A fusion model of HMM, ANN and GA for stock market forecasting

Science Direct
Expert Systems with Applications 33 (2007) 171–180

Md. Rafiul Hassan, Baikunth Nath, Michael Kirley

Presenter: Chung, Wei-Chun
Date: July 23, 2008

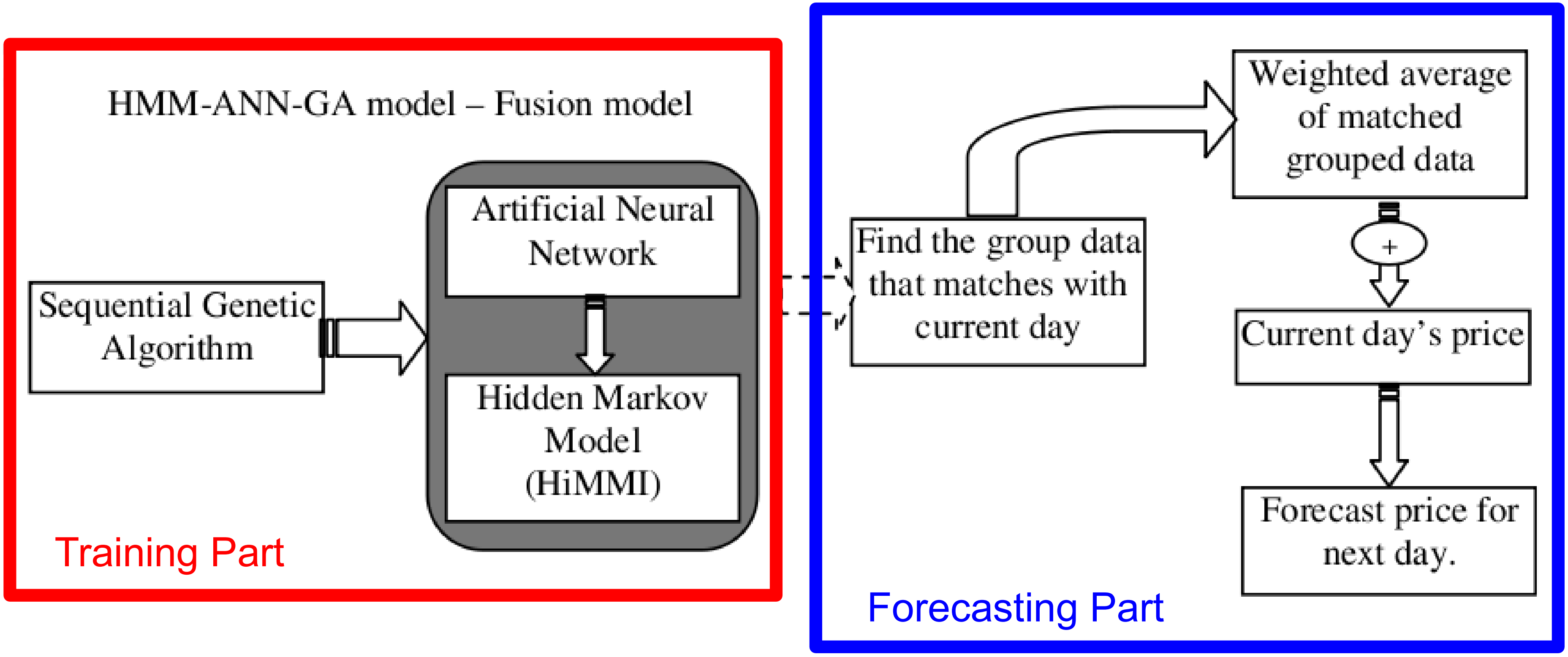
Abstract (1/2)

- In this paper we propose and implement a fusion model by combining the Hidden Markov Model (HMM), Artificial Neural Networks (ANN) and Genetic Algorithms (GA) to forecast financial market behaviour. The developed tool can be used for in depth analysis of the stock market. Using ANN, the daily stock prices are transformed to independent sets of values that become input to HMM. We draw on GA to optimize the initial parameters of HMM.

Abstract (2/2)

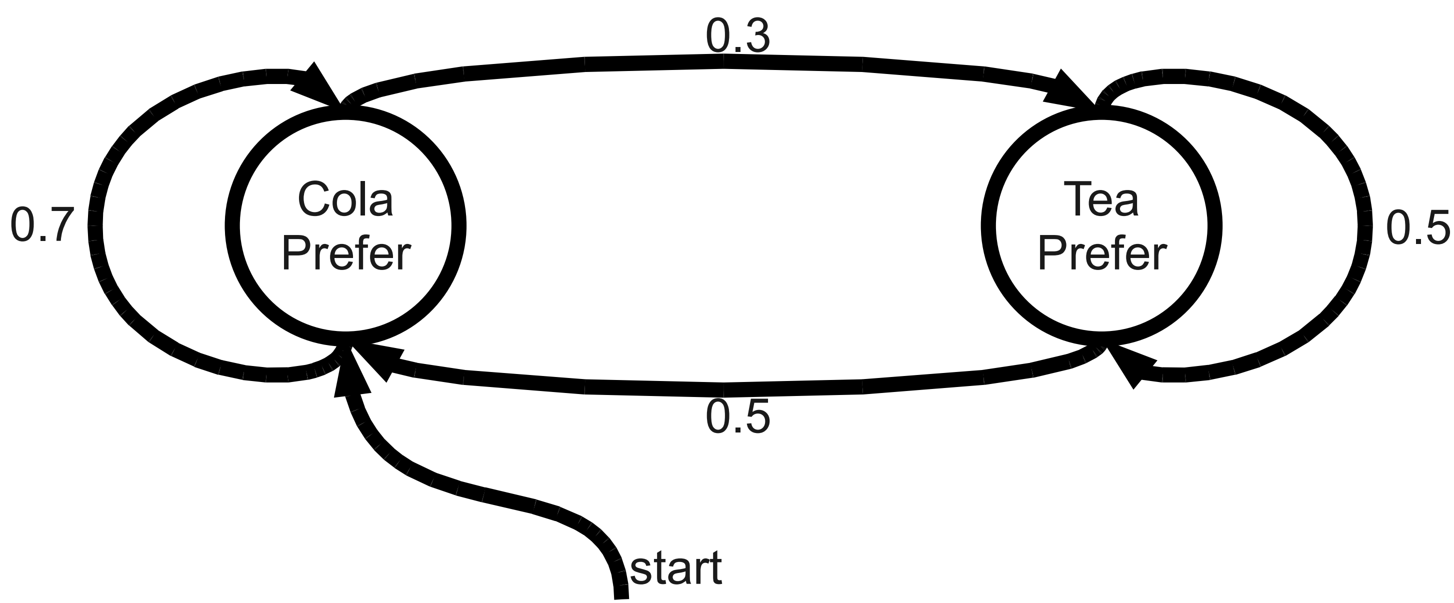
- The trained HMM is used to identify and locate similar patterns in the historical data. The price differences between the matched days and the respective next day are calculated. Finally, a weighted average of the price differences of similar patterns is obtained to prepare a forecast for the required next day. Forecasts are obtained for a number of securities in the IT sector and are compared with a conventional forecast method.

Block Diagram



Markov Model (1/2)

- Example: A crazy soft drink machine



Markov Model (2/2)

- What is the probability if the states={CP, IP}, and Output={lem_t,ice_t}?
- The solution is:
  0.7*0.3*0.7*0.1 + 0.7*0.3*0.3*0.1 +

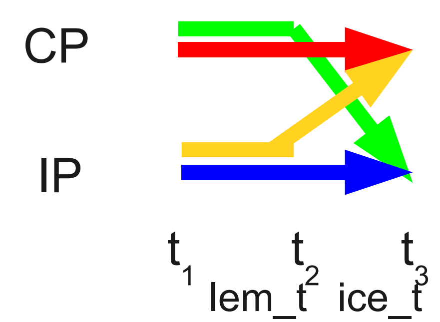
  0.3*0.3*0.5*0.7 + 0.3*0.3*0.5*0.7 = 0.084

Output probability

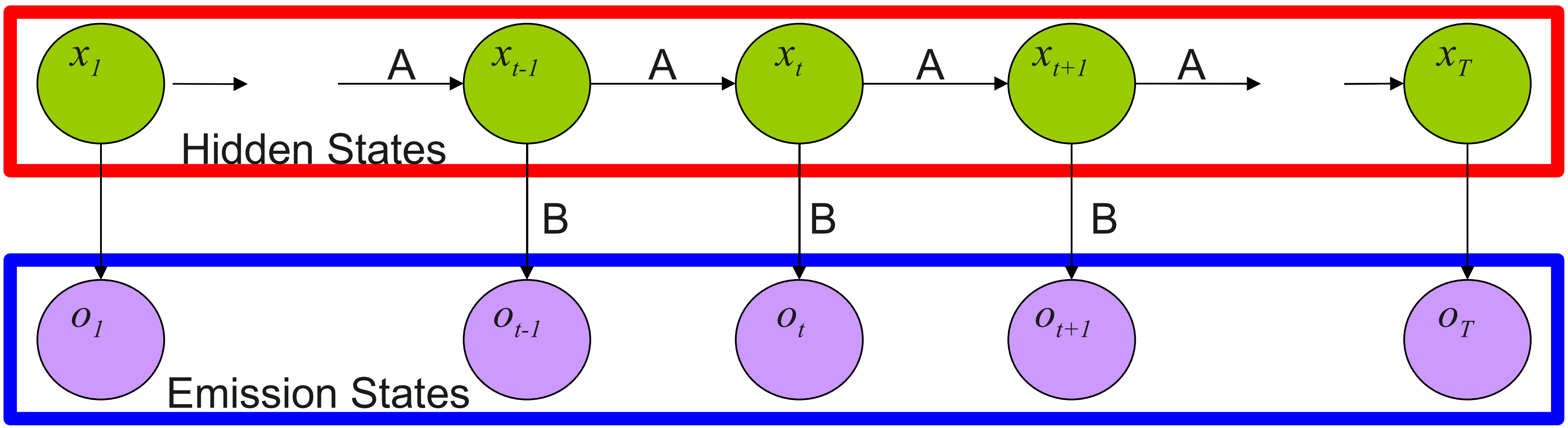
	cola	ice_t	lem_t
CP	0.6	0.1	0.3
IP	0.1	0.7	0.2

State transition probability

	CP	IP
CP	0.7	0.3
IP	0.5	0.5



Hidden Markov Model (1/2)



- HMM is characterized by
  - Number of states in the model (*S*)
  - Number of observation symbols (*K*)
  - The three main parameter matrices

Hidden Markov Model (2/2)

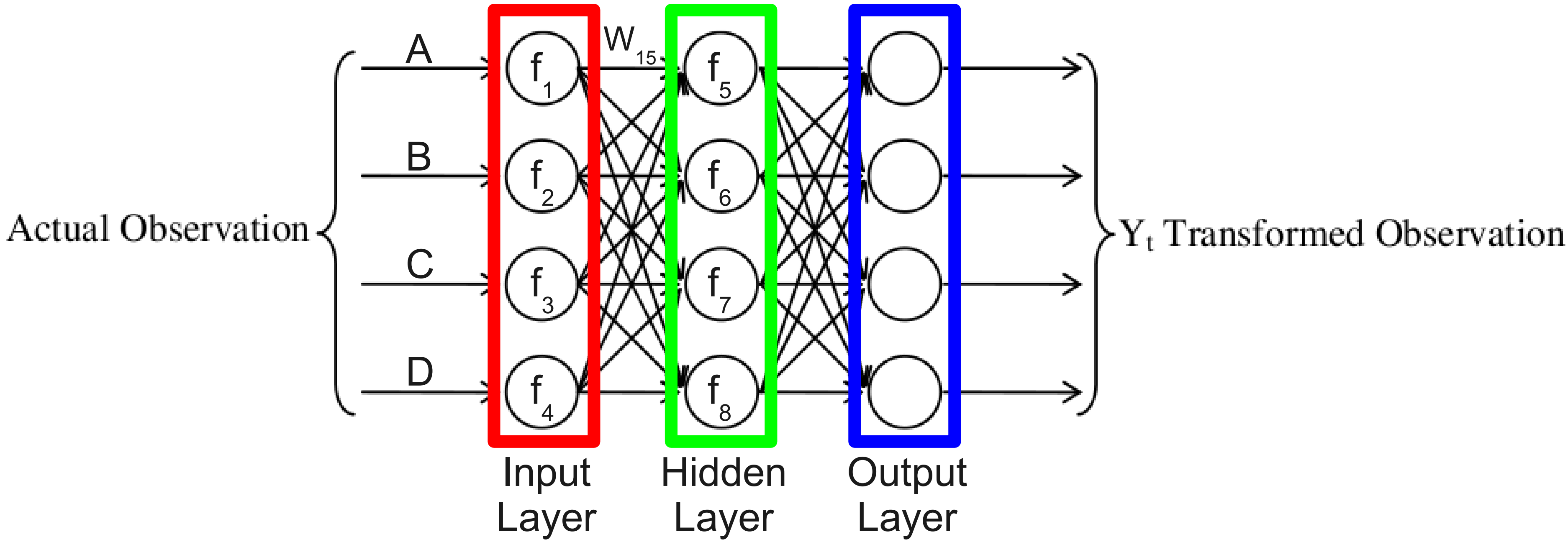
- Three main parameters
  - States transition probabilities (A)
  - Emission states probabilities (B)
  - Prior probabilities (π)
- Re-estimation (Baum-Welch algorithm): To find the values to best explain what we observed.
  - Forward algorithm (α)
  - Backward algorithm (β)
  - Best state sequence (γ)

HiMMI

- HiMMI [Hassan & Nath, 2005]
  - 3-D Gaussian distribution
  - Input: opening, high, low, closing price
  - Output: next day's closing price

	Opening price	High price	Low price	Closing price	Predicted closing price (30 Sep 2004)	Actual closing price (30 Sep 2004)
Today's data 29 Sep 2004	\$13.63	\$13.73	\$13.49	\$13.62	\$13.85	\$13.85
Matched data pattern using HMM 01 Jul 2003	\$17.1	\$17.2	\$16.83	\$17.13		
Next day's data 02 Jul 2003				\$17.36	13.85 = 13.62 + (17.36-17.13)	

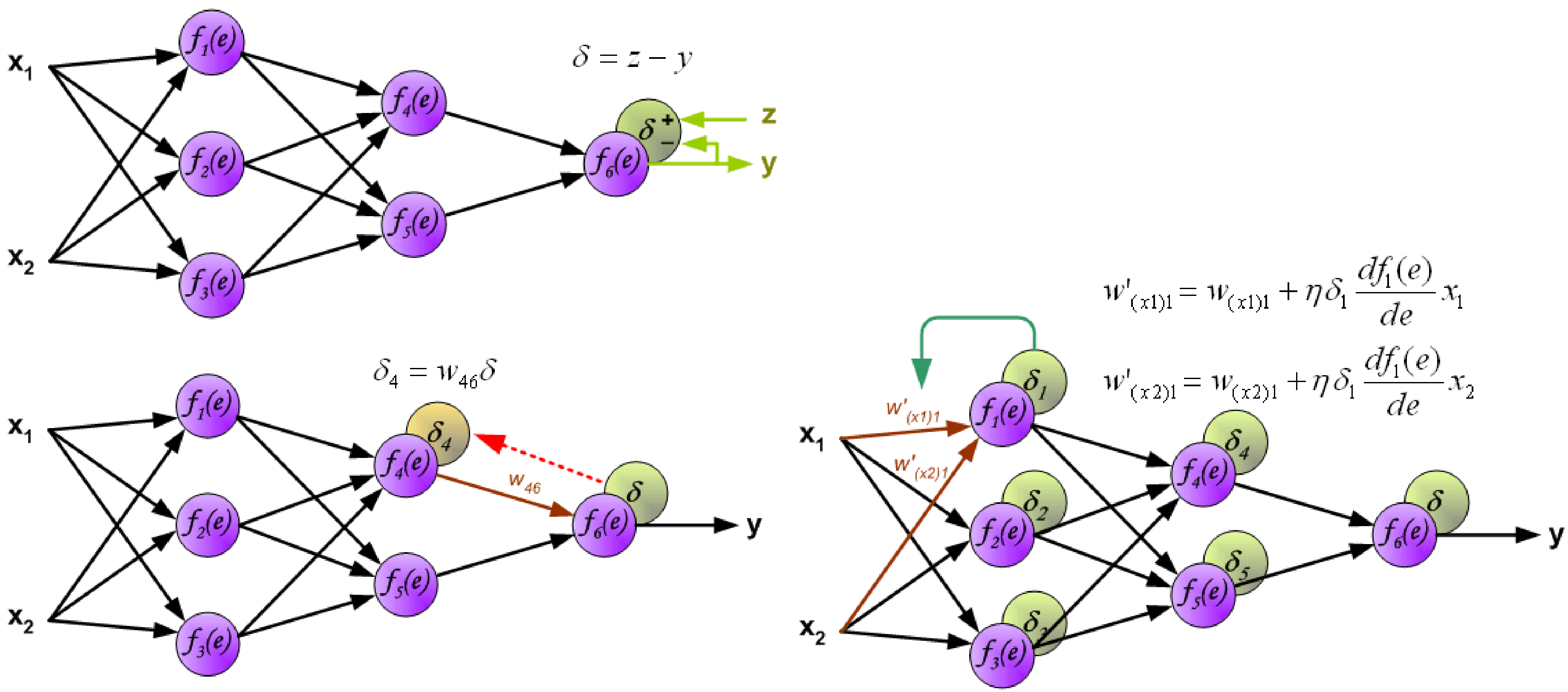
Artificial Neural Network (1/3)



- There are weights for each arc between layers.
- The output of node 5 is
  $y_5 = f_5(w_{15}f_1(A)+w_{25}f_2(B)+w_{35}f_3(C)+w_{45}f_4(D))$

Artificial Neural Network (2/3)

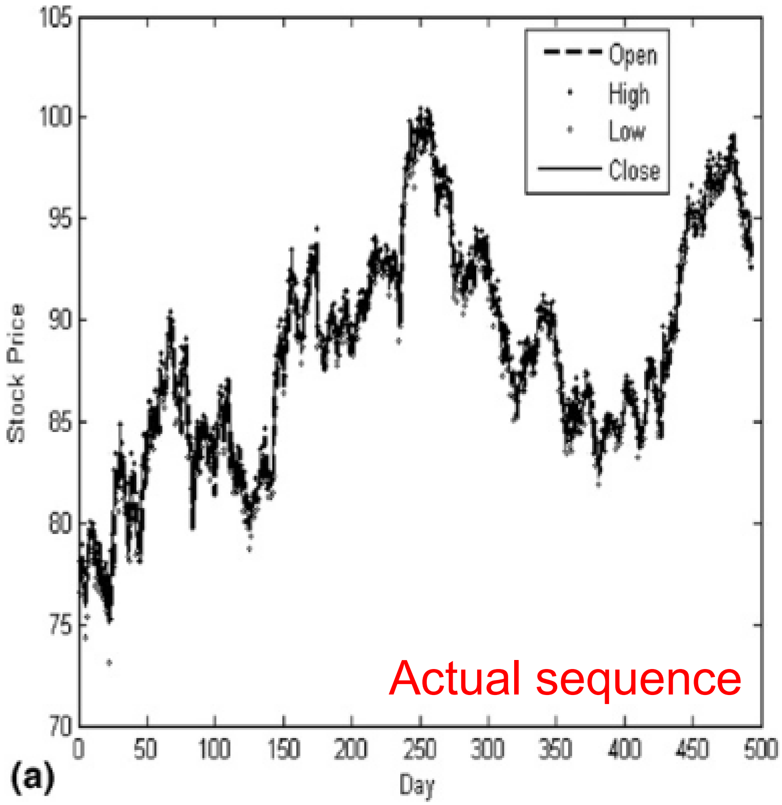
- Re-estimation: Back-propagation Algorithm

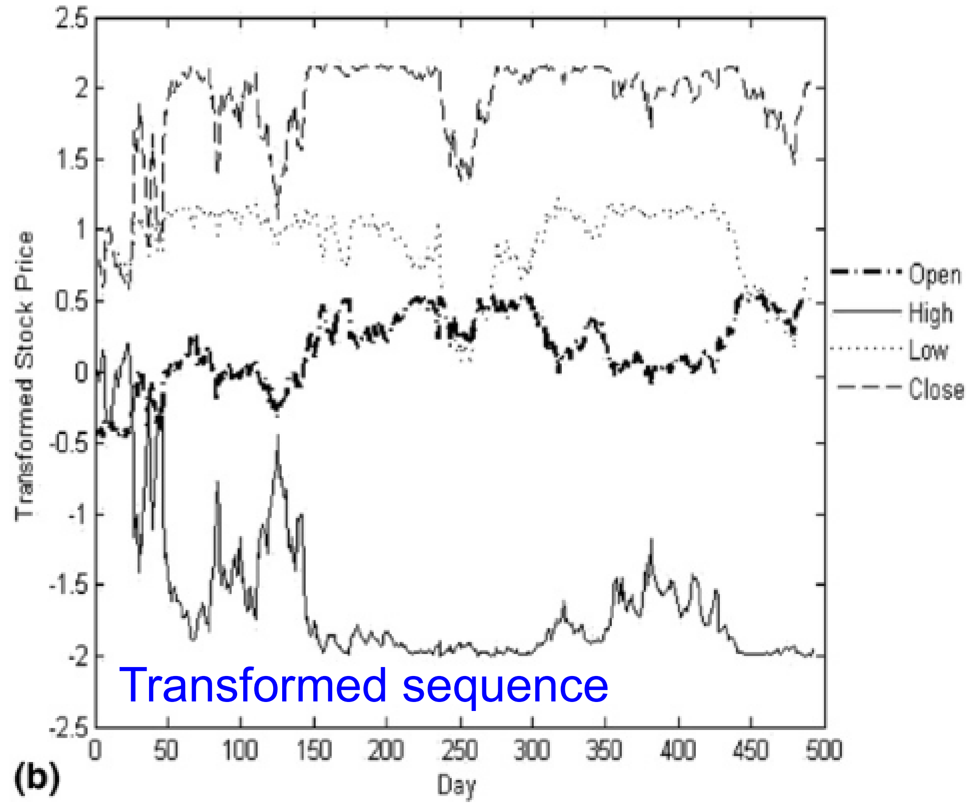


http://galaxy.agh.edu.pl/~vlsi/AI/backp_t_en/backprop.html

Artificial Neural Network (3/3)

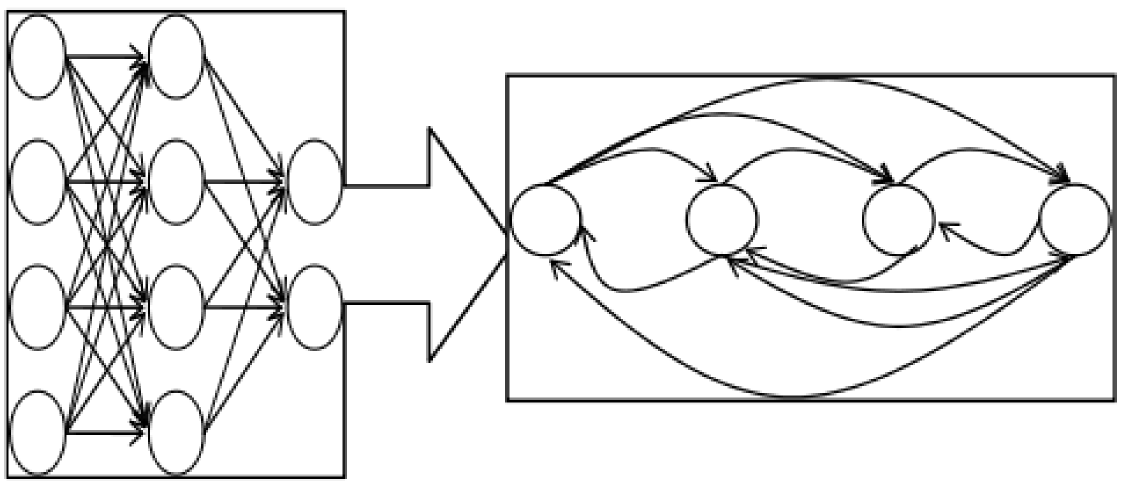
- The comparison of results using ANN





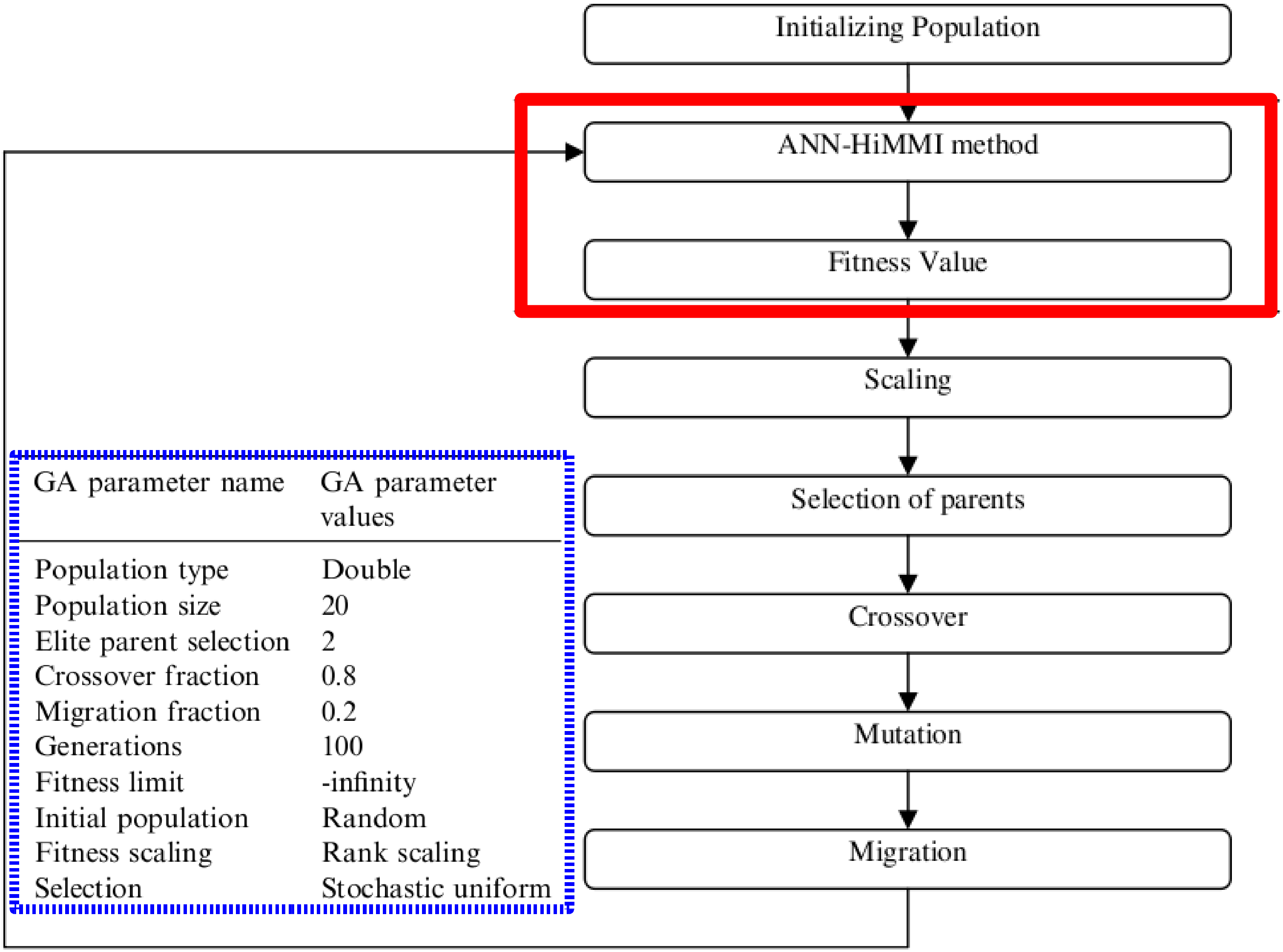
ANN-HMM Model

- The integrated model:
  - A feed-forward neural network
  - A 4-state hidden markov model



- But how to find a better model of HMM?
  - The parameters: *A*, *B* and π

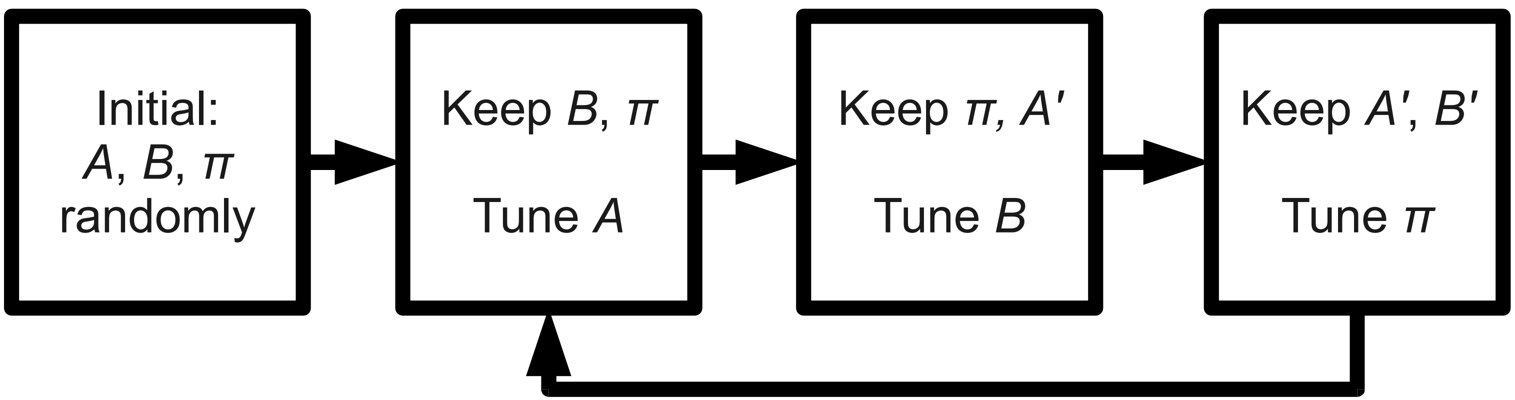
Genetic Algorithm (1/2)



GA parameter name	GA parameter values
Population type	Double
Population size	20
Elite parent selection	2
Crossover fraction	0.8
Migration fraction	0.2
Generations	100
Fitness limit	-infinity
Initial population	Random
Fitness scaling	Rank scaling
Selection	Stochastic uniform

Genetic Algorithm (2/2)

- For the performance and loading of calculation, the 3 matrices are tuned separately.



- For this 4-state HMM, the size of chromosomes are A=16, B=16, π=4.

Performance Criteria

- MAPE (Mean Absolute Percentage Error)

$$= \frac{\sum_{i=1}^{r} \left(\frac{\text{abs}(y_i - p_i)}{y_i} \right)}{r} \times 100\%$$

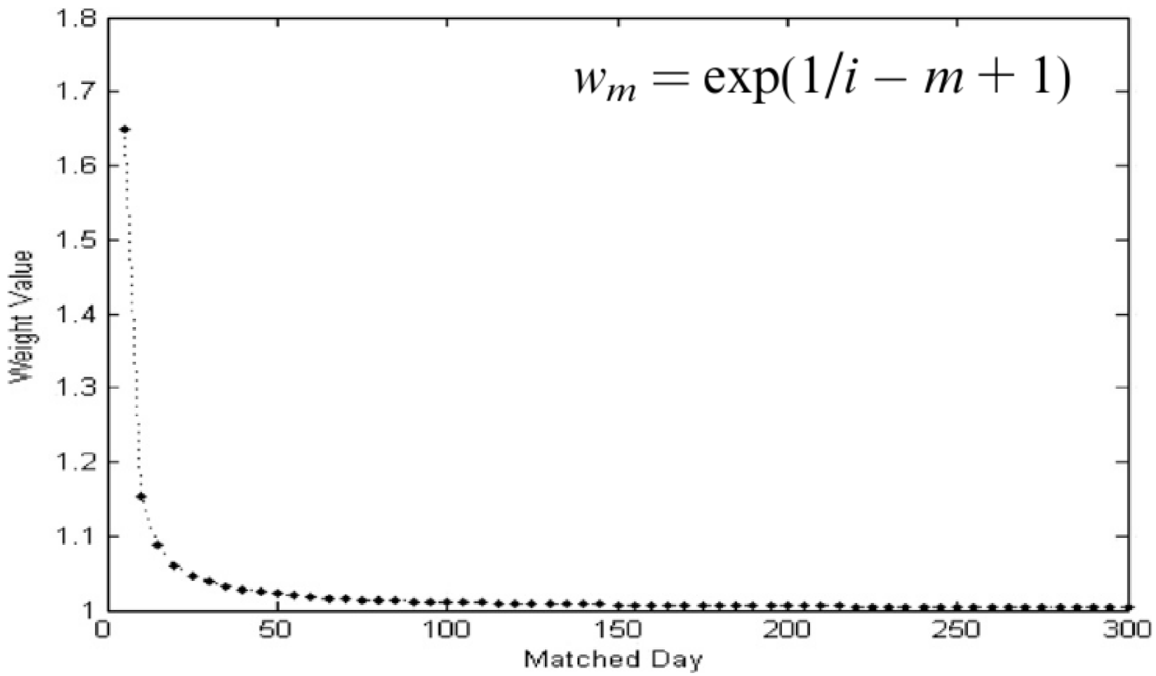
where:

r	total number of test data sequences
y_i	actual stock price on day i
p_i	forecast stock price on day i

Forecasting

- If the sequence is *a, b, c, d, e, f, g, h, i*
  MatchedDay *CurrentDay*
- Method: Weighted Average

$$\mathrm{wd}_i = \frac{\sum_m w_m^* \mathrm{diff}_m}{\sum_m w_m}$$



EX: $\mathrm{diff}_a = b - a$

Experiment Result (1/2)

Training and test data information

Stock name	Training data		Test data	
	From	To	From	To
Apple Computer Inc.	10 February 2003	10 September 2004	13 September 2004	21 January 2005
IBM Corporation	10 February 2003	10 September 2004	13 September 2004	21 January 2005
Dell Inc.	10 February 2003	10 September 2004	13 September 2004	21 January 2005

The performance improvement of the fusion models

Stock name	Mean absolute % error (MAPE) in forecast for 91 sequential test dataset		
	HiMMI	The proposed fusion model (ANN-GA-HMM-Interpolation)	The proposed fusion model with weighted average (ANN-GA-HMM-WA)
Apple Computer Inc.	2.8373	2.16492	1.9247
IBM Corporation	1.2186	1.0555	0.84871
Dell Inc.	1.01173	0.84463	0.699246

Experiment Result (2/2)

Training and test data information

Stock name	Training data		Test data	
	From	To	From	To
Apple Computer Inc.	10 February 2003	10 September 2004	13 September 2004	21 January 2005
IBM Corporation	10 February 2003	10 September 2004	13 September 2004	21 January 2005
Dell Inc.	10 February 2003	10 September 2004	13 September 2004	21 January 2005

Forecast accuracy comparison with the ARIMA (Autoregressive moving average model)

Stock name	Mean absolute % error (MAPE) in forecast for the 91 sequential test dataset	
	The proposed fusion model with weighted average	ARIMA
Apple Computer Inc.	1.9247	1.8009
IBM Corporation	0.84871	0.9723
Dell Inc.	0.699246	0.66035

Conclusion

- The comparison shows the forecasting ability of the fusion model is as good as that of ARIMA model which is a popular statistical forecasting tool.
- The proposed fusion model can be used without analysing the dataset prior to the forecast that users do not analysis before adopting the model.
- The number of states we choose as the number attributes in the observation vectors may not be suitable for some instances. We plan to employ another GA to find the best HMM architecture for a given dataset.