# NEW METHOD FOR STRUCTURAL CHANGE DETECTION OF TIME SERIES AS AN OPTIMAL STOPPING PROBLEM

Hiromichi Kawano, Ken Nishimatsu
*NTT Service Integration Laboratories, Musashino-shi, Tokyo, 180-8585 Japan*

Tetsuo Hattori
*School of Enginieering, Kagawa University, Takamatsu City, Kagawa, 761-0396 Japan*

Keywords: Time series, Structural change, Dynamic Programming, Optimal stopping problem

Abstract: In general, an appropriate prediction expression and/or model is constructed to fit a time series though, the model begins to unfit (or not to fit) the time series from some time point, especially in the field that relates to human activity and social phenomenon. In such case, it will be important not only to quickly detect the unfitting situation but also to rebuild the prediction model after the detection as soon as possible. In this paper, we formulate the structural change detection problem in time series as an optimal stopping problem, using the concept of DP (Dynamic Programming) with a cost function that is the sum of unfitting (or not fitting) loss and action cost to be taken after detection. And we propose a method for optimal solution and show the correctness by proving a theorem. Also we clarify the effectiveness by showing the numerical experimentation.

## 1 INTRODUCTION

Change point detection (CPD) problem in time series is to find that a structure of generating data has changed at some time point by some cause. We consider that the problem is very important and that it can be applied to a wide range of application fields.
For example, degradation detection in communication system (R.Jana and S.Dey, 2000), object detection on a radar screen ( R.M.Gagliardi and I.S.Reed, 1965), speech processing (R.J.Di Francesco, 1990), and fault detection (A.S.Willsky, 1996), (D.Kauame, et al., 1996) are such application examples of the CPD problem.

The processing method for the CPD problem is roughly divided into two types: one is batch processing that checks all generated data in the past and another is sequential processing that checks if the structure has changed or not at every new data generation.

As the former representative method, Chow test is well known and is often used in econometrics (Chow,G.C., 1960). It does a statistical test by setting the hypothesis that the change has occurred at time t. However, the problem of Chow test exits in that we have to give the change time t for the hypothesis setting, and also in that the test lacks the rapidity to detect the change point.

As the latter representative method, there are Bayes' method (S.MacDougall, A.K.Nandi and R.Chapman, 1998), (V.V.Veeravalli and A.G.Tartakovsky, 2002) and CUSUM one (E.S.Page, 1954), (C.Han, P.k.Willet and D.A.Abraham, 1999), (S.D.Blostein, 1991), (Y.Liu and S.D.Blostein, 1994), (M.Basseville and I.V.Nikiforov, 1993), (M.Basseville, 1988), based on sequential probability ratio test. The Bayes' method can detect the change point, based on the sequential estimation of posterior probability, if the generation distribution of time series data is known at the time before and after the change point. So the Bayes' method can solve the problem in the Chow test, but it requires that the generation distribution for the time series data is already known.



Kawano H., Nishimatsu K. and Hattori T. (2004).
NEW METHOD FOR STRUCTURAL CHANGE DETECTION OF TIME SERIES AS AN OPTIMAL STOPPING PROBLEM.
In *Proceedings of the First International Conference on Informatics in Control, Automation and Robotics*, pages 24-31
DOI: 10.5220/0001147500240031

Moreover, in practical situation, we have to consider not only that a loss cost is involved with prediction error but also that an action to be taken after the change detection will need a cost. Conversely, the CPD is necessary in order to judge when to take the action.

Taking the field of network management for example, time series data (e.g. error rate and delay) of the quality are always monitored, and when the structural change is detected, some action for the quality improvement is taken.

In the structural change detection under such situations, we must consider the trade-off between loss by the degradation and cost for the quality reformation.

However, as far as the authors can know, no such conventional CPD method considering the action cost has been proposed, in spite of the fact that such method is very useful at practical level.

In this paper, in order to solve difficulties in conventional methods for structural change detection in time series, we propose a new and practical method based on an evaluation function of loss cost. And we formulate the CPD problem as an optimal stopping problem using the concept of DP (Dynamic Programming) and give the optimum solution in the formulation. We consider that our method is effective in the sense as follows.

1. Differently from the Chow test, it does not need to set the change point in a priori.
2. Unlike the Bayes' method, it does not need to give the generation distribution of time series data.
3. It can quickly detect the structural change point by the sequential processing.
4. It minimizes the evaluation function that sums up the loss involved with prediction error and action cost to be taken after the change detection.
5. It is a meta-level method so that we can apply it to any prediction model in the evaluation function.

Also in this paper, we present the correctness of our solution by proving a theorem and show the effectiveness by numerical experimentation results.

# 2 FORMULATION

## 2.1 Evaluation Function

We formulate the CPD problem as an optimal stopping one based on an evaluation function that sums up the cost involved by prediction error and action cost to be taken after the change detection.

For example, a prediction expression is given in the following equation as a function of time **t**, where $y_t$, $\beta_1$, β0, ε mean the function value, two constant coefficients, and error term, respectively.

$$y_t = \beta_1 \cdot t + \beta_0 + \varepsilon \qquad (1)$$

The error term ε is given as a random variable of the normal distribution of variance σ and average of 0, i.e., $\varepsilon \sim N(0, \sigma)$.

A time series data based on the Equation (1) is shown in Figure 1, that is generated by making normal random numbers of average 0 and variance 1 for ε, and by setting $\beta_1$=0.2, β0=1 for the time t=1,2,…,70, and $\beta_1$=0.8, β0= - 41 for the time after t=71.

The tolerant error interval or tolerance zone between two broken lines as shown in Figure 1 is decided using the first time series data from t=1 to t=20.

Using those data, the prediction expression is made by the least squares method, and the tolerant interval of error is calculated as 95% confidence interval of the sample variance of residual ε.

Note that the tolerant error interval is not based on the confidence interval of regression formula given in the following Equation (2), but is defined based on the distribution of error term ε of Equation (1)(N.R.Draper, H.Smith, 1996).

$$y = \hat{\beta}_0 + \hat{\beta}_1 x$$

$$\pm t(n-2,\alpha)\sqrt{\left(1+\frac{1}{n}+\frac{(x-\bar{x})^2}{S_{xx}}\right)V_e} \qquad (2)$$

where $t$, α, $\bar{x}$, and $n$ denote $t$ distribution, significance level, average, and the number of data, respectively, and let $e_i$ be prediction error in time $t_i$, $V_e$ and $S_{xx}$ are defined as follows.

$$V_e = \sum_{i=1}^{n} e_i^2 \Big/ n \qquad S_{xx} = \sum_{i=1}^{n} (x_i - \bar{x})^2 .$$

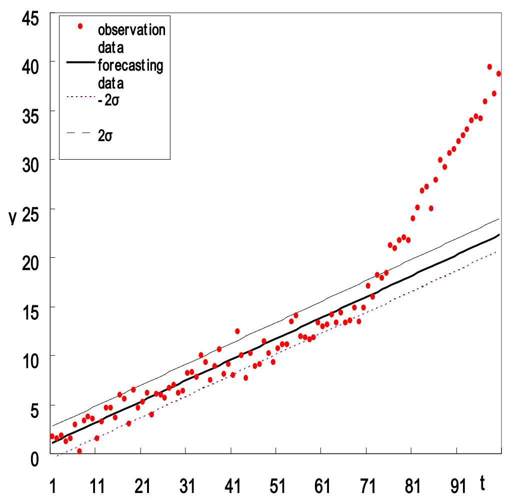


Figure 1: Example of time series data.

In Figure 1, it can be read that the generated data runs out frequently from the tolerance zone since after t=70. From the fact, when the difference (or observation error) between the observed data and forecasted value exceeds a specified tolerance (i.e., when the observed data goes out from the tolerance zone), we can think that that there is a high possibility that the structural change has occurred.

For simplicity, we think two situations; one is the situation that the observed data is out from the tolerance zone, and another the situation that the observed data is in the zone. Then we call the former situation "*unfitting*" and the latter "*fitting*". Based on this discussion, we consider that the structure has changed, when the unfitting occurs between sequentially observed data and forecasted value by continuing *N* times. This specified tolerance is defined as, e.g., 2σ of the distribution on error ε that is estimated at the time when the prediction expression is made.

The evaluation function is given in (3) as the sum of two kinds of cost: the damage caused by the unfitting (i.e., unfitting loss) and action cost to be taken after the change detection.

$$\text{Total\_cost} = \text{cost}(A) + \text{cost}(n) \quad (3)$$

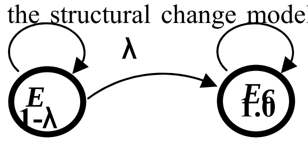
where cost($n$) is the sum of the loss by continuing $n$ times unfitting before the structural change detection, and cost($A$) is the cost involved by the action after the change detection.

Taking the quality control problem for example, the above cost($n$) means the loss caused by the quality degradation and superfluous quality. And the cost($A$) means the cost involved by some facility replacement.

Since the observed time series data is a random variable and the unfitting event is stochastic, the value of the evaluation function Total_cost also becomes a random variable. Then we have to find the number of times *N* that minimizes the expectation value of Total_cost, under the assumption that the structural change occurs randomly. Note that the evaluation function can be defined if only the distribution of error ε is given, so there is no need for the prediction expression to be such a form like equation (1).

## 2.2 Structural Change Model

We assume that the structural change is Poisson occurrence of average λ , and that, once the change has occurred during the observing period, the structure does not go back to the previous one. The reason why we set such a model is that we focus on the detection of the first structural change in the sequential processing (or sequential test). The concept of the structural change model is shown in Figure 2.




***Ec*** : State that the structural change occurred.
***E*** : State that the structure is unchanged.
**λ** : Probability of the structural change occurrence. (Poisson Process.)

Figure 2: Structural change model

Moreover, we introduce a more detailed model. Let *R* be the probability of the unfitting when the structure is unchanged. Let *Rc* be the probability of the unfitting when the structure change occurred. We can consider that *Rc* is greater than *R*, i.e., *Rc*>*R*.

The detailed internal model for the State ***Ec*** and ***E*** are illustrated as similar probabilistic finite state automatons in Figure 3 and 4, respectively.

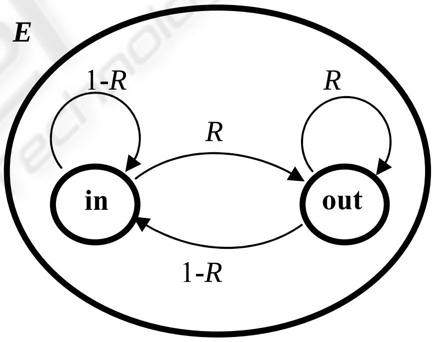



Figure 3: Internal model of the State ***E***.

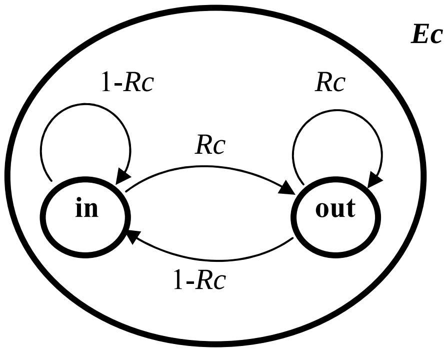



Figure 4: Internal model of the State ***Ec***.

## 2.3 Evaluation Function Using DP

Let the cost($n$) be $a\cdot n$ as a linear function for $n$, where $a$ is the loss caused by the unfitting in one time. And for simplicity, let $C$ and $A$ denote the Total_cost and cost($A$), respectively. Then, the evaluation function in (3) is denoted as the following equation (4).

$$C = A + a \cdot n \tag{4}$$

Using the concept of the DP (Dynamic Programming), we introduce a function $EC(n,N)$ to obtain the optimum number of times $n$ that minimizes the expectation value of the evaluation function of Equation (4).

Let $N$ be the optimum number. Let the function $EC(n,N)$ be the expectation value of the evaluation function at the time when the unfitting has occurred in continuing $n$ times, where $n$ is less than or equal to $N$, i.e., $0 \leqq n \leqq N$. Then the function $EC(n,N)$ is recursively defined as in the following equation.

$$(\text{if } n=N)\quad EC(N,N) = A + a\cdot N \tag{5}$$

$$(\text{if } n<N)\quad EC(n,N) = P(\bar{S}_{n+1} \mid S_n)\, a\cdot n + (1-P(\bar{S}_{n+1} \mid S_n))EC(n+1,N) \tag{6}$$

where $S_n$ is the state of unfitting in continuing $n$ times, $\bar{S}_{n+1}$ is the state of fitting for the ($n$+1)-th time observed data, and $P(\bar{S}_{n+1} \mid S_n)$ is the conditional probability that the state $\bar{S}_{n+1}$ occurs after the state $S_n$ occurred.

The first term in the right-hand side (RHS) of Equation (6) indicates the expectation value of the evaluation function at the time when the fitting happens for the ($n$+1)-th time observed data after the unfitting occurred for continuing $n$ times.

The second term in the RHS of Equation (6) indicates the expectation value of the evaluation function at the time when the unfitting happens for the ($n$+1)-th time observed data after the unfitting occurred for continuing $n$ times.

Note that, from the definition of the function $EC(n,N)$, the $N$ that minimizes $EC(0,N)$ is the same as $n$ that minimizes the expectation value of the evaluation function of (4).

## 2.4 Minimization

For the aforementioned $EC(0,N)$, the following theorem holds, and gives the $n$ that minimizes the expectation value of the evaluation function of (4).

Theorem.

The $N$ that minimizes $EC(0,N)$ is given as the largest number $n$ that satisfies the following Inequality (7).

$$a < (A+a)\cdot P(\bar{S}_n \mid S_{n-1}) \tag{7}$$

where the number $N$+1 can also be the optimum one that minimizes $EC(0,N)$, i.e., $EC(0,N) = EC(0,N+1)$, only if $a = (A+a)\cdot P(\bar{S}_{N+1} \mid S_N)$.

Proof (Outline).

Since the strict detailed proof needs many pages, we present the outline of the proof for the Theorem.

In order to prove this Theorem, we derive a contradiction with two assumptions under a premise as follows.

Premise: a number $N'$ is the largest number $n$ that satisfies the Inequality (7).

Assumption 1: There exists a number $N''$ such that $N'' < N'$ and $EC(0,N'') < EC(0,N')$

Assumption 2: There exists a number $N''$ such that $N' < N''$ and $EC(0,N') > EC(0,N'')$.

We can derive the above contradiction by three steps, as described below. At Step 1, we prove the following fundamental lemmas: Lemma 1-1 and Lemma 1-2. At Step 2, two lemmas, Lemma 2-1, and Lemma 2-2, are proved.

Using those lemmas, we can show that the above Assumption 1 contradicts the Premise. Similarly, at Step 3, it is proved that the Assumption 2 contradicts the Premise, using two lemmas: Lemma 3-1 and Lemma 3-2.

### (A) Lemmas in Step1

Lemma 1-1:

Let $E_{cn}$ be the event that the structural change occurs once during the period of observation in continuing $n$ times. Let $P(E_{cn} \mid S_n)$ be the conditional probability that the $E_{cn}$ occurs under the condition that failing occurs in continuing $n$ times. Then, $P(E_{cn} \mid S_n)$ is an increase function for $n$.

Lemma 1-2:

The conditional probability $P(\bar{S}_{n+1} \mid S_n)$ is a decrease function for $n$.

Those Lemmas are strictly proved subsequently in the Appendix.

### (B) Lemmas in Step2

Lemma 2-1:

If $N'' < N'$, then $EC(N'',N') < EC(N'',N'')$.

Lemma 2-2:

If $N'' < N'$, then, for $m$ ($0 < m \le N''$),

$$EC(N''-m, N') < EC(N''-m, N'')$$

By putting $m = N''$ in the Lemma 2-2, we have $EC(0, N') < EC(0, N'')$ in case of $N'' < N'$. This inequality contradicts the Assumption 1: There exists a number $N''$ such that $N'' < N'$ and $EC(0, N'') < EC(0, N')$.

**(C) Lemmas in Step3**

Lemma 3-1:

If $N' < N''$, then $EC(N', N'') \ge EC(N', N')$

where the equality holds only if $N'' = N'+1$ and $a = (A+a) \cdot P(\bar{S}_{N'+1} | S_{N'})$

Lemma 3-2:

If $N' < N''$, then, for $m$ ($0 < m \le N'$),

$$EC(N'-m, N'') \ge EC(N'-m, N'),$$

where the equality holds only if $N'' = N'+1$ and $a = (A+a) \cdot P(\bar{S}_{N'+1} | S_{N'})$.

By putting $m = N'$ in the Lemma 3-2, we have $EC(0, N'') \ge EC(0, N')$ in case of $N' < N''$. This contradicts the Assumption 2: There exists a number $N''$ such that $N' < N''$ and $EC(0, N') > EC(0, N'')$.

After all, $EC(0, N') \le EC(0, N'')$ ($N'' < N'$ or $N' < N''$), where the equality holds only if $N'' = N'+1$ and $a = (A+a) \cdot P(\bar{S}_{N'+1} | S_{N'})$. It means that $N'$ minimizes $EC(0, N)$. And, when $a = (A+a) \cdot P(\bar{S}_{N'+1} | S_{N'})$, $N'+1$ also minimizes $EC(0, N)$, i.e., $EC(0, N') = EC(0, N'+1)$.

This completes the proof of the aforementioned Theorem.

# 3 EXPERIMENTATION

## 3.1 Feature of Evaluation Function

We have experimented the proposed method, and evaluated the feature of the evaluation function, using the probability of the structural change occurrence λ and each constant of the Equation (4), i.e. $A$ and $a$, as parameters.

First, by numerical computing, we show the decreasing situation of the probability $P(\bar{S}_n | S_{n-1})$ for $n$ in Figure 5. In this case, the probability approaches to 0.05 (5 %) by letting $n$ become greater. It meets to the aforementioned Lemma 1-2.

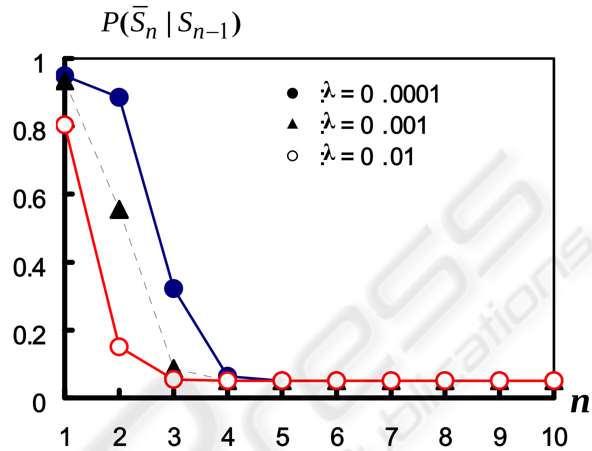



Figure 5: The probability $P(\bar{S}_n | S_{n-1})$ for three kinds of λ (occurrence probability of structural change) in the case of $Rc$=0.95

That is, we have examined the relation between the ratio of $(a+A)/a$ and the optimum number of times n (that is the same as aforementioned $N$), by varying the probability λ .

**Experimental condition:**

(i) Structural change probability λ : (three types) 0.1, 0.05, and 0.01.

(ii) $(a+A)/a$ : 1.5 ~ 10.0.

(iii)Tolerance of prediction: 2σ of the distribution on error ε.

(iv)Unfitting probability when the structure is unchanged: $R$=0.1.

(v)Unfitting probability when the structure has changed: $Rc$=0.9.

**Result:**

The result is shown in Figure 6, where horizontal axis is $(a+A)/a$ and vertical axis is $n$. We can see that the tendency meets our intuition, as follows.

(i) The optimum number of times $n$ tends to be larger when the action cost A after the CPD is bigger than the unfitting cost $a$. That is, the $n$ grows in the case of $A$ > $a$, because the action cost A after the change detection becomes dominant over the loss cost $a$ by prediction error, and the $n$ decreases in the case of $A$ < $a$ for the reverse reason.

(ii) The $n$ for the CPD increases when the probability λ of the structural change occurrence becomes smaller.

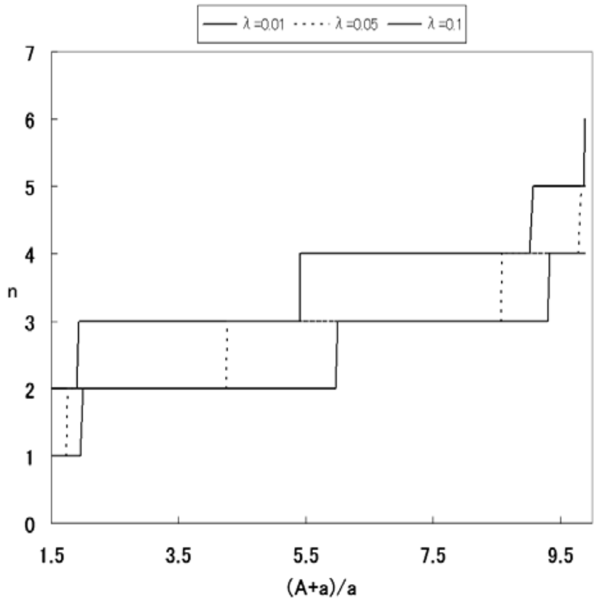



Figure 6: Relation between the ratio $(a+A)/a$ and the optimum number of times $n$

## 3.2 Application to Time Series Data

Since the tolerant error in the proposed method is decided based on the residual sample distribution when the prediction expression is estimated, the accuracy of CPD depends on the accuracy of prediction (or the prediction model). We examine the fact using the time series data shown in Figure 1.

**Outline of experimentation:**

(i)Generate the time series data (Figure 1) based on the Equation (1) as aforementioned in the Section 2.1, by making normal random numbers of average 0 and variance 1.0 for ε, and by setting $\beta_1$=0.2, β0=1 for the time t=1,2,…,70, and $\beta_1$=0.8, β0= - 41 for the time after t=71.

(ii)Make prediction expression, using a sequence of data at the time t=1,…,k from the above generated time series.

(iii)Decide the tolerant error interval.

(iv)Based on the proposed method, measure the number of times when the observed data goes out from the tolerance zone (or tolerant error interval) for observation data after the time at k+1, and detect the structural change point.

(v)Perform the above things repeatedly by M times, and calculate the average of the structural change point.

**Experimental condition:**

(i)Tolerant error interval: ± 2σ of the distribution on error ε.

(ii)The number of data for the decision of prediction expression: k=20, 40 (2 types).

(iii)Parameter value of the evaluation function: λ =0.01, and $(a+A)/a$ is changed in a range of 1.5 ~ 10.0.

(iv)Repeating times: M=100.

**Result:**

The result is illustrated in Figure 7, where horizontal axis shows $(a+A)/a$ and vertical axis shows the detected change point n that is the average of 100 times computation.

Although the detection of the change point depends on the value of $(a+A)/a$, it is expected that the change point will be detected around the time at t=70, because the structure of the time series is changed at t=70. We have verified that the result meets our intuition very well as follows.

(A) In case of k=20, because the number of the data for the prediction expression is less than the case of k=40, the prediction accuracy is considered to be so much worse. Therefore, the unfitting frequency increases and the change point tends to be detected early.

(B) In case of k=40, the change is detected within the time at t=70 ~ 80. We consider that the proposed method has appropriately detected the change point.

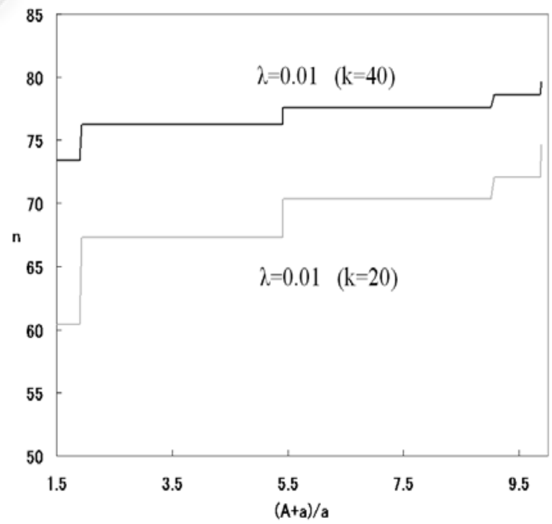



Figure 7: Detected change point $n$ for the time series in Figure 1

## 4 CONCLUSION

We have proposed a sequential processing method for structural change detection of time series data, which we formulated as an optimal stopping problem with a cost evaluation function. We have presented the algorithm for the optimum solution, and have shown the correctness by proving a theorem. The proposed method is effective in the sense as follows.

1. Differently from the Chow test, it does not need to set the change point in a priori.
2. Unlike the Bays' method, it does not need to give the generation distribution of time series data.
3. It can quickly detect the structural change point by the sequential processing.
4. It minimizes the evaluation function that sums up the loss involved with prediction error and action cost to be taken after the change detection.
5. It can be applied to any prediction model.

Moreover, we have shown some numerical experimentation results, where the resultant situations by obtaining optimum solutions well meet our intuition and the change point of artificially generated time series data.

## APPENDIX: PROOF OF LEMMA IN THE STEP 1

Lemma 1-1.
The conditional probability $P(E_{cn} \mid S_n)$ is an increase function for *n*.

Proof. Based on the model (see Figure 2-4), the event $E_{cn}$ is given in (8).

$$E_{cn} = \bigcup_{i=0}^{n-1} \left( E^i \cap E_c^{n-i} \right) \tag{8}$$

where $E$ is the event that there is no structural change, $E_c$ is the event that the structural change occurred, and $E^n$ is defined as $E^n = \bigcap_{i=1}^{n} E^i$ .

The probability of the event $E_{cn}$ defined in (8) is given as follows.

$$P(E_{cn}) = P\left( \bigcup_{i=0}^{n-1} \left( E^i \cap E_c^{n-i} \right) \right) = \sum_{i-0}^{n-1} P\left( E^i \cap E_c^{n-i} \right)$$

$$= \sum_{i=0}^{n-1} (1-\lambda)^i \lambda \tag{9}$$

Then the joint event between $E_{cn}$ and $S_n$, and the probability are given by (10) and (11), respectively.

$$S_n \cap E_{cn} = S_n \cap \left( \bigcup_{i=0}^{n-1} \left( E^i \cap E_c^{n-i} \right) \right) \tag{10}$$

$$P\left( S_n \cap E_{cn} \right) = \sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i} \tag{11}$$

Therefore, using (9) and (11), we have

$$P(S_n \mid E_{cn}) = \frac{P(S_n \cap E_{cn})}{P(E_{cn})}$$

$$= \frac{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}{\sum_{i=0}^{n-1} (1-\lambda)^i \lambda} \tag{12}$$

According to the Bayes' theorem, the posterior probability $P(E_{cn} \mid S_n)$ is given by the following (13).

$$P(E_{cn} \mid S_n) = \frac{P(S_n \cap E_{cn})}{P(S_n \cap E_{cn}) + P(S_n \cap E^n)}$$

$$= \frac{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i} + (1-\lambda)^n R^n}$$

$$= \frac{1}{1 + \dfrac{(1-\lambda)^n R^n}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}}$$

$$= \frac{1}{1 + D(n)} \tag{13}$$

where $D(n) = \dfrac{(1-\lambda)^n R^n}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}$ .

The $D(n)$ is also expressed as the following (14).

$$D(n)=\frac{(1-\lambda)^n\left(\frac{R}{R_c}\right)^n}{\lambda\sum_{i=0}^{n-1}(1-\lambda)^i\left(\frac{R}{R_c}\right)^i}=\frac{X^n}{\lambda\sum_{i=0}^{n-1}X^i}$$

$$=\frac{1}{\lambda\left(\frac{1}{X^n}+\frac{1}{X^{n-1}}+\cdots+\frac{1}{X}\right)} \qquad (14)$$

where $X=(1-\lambda)\frac{R}{R_c}$.

Since $0\le\lambda<1$, $0<1-\lambda\le 1$, and $R_c>R$, then $0 < X < 1$. So, the $D(n)$ becomes a monotonous decrease for $n$. Therefore, the probability $P(E_{cn}|S_n)$ of (13) is a monotonous increase function for $n$. Lemma 1-1 is proved.

Remark: Lemma 1-1 indicates that, if the number of times of the unfitting $n$ increases, the probability that the structural change has occurred increases. This meets our intuition clearly.

Lemma 1-2.
The conditional probability $P(\bar{S}_{n+1}|S_n)$ is a decrease function for $n$.

Proof. Based on the model in Fig. 3, we have

$$P(\bar{S}_{n+1}|S_n)=(1-R)(1-P(E_{cn}|S_n)) +(1-R_c)P(E_{cn}|S_n) \qquad (15)$$

The first term in the RHS of (15) shows the probability that the fitting occurs for the ($n$+1)-th time observed data when the structure is unchanged. The second term shows the probability that the fitting occurs for the ($n$+1)-th time observed data when the structure changed.

From (15), we have

$$P(\bar{S}_{n+1}|S_n)=1-R+P(E_{cn}|S_n)(R-R_c) \qquad (16)$$

By Lemma 1-2, $P(E_{cn}|S_n)$ is an increase function, and $R<R_c$, therefore, $P(\bar{S}_{n+1}|S_n)$ is a decrease function for $n$. Lemma 1-2 is proved.

Remark: Lemma 1-2 indicates that, if the number of times of continuous unfitting increases, the probability of the fitting for the next observed data after those continuous unfitting decreases. This is intuitively clear, because, by Lemma 1-1, the probability of the structural change increases if the number of times of the continuous unfitting increases.

# REFERENCES


R.Jana and S.Dey, 2000. Change detection in Teletraffic Models. IEEE Trans. Signal Processing, vol.48, No.3, pp.846-853, March 2000.

R.M.Gagliardi and I.S.Reed, 1965. On the Sequential detection of Emerging Targets. IEEE Trans. Information Theory, vol.6, pp.260-263, April 1965.

R.J.Di Francesco, 1990. Real-Time Speech segmentation Using Pitch and Convexity Jump Models. IEEE Trans. Acoustics, Speech, Signal Processing, vol.38, pp.741-748, May 1990.

A.S.Willsky, 1996. A survey of design methods for failure detection in dynamic systems. Automatica, Vol.12. pp. 601-611, Sept. 1996.

D.Kauame and J.P.remenieras et al., 1996. Effect of The Compensation In Abrupt Model Change Detection Problem. *Proc. of IEEE Digital Signal Processing Workshop, pp.311-314, Sept. 1996.*

Chow,G.C., 1960. Tests of Equality Between Sets of Coefficients in Two Linear Regressions. Econometrica, Vol.28, No.3, pp.591-605, July 1960.

S.MacDougall, A.K.Nandi and R.Chapman, 1998. Multisolution and hybrid Bayesian algorithms for automatic detection of change points. *Proc. of IEEE Visual Image Signal Processing, vol.145, No.4, pp.280-286, August 1998.*

V.V.Veeravalli and A.G.Tartakovsky, 2002. Asymptotics of quickest change detection procedures under a Bayesian criterion. *ITW2002, pp.100-103, Bangalore, India, Oct. 2002.*

E.S.Page, 1954. Continuous inspection schemes. Biometrika, Vol. 41, pp.100-115, 1954.

C.Han,P.k.Willet and D.A.Abraham, 1999. Some methods to evaluate the performance of Page's test as used to detect transient signals. IEEE Trans. Signal processing, vol.47, No.8, pp.2112-2127, August 1999.

S.D.Blostein, 1991. Quickest detection of a time-varying change in distribution. IEEE Trans. Information Theory, vol.37, No.4, pp.1116-1122, July 1991 .

Y.Liu, S.D.Blostein, 1994. Quickest detection of an abrupt change in a random sequence with finite change-time. IEEE Trans. Information Theory, vol.40, No.6, pp.1985-1993, Nov. 1994.

M.Basseville and I.V.Nikiforov, 1993. *Detection and Application,* Prentice Hall, Inc., April 1993.

M.Basseville, 1988.Detecting Changes in signals and systems– A survey. Automatica,Vol.24, pp.309-326, 1988.

N.R.Draper, H.Smith, 1996. *Applied Regression Analysis,* John Wiley & Sons Inc, 1966.