

Military Technical College
Kobry El-Kobba
Cairo, Egypt



12-th International Conference
on
Aerospace Sciences &
Aviation Technology

IDENTIFICATION OF PULSED RADARS USING HIDDEN MARKOV MODEL

HOSSAM E. ABOU-BAKR HASSAN*, KHAIRY ELBARBRY*
and MOHAMED ASAAD ABD-ELRAZEK*

ABSTRACT

In this paper, we propose radar system recognition approach using a HMM. We employ the Baum-Welsh algorithm to search for an HMM which best explains the observed radar signals represented by sequences of 0's and 1's. Deterministic periodic sequences (stable PRI radars) are considered. We can obtain an HMM which yields the globally biggest training probability. We can modify the results, to some degree, to create models more robust to observation errors. Preliminary results in combination with either the forward back-ward procedure or the Viterbi algorithm may be adequate for carrying out radar system recognition.

KEY WORDS

Signal processing, radar, Hidden Markov Model.

* Egyptian Armed Forces

1. INTRODUCTION

As radar technology advances, the complexity of electromagnetic signals is growing [1]. Thus, conventional electronic warfare (EW) assumptions that radar signals are stable and correspond to emitter modes can no longer be valid. This in turn implies that approaches to radar system recognition based on conventional techniques are no longer adequate.

In this paper, we propose radar system recognition approach using Hidden Markov Models (HMMs). Our motivation for using HMMs is as follows. First, HMMs provide flexibility for modeling dynamic behavior of radar systems. Second, HMMs are compatible with conventional receiver front-ends, and could be inserted into existing EW systems. Third, HMMs appear promising for enhancing the functions of threat libraries since our approach does not depend on the conventional assumptions. Fourth, HMMs have been extensively and successfully employed in speech recognition, control and communication applications [2, 3, 4]. The only hardware requirement of the proposed approach is the availability of a fast processor with sufficient memory.

Specification of an HMM involves the choice of the number of states, and the number of observation symbols, and the specifications of a model parameters as transition probability matrix, an observation symbol probability matrix and an initial state probability vector. Given the number of states as well as training sequences, we employ the Baum-welsh algorithm [2] to choose the HMM model that gives us the best results which explains the observed radar signals represented by sequences of 1's and 0's where a '1' represents a pulse transmitted by the radar and the number of 0's depends on the pulse repetition interval of the transmitting radars. The training results depend on the length of training sequences and observation errors (dropping and spurious pulses) in the training sequences.

For periodic deterministic signal sequences with uncorrupted radar signals we determine the model size using the periodic information of signals. In this case, the Baum-Welsh algorithm guarantees to converge to an HMM with yields the globally highest training probability. When observation errors are added, the foregoing results can be further modified by adjusting the observation symbol probability matrix to enhance the model's recognition capability. Unlike the modeling procedure in HMM theory for general problems, our methods to search for an appropriate model for radar signal recognition in the case of periodic deterministic sequence utilize the prior information pertaining to uncorrupted radar signals and therefore do not have to completely rely on trial and error.

Based on our modeling result, we then employ either the forward-backward (F-B) procedure [2] or the Viterbi algorithm [4] to carry out the task of recognition radar systems. Given sufficient observations, preliminary results suggest that the proposed approach may be adequate for radar system recognition.

The rest of this paper is organized as follows. The F-B procedure, the Viterbi and B-W algorithms are briefly reviewed in section 2. The training of HMMs for error-free and error-corrupted deterministic signal sequences are considered in Section 3. Radar signal recognition using HMMs is illustrated in Section 4. With an example consisting of three radars. Finally, we summarize this work in Section 5.

2- ALGORITHMS

Following [2], we define the parameters of discrete HMM as follows:

N: Number of states in the model.

M: Number of observation symbols.

$Q = \{q_1, q_2, \dots, q_N\}$, set of states

$V = \{v_1, v_2, \dots, v_M\}$, set of observation symbols

$O = \{o_1, o_2, \dots, o_T\}$, observation sequence

T: length of observation sequence.

$A = \{a_{ij}\}$ with $a_{ij} = P(q_j(t+1)|q_i(t))$, state transition probability matrix.

$B = \{b_j(l)\}$ with $b_j(l) = P(v_l(t)|q_j(t))$, observation symbol probability matrix.

$\pi = \{\pi_i | 1 \leq i \leq N\}$ with $\pi_i = p(q_i(t=1))$, initial state probability vector

$\lambda = (A, B, \pi)$, an HMM

A discrete HMM works as follows: we begin with choosing an initial state q_1 according to the initial state distribution π , set the clock $t=1$, choose an initial observation symbol according to b_1 , make a state transition to q_2 according to a_{ij} , increment the clock by one choose an observation symbol according to b_j and repeat this process until $t=T$.

2-1 Scoring

To determine which model is most likely to have generated O , a sequence of observation symbols, we need to calculate the probability of O given the model λ , i.e., $P(O|\lambda)$. This can be achieved using the F-B procedure.

2-2 F-B Procedure

To compute $P(O|\lambda)$ we can utilize forward probabilities $\alpha_t(i) = P(o_1, \dots, o_t, q_i | \lambda)$ defined by

$$(a) \quad \alpha_1(i) = \pi_i b_i(o_1) \quad 1 \leq i \leq N \quad (1)$$

For $t = 1, 2, 3, \dots, T-1$, $1 \leq i \leq N$

$$(b) \quad \alpha_{t+1}(j) = \sum_{i=1}^N \alpha_t(i) a_{ij} b_j(o_{t+1}) \quad (2)$$

Thus
$$p(O|\lambda) = \sum_{i=1}^N \alpha_T(i) \tag{3}$$

Similarly, we consider backward probabilities $\beta_t(i)$ defined by

(a)
$$\beta_T(i) = 1 \tag{4}$$

For $t = T - 1, T - 2, \dots, 2, 1$, $1 \leq i \leq N$

(b)
$$\beta_t(i) = \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \beta_{t+1}(j) \tag{5}$$

Thus
$$P(O|\lambda^{(l)}) = \sum_{i=1}^N \pi_i b_i(o_1) \beta_1(i) \tag{6}$$

Instead of the F-B procedure, the Viterbi can yield an approximation $P^*(O|\lambda)$ of $P(O|\lambda)$.

2-3 Viterbi algorithm

Step (1) Initialization:

$$\delta_1(i) = \pi_i b_i(o_1) \tag{7}$$

$$1 \leq i \leq N$$

$$\psi_1(i) = 0 \tag{8}$$

Step (2) Recursion:

For
$$\begin{aligned} 2 \leq t \leq T \\ 1 \leq j \leq N \end{aligned} \tag{9}$$

$$\delta_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] b_j(o_{t-1}) \tag{10}$$

$$\psi_t(j) = \max_{1 \leq i \leq N} [\delta_{t-1}(i) a_{ij}] \tag{11}$$

Step (3) Termination:

$$P^*(O|\lambda) = \max_{1 \leq i \leq N} [\delta_T(i)] \tag{12}$$

$$i_t^* = \arg \max_{1 \leq i \leq N} [\delta_t(i)] \quad (13)$$

Step (4) Path back tracking

For $t = T - 1, T - 2, \dots, 1$

$$i_t^* = \arg \max [\psi_{t+1}(i_{t+1}^*)] \quad (14)$$

2-4 B-W Algorithm

The B-W iterative algorithm is utilized for estimating various probability distribution for HMM as follows:

(a) $\pi_i = \gamma_1(i) \quad 1 \leq i \leq N \quad (15)$

(b)
$$a_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(j)} \quad (16)$$

(c)
$$b_j(l) = \frac{\sum_{t=1, o_t=l}^T \gamma_t(j)}{\sum_{t=1}^T \gamma_t(j)} \quad (17)$$

The quantity $\xi_t(i, j)$ is defined as the probability of being in state q_i at time t and being in state q_j at time $t+1$, given the observation sequence and the model. We can write $\xi_t(i, j)$ as

$$\xi_t(i, j) = \frac{\alpha_t(i) a_{ij} b_j(o_{t+1}) \beta_{t+1}(j)}{P(O|\lambda)} \quad (18)$$

The quantity $\gamma_t(i)$ is defined as the probability of being in state q_i at time t , given the observation and the model. We can write $\gamma_t(i)$ as

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j) = \frac{\alpha_t(i) \beta_t(i)}{P(O|\lambda)} \quad (19)$$

As t increases, $\alpha_t(i)$ and $\beta_t(i)$ decreases and can cause numerical underflow in the above iterative procedure. Thus, a scaling scheme is preferable. In the following, we provide the steps for incorporating the scaling scheme to the B-W algorithm.

$$\hat{\alpha}_t(i) = \prod_{\tau=1}^t c_\tau \alpha_\tau(i) \quad (20)$$

Where c_τ is obtained by means of the following steps:

$$1. \quad c_1 = \left[\sum_{i=1}^N \alpha_1(i) \right]^{-1} \quad (21)$$

$$\hat{\alpha}_1(i) = c_1 \alpha_1(i) \quad (22)$$

$$2. \quad \alpha_{t+1}(j) = \sum_{i=1}^N \hat{\alpha}_t(i) a_{ij} b_j(o_{t+1}) \quad (23)$$

$$c_t = \left[\sum_{i=1}^N \alpha_t(i) \right]^{-1} \quad (24)$$

$$\text{And} \quad \hat{\alpha}_t(i) = c_t \alpha_t(i) \quad (25)$$

$$\text{Then,} \quad \hat{\alpha}_{t+1}(i) = \prod_{\tau=1}^{t+1} c_\tau \alpha_{t+1}(i) \quad (26)$$

$$\alpha_{t+1}(i) = c_{t+1} \sum_{i=1}^N \hat{\alpha}_t(i) a_{ij} b_j(o_{t+1}) \quad (27)$$

$$\text{Let} \quad \hat{\beta}_{t+1}(i) = \prod_{\tau=t+1}^T c_\tau \beta_{t+1}(i) \quad (28)$$

$$\text{Then,} \quad \hat{\beta}_t(i) = \prod_{\tau=t}^T c_\tau \beta_t(i) \quad (29)$$

$$\hat{\beta}_t(i) = c_t \sum_{j=1}^N a_{ij} b_j(o_{t+1}) \hat{\beta}_{t+1}(j) \quad (30)$$

Based on the above results, we obtain

$$P(O|\lambda) = \sum_{i=1}^N \alpha_T(i) = \frac{\sum_{i=1}^N \hat{\alpha}_T(i)}{\prod_{\tau=1}^T c_\tau} \quad (31)$$

$$P(O|\lambda) = \left[\prod_{\tau=1}^T c_\tau \right]^{-1} \quad (32)$$

Taking the logarithm for both sides gives

$$\ln P(O|\lambda) = -\sum_{\tau=1}^T \ln(c_\tau) \quad (33)$$

Let
$$C_T = \prod_{\tau=1}^T c_\tau \quad (34)$$

$$\hat{\xi}_t(i, j) = \frac{\hat{\alpha}_t(i) a_{ij} b_j(o_{t+1}) \hat{\beta}_{t+1}(i)}{C_T P(O|\lambda)} \quad (35)$$

$$\hat{\gamma}_t(i) = \frac{\hat{\alpha}_t(i) \hat{\beta}_t(i)}{c_i C_T P(O|\lambda)} \quad (36)$$

Thus, the B-W reestimation procedure of the parameters can be rewritten as follows:

(a)
$$\pi_i = \hat{\gamma}_1(i) \quad (37)$$

(b)
$$a_{ij} = \frac{\sum_{t=1}^{T-1} \hat{\xi}_t(i, j)}{\sum_{t=1}^{T-1} \hat{\gamma}_t(i)} \quad (38)$$

(c)
$$b_j(k) = \frac{\sum_{t=1, o_t=v_k}^T \hat{\gamma}_t(j)}{\sum_{t=1}^T \hat{\gamma}_t(j)} \quad (39)$$

3- MODELING DETERMINISTIC SEQUENCE

We consider deterministic sequence. It can be shown that there exist an HMM λ for an observed deterministic sequence O such that the training probability $p(O|\lambda)=1$. We refer to such an HMM as a perfect model for sequences O .

To find a perfect model for sequence O , we begin searching for a state machine that can track the sequence as long as its state output and transitions follow the evolution of the sequence. This suggests that as the length of the sequence grows, so does the size of the state machine to track the sequence. Fortunately, when a sequence is periodic, the size of the state machine only needs to be at least the length of the cycle of the sequence. In the following, we begin with an example to illustrate the existence of a perfect model for any given sequence.

The sequence 00101 denoted as O is non-periodic. One of its periodic counter parts for example, is the sequence 0010100101 with period 2. To track the sequence 00101, it suffices to use a state machine with 5 state depicted in the following figure where the first state is the initial state.

In Figure 1, if the states are labeled in the ascending order 1, 2, 3, 4 and 5 then the parameters of the perfect model for the non-periodic sequence 00101 and its periodic counter parts are given as follows.

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (40)$$

$$B = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \end{bmatrix} \quad (41)$$

$$\pi = [1 \ 0 \ 0 \ 0 \ 0] \quad (42)$$

To verify that the HMM λ , given in Figure 1, is a perfect model for O we ran the Baum-Welsh algorithm using some 5-state initial guesses A , and B . Indeed, the Baum-Welsh algorithm yield the HMM given Figure 1 such that $P(O|\lambda)=1$. Even if a perfect model exists, the training process with respect to different initial guesses was perfectly chosen. As there are $5!$ Permutations for state labels 1, 2, 3, 4 and 5, there are $5!$ 5-state perfect models that is independent of c to one another. Also note that there exist many n -state ($n \geq 6$) perfect models for the sequence 00101. This is shown in Figure2.

The following proposition essentially states that any perfect model for non-periodic sequence is also a perfect model for its periodic counterparts.

Given a non-periodic sequence S with length N and a perfect model λ for S (i.e., $P(s|\lambda)=1$), λ is a perfect model for sequence O with length kN where $k > 1$ whenever O is a sequence with S periodically k times.

In practice, a radar signal may be corrupted due to observations errors. When a perfect model is used for recognizing any erroneous set sequences, it yields a zero recognition probability. This means that a perfect model cannot tolerate any error in the testing sequence. To recognize erroneous sequence to some extent we need to use an HMM robust to errors to obtain such an HMM. We use some erroneous sequence as a training sequence and take the parameters π , A and B of a perfect model as initial guesses. It turns out that the training simply adjusted the matrix B of the perfect model to accommodate errors.

This phenomenon is illustrated in the following.

Let O be the error-free sequence 0010100101 its length is 10, thus there are $\binom{10}{x}$ x -symbol error sequence for O . A perfect model O is given in Figure1 we use the vector π , the matrix A of the perfect model and take the matrix B

$$B = \begin{bmatrix} 0.9999 & 0.9999 & 0.0001 & 0.9999 & 0.0001 \\ 0.0001 & 0.0001 & 0.9999 & 0.0001 & 0.9999 \end{bmatrix} \quad (43)$$

as initial guesses π , A , B to conduct the training of an HMM robust to errors. Using $\binom{10}{x}$ x -symbol error sequence as observation sequences for training, we obtain that the vector π , the matrix A of the perfect model were unaltered. But the matrix B of the perfect model was adjusted to accommodate errors and became

$$B = \begin{bmatrix} 1 - \frac{x}{10} & 1 - \frac{x}{10} & \frac{x}{10} & 1 - \frac{x}{10} & \frac{x}{10} \\ \frac{x}{10} & \frac{x}{10} & 1 - \frac{x}{10} & \frac{x}{10} & 1 - \frac{x}{10} \end{bmatrix} \quad (44)$$

In this case, the elements $B(1, j)$ and $B(2, j)$ of the perfect model, $1 \leq j \leq 5$ were adjusted linearly with respect to the ratio $\frac{x}{10}$.

4- RECOGNITION

In this section, we illustrate with an example that modeling using HMM given in section 3 may be appropriate for recognition of pulsed radars. Assume the uncorrupted sequence 1011 is repeatedly emitted by radar 1, and so are 01011 by radar 2 and 011101 by radar 3. Based on the results in section 3 we can obtain a perfect model for error-free periodic sequence emitted by each radar. Assume 10% dropping pulses and 5% spurious pulses may occur in a given error-free sequence. In this case we trained robust HMMs in order to ensure that the recognition task can be performed adequately. First, the length of the training sequences was fixed at 120. Second, twenty erroneous training sequences were randomly generated each time with respect to a given periodic error-free sequence using B-W algorithms yield HMMs. $\lambda_i = (A_i, B_i, \pi_i)$ corresponding to radars, $i = 1, 2, 3$

$$A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \quad (45)$$

$$B_1 = \begin{bmatrix} 0.1000 & 0.8000 & 0.1333 & 0.1667 \\ 0.9000 & 0.2000 & 0.8667 & 0.8333 \end{bmatrix} \quad (46)$$

$$\pi_1 = [1 \quad 0 \quad 0 \quad 0] \quad (47)$$

$$A_2 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (48)$$

$$B_2 = \begin{bmatrix} 0.8750 & 0.1250 & 0.8750 & 0.1667 & 0.2083 \\ 0.1250 & 0.8750 & 0.1250 & 0.8333 & 0.7917 \end{bmatrix} \quad (49)$$

$$\pi_2 = [1 \quad 0 \quad 0 \quad 0 \quad 0] \quad (50)$$

$$A_3 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad (51)$$

$$B_3 = \begin{bmatrix} 0.8500 & 0.2500 & 0.2000 & 0.0000 & 0.8500 & 0.1500 \\ 0.1500 & 0.7500 & 0.8000 & 1.0000 & 0.1500 & 0.8500 \end{bmatrix} \quad (52)$$

$$\pi_3 = [1 \ 0 \ 0 \ 0 \ 0 \ 0] \quad (53)$$

We generated a testing sequence by letting 10% dropping pulses and 5% spurious pulses to occur in the periodic error-free sequence emitted by radar. In this way, we have three testing sequences as follows:

O_1 : 101110111011.....

O_2 : 010110101101011.....

O_3 : 011101011101011101.....

Where O_i are corrupted sequences corresponding to the error-free sequences emitted by radars i , $i=1, 2, 3$.

The Forward Backward procedure yielded Tables 1-3

Table 1

Results	$\log P(O_1 \lambda_1)$	$\log P(O_1 \lambda_2)$	$\log P(O_1 \lambda_3)$
Without error	-0.0052	-216.0029	-280.0022
With error	-72.0044	-212.8029	-264.0023

Table 2

Results	$\log P(O_2 \lambda_1)$	$\log P(O_2 \lambda_2)$	$\log P(O_2 \lambda_3)$
Without error	-216.0029	-0.0052	-224.0028
With error	-222.4028	-72.0044	-236.8026

Table 3

Results	$\log P(O_3 \lambda_1)$	$\log P(O_3 \lambda_2)$	$\log P(O_3 \lambda_3)$
Without error	-280.0022	-224.0028	-0.0052
With error	-265.6023	-233.6027	-72.0044

The Viterbi algorithm yields similar results to Tables 1-3. Based on these results, we conclude that HMM1 recognizes O_1 since $\log P(O_1|\lambda_1)$ is the largest in Table 1, we conclude that HMM2 recognizes O_2 since $\log P(O_2|\lambda_2)$ is the largest in Table 2, we conclude that HMM3 recognizes O_3 since $\log P(O_3|\lambda_3)$ is the largest in Table 3

5- CONCLUSIONS

This paper presents preliminary results for recognition of radars using HMM. A perfect model indeed exists for any given deterministic sequences. To yield HMM more robust to errors, we use erroneous sequences as a training sequences and the parameters of a given perfect model as initial guesses. It is noted that the training simply adjusts the B matrix of the perfect model to accommodate errors. The experimental results show that using HMM may be a good approach for radar recognition.

6- REFERENCES

- [1] Natharson, F.E., Reilly, J.P. and Cohen, M.N., Radar Design principles: Signal Processing and the Environment, New York: McGraw-Hill, 1991.
- [2] Rabiner, L.R., "A Tutorial on Hidden Markov Models and Applications in Speech Recognition," Proc. IEEE, Vol.77, NO.2 Feb. pp 257-285, 1989.
- [3] Elliort, R.J. Aggonu, L., and Moore, J.B., Hidden Markov Models: Estimation and Control, New York: Springer-Verlag, 1995.
- [4] Forneg, G, D., Jr. "The Viterbi Algorithm" Proc. IEEE, Vol. 61 pp 268-278, Mar., 1973.

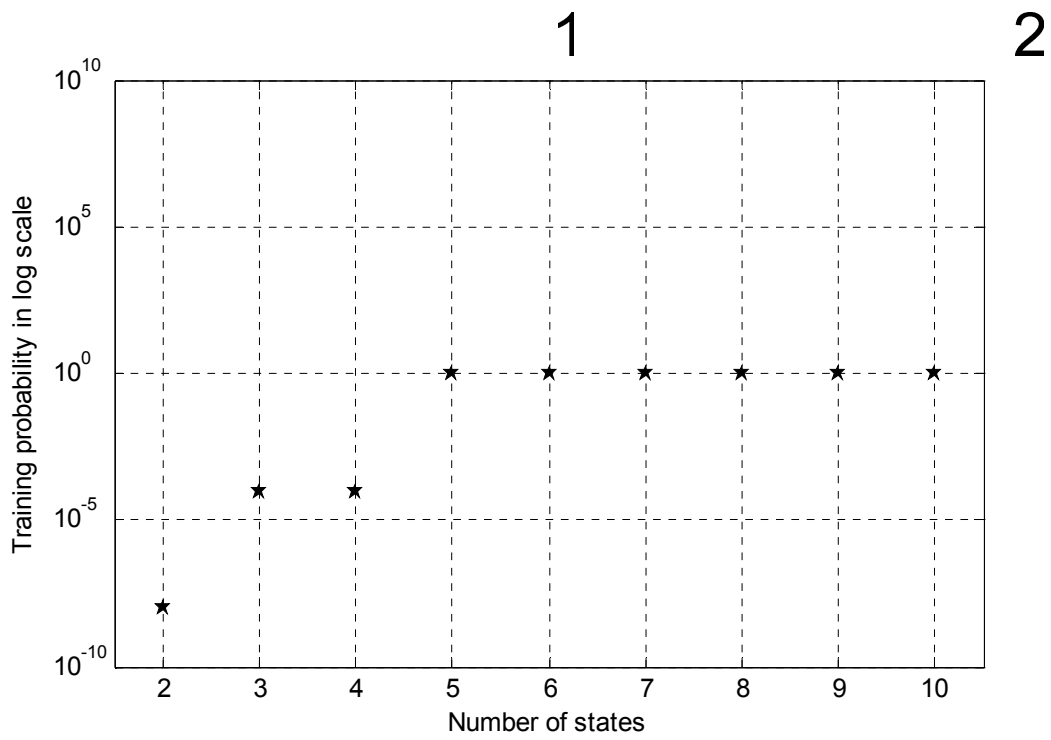
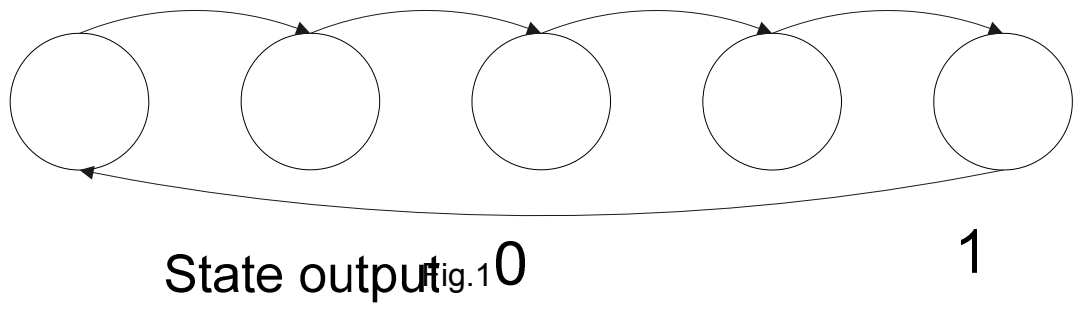


Fig.2. Probability of perfect model