMM for this cluster which is capable of escaping its backbone at anytime as shown left-center in Fig. 1.

The means and covariances of Gaussians in backbones are initialized as the results of clustering. The transition probabilities from first level except `r` to second level is initialized as priors of Fig. 3.

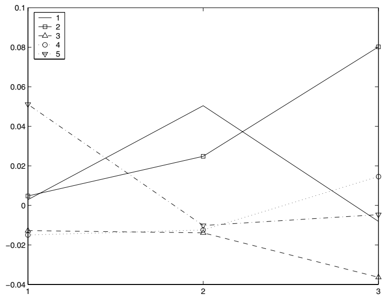

Fig. 2. Centers of five clusters.

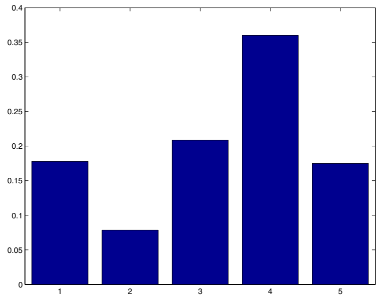

Fig. 3. Priors of five clusters.

2.4 Training the Model

To train our HHMM model, we adapt a semi-supervised learning techniques. Rather than learning HHMM from only observed RC sequences, we manually labelled the hidden states in the first level. We consider in this paper the trend as the interval to which the gradient of moving average of future price including current price belongs. Moving average (MA) is defined as

$$MA(i)^W = \sum_{t=0}^{W-1} \mathrm{close}(i+t)$$

where i is the i-th day and W is the length of window to be considered. The trend of i-day is labelled to one of five hidden states according to the gradient of the moving average line. Using the labelled data, the HHMM is trained by EM algorithm for dynamic Bayesian Networks.

3 Experiment

In Korean stock market, there are about 2,000 companies. Among those, we target the 20 candidates included in Korea Leading Company Index (KLCI) announced by Daewoo Securites Co., Ltd. The company names are listed in the first column of Table 1. The training data set is constructed from January 1998 to June 2003. From July 2003 to January 2004 is used as test set.

To construct, inference, and learn HHMM, we use *Bayes Net Toolbox* and slightly modify it [5]. Among several inference algorithms, `jtree_2TBN_inf_engine`, one of online smoothing algorithms is used.

Table 1 summarizes the trading results on the test period. Baseline 1 is the imaginary profit as if we would hold a stock during entire test period. Baseline 2 is the cumulative profit when we use `TRIX`, a technical indicator, of which gold cross is used as bid signal and dead cross is used as ask signal. HHMM is the cumulative profit of proposed method. We take a bid signal if the sum of probabilities of `su` and `wu` gets greater than 0.5 and a ask signal if the sum for `sb` and `wb` gets lower than 0.5.

Table 1. Comparison of the trading performances.

Company	Baseline 1	Baseline 2	HHMM
Samsung Electronics	47.38	14.67	29.78
SK Telecom	-1.39	-4.68	6.33
POSCO	25.87	2.08	18.66
KEPCO	10.96	1.07	4.91
KT	-8.80	-5.91	3.28
Hyundai Motors	52.21	27.61	32.14
LG Electronics	30.00	17.70	19.66
Samsung SDI	75.00	14.95	30.22
Hyundai MOBIS	52.55	49.78	53.83
Shinsegae	44.53	10.02	18.76
KT&G	32.58	16.08	17.12
LG Chemical	19.56	22.80	25.01
DSME	30.85	-2.20	33.12
KOGAS	0.74	-1.24	-0.50
Hankook Tire	74.49	29.35	24.53
NC Soft	13.49	3.76	15.77
Hanlla Climate Control	55.65	16.94	27.01
Cheil Industries	5.71	17.87	9.66
Daum Communication	-32.75	-11.42	-0.71
Handsome	2.91	4.07	21.09

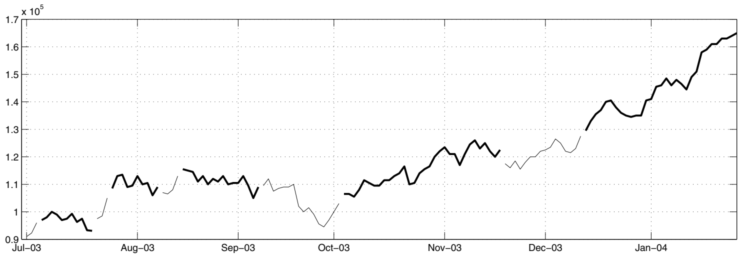

Fig. 4. Simulated trading on Samsung SDI.

Fig. 5. Hinton diagram.

For the test period, Korean stock market is in its bull market. So in several cases, Baseline 1, simple buy-and-wait strategy, wins other two methods. But HHMM outperforms Baseline 2 in most cases. Furthermore, HHMM minimizes the loss when the stock is declined in the whole period such as Daum Communication.

Figure 4 shows the simulated trading of Samsung SDI with HHMM in detail. The thick lines represent holding periods of stock. Figure 5 shows a Hinton diagram for filtered distribution of the trend states, that is $P(s_t|y_{t-19:t})$. Each row means `sb, wb, r, wu,` and `su` respectively from top to bottom. The larger white box is, the higher the probability is.

4 Conclusion

In this paper, a hierarchical hidden Markov model is designed to model the dynamics of trend of stock prices. The learned HHMM captures some characteristics of dynamics and the trend predicted by the model can be used as an effective indicator for trading signal.

There are some rooms to expand our HHMM model. The trading volume is generally assumed more chaotic than price series but it might be a key factor of the price trend through more complicated relations. We are trying to find a way to incorporate the information of trading volume.

Acknowledgement

This research was supported in part by the Ministry of Education and Human Resources Development under the BK21-IT Program. The RIACT at Seoul National University provides research facilities for this study.

References

1. Fama, E. F. and K. R. French. Dividend Yields and Expected Stock Returns. *Journal of Financial Economics*, 22, pp. 3-26, 1988.
2. Fine, S., Y. Singer, and N. Tishby, The Hierarchical Hidden markov Model Analysis and Applications, *Machine Learning*, 32, pp. 41-62, 1998.
3. Kendall, S. M. and K. Ord. *Time Series*. Oxford, New York, 1997.
4. Malkiel, B. G. *A Random Walk Down Wall Street*. Norton, New York, 1996.
5. Murphy, K. P. *Bayes Net Toolbox for Matlab*. http://www.ai.mit.edu/˜murphyk/Software/BNT/bnt.html, 2004.
6. Murphy, K. P. and M. Paskin. Linear Time Inference in Hierarchical HMMs. *In Proceedings of Neural Information Processing Systems*, 2001.