The Sixteenth International Symposium on Artificial Life and Robotics 2011 (AROB 16th '11),
B-Con Plaza, Beppu,Oita, Japan, January 27-29, 2011

# DP method for Structural Change Detection as Optimal Stopping --- Verification and Extension ----

*Tetsuo Hattori, *Katsunori Takeda, **Hiromichi Kawano, *** Tetsuya Izumi

*Graduate School of Engineering, Kagawa University / 2217-20 Hayashi, Takamatsu City, Kagawa 761-0396, Japan
** NTT Advanced Technology / Musashino-shi Nakamachi 19-18, Tokyo 180-0006, Japan
*** Micro TechnicaCo., Ltd./ Yamagami BLD. 3-12-2 Higashi-ikebukuro, Toshima-ku, Tokyo 170-0013, Japan
(Tel : 81087-864-2221; Fax : 81-087-864-2262)
(hattori@eng.kagawa-u.ac.jp)

***Abstract*:** Previously, we have formulated the structural change detection method in time series as an Optimal Stopping Problem with an action cost, using the concept of DP (Dynamic Programming). Then we have proved a theorem that the solution satisfies an inequality. In this paper, we verify the solution by numerical computation and gives the extension of the method by clarifying the notions of estimated structural change point and detection time point that the structure has changed so far.

***Keywords*:** time series, structural change, optimal stopping problem

## I. INTRODUCTION

For ongoing time series analysis, three stages are considered: prediction model construction, structural change detection (and/or disparity detection between the model and observing data), and renewal of prediction model. Especially in the second stage, it is important to detect the change point as quickly and also correctly as possible, in order to renew the accurate prediction model as soon as possible after the detection.

As the structural change detection, or change point detection (CPD), some methods have been proposed [1]-[4]. The standard well known method is Chow Test that is used in econometrics [2]. It does a statistical test by setting the hypothesis that the change has occurred at time t for all of data obtained so far.

Previously, we have formulated the structural change detection method in time series as an Optimal Stopping Problem with an action cost, using the concept of DP (Dynamic Programming) [5],[6]. Moreover, we have shown a theorem [6]. This paper presents the verification of the theorem and shows the extension.

## II. OPTIMAL STOPPING DP METHOD

### 1. Formulation ([5],[6])

According to the previously presented description ([5],[6]), we formulate the DP method for the change point detection problem (CPD), using 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 $y_t$, $\beta_1$, $\beta_0$, $\varepsilon$ mean the function value, two constant coefficients, and error term, respectively.

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

The error term $\varepsilon$ is given as a random variable of the normal distribution of variance $\sigma^2$ and average of 0, i.e., $\varepsilon \sim N(0, \sigma^2)$. For a time series data based on the equation (1), we think of two situations: one is the situation that the observed data goes out from the tolerance zone that means missing the range of $\pm 2\sigma$ from the predicted value. And, another is the situation that the observed data goes in the zone. We call the former situation "failing" (or "Out") and the latter "hitting" (or "In"). We assume that the structure changes when the failing occurs for continuing *N* times.

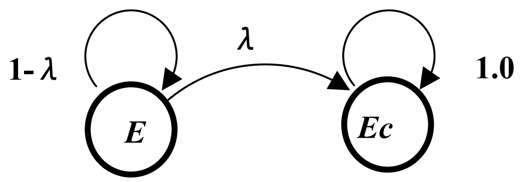
The evaluation function is given by (2) as the sum of two kinds of cost: the damage caused by the failing (i.e., failing loss) and action cost to be taken after the change detection.

$$\text{Total\_cost} = \text{cost}(A) + \text{cost}(n) \quad (2)$$

where cost($n$) is the sum of the loss by continuing $n$ times failing before the structural change detection, and cost($A$) denotes the cost involved by the action after the change detection. Then we have to find the number of $N$ that minimizes the expectation value of Total_cost, assuming that the structural change occurs randomly.

### 2. Structural change model ([5],[6])

We can 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 Fig. 1.




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

Fig.1. Structural change model.

Moreover, we introduce a more detailed model. Let *R* be the probability of the failing when the structure is unchanged. Let *R*c be the probability of the failing when the structure change occurred. We consider that *R*c is greater than *R*, i.e., *R*c>*R*. The detailed model for the State *E*c and *E* are illustrated as similar probabilistic finite state automatons in Fig.2 and Fig.3, respectively.

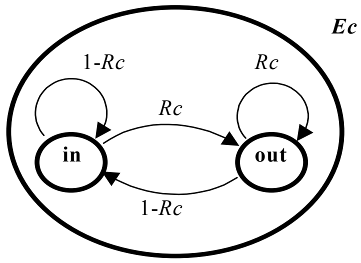



Fig.2. Internal model of the State *E*.

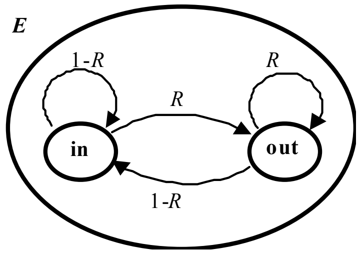



Fig.3. Internal model of the State *E*c.

### 3. Definition ([5],[6])

Let the cost(*n*) be $a \cdot n$ as a linear function for *n*, where $a$ is the loss caused by the failing in one time. And for simplicity, let *T* and *A* denote the Total_cost and cost (*A*), respectively. Then, the evaluation function in (2) is denoted as the following equation (3).

$$T = A + a \cdot n \qquad (3)$$

We recursively define a function $ET(n, N)$ to obtain the optimum number of times *n* that minimizes the expectation value of the evaluation function of Equation (3), using the concept of DP (Dynamic Programming). 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 failing has occurred in continuing n times, where *n* is less than or equal to *N*, i.e., $0 \le n \le N$.

Thus the function is recursively defined as follows.

$$(\text{if } n = N) \quad ET(n, N) = A + a \cdot N \qquad (4)$$

$$(\text{if } n < N) \quad ET(n, N) = P(\overline{S}_{n+1} \mid S^n) \cdot a \cdot n + (1 - P(\overline{S}_{n+1} \mid S^n)) ET(n+1, N) \qquad (5)$$

where $S^n$ means the state of failing in continuing *n* times, $\overline{S}_{n+1}$ the state of hitting at the (*n*+1) th observed data, and $P(\overline{S}_{n+1} \mid S^n)$ means the conditional probability that the state $\overline{S}_{n+1}$ occurs after the state $S^n$.

The first term in the right-hand side (RHS) of the equation (4) indicates the expectation value of the evaluation function at the time when hitting happens at the (*n*+1)th data after the continuing *n* times failing. The second term in the RHS of the equation (5) indicates the expectation value of the evaluation function for the time when failing happens at the (*n*+1)th data after continuing *n* times failing.

Then, from the definition of the function $ET(n, N)$, the goal is to find the *N* that minimizes $ET(0, N)$, because the *N* is the same as *n* that minimizes the expectation value of the evaluation function of (4).

### 4. Minimization of the evaluation function

The analytical solution *N* that minimizes $EC(0, N)$ can be deduced. The strict proof needs many pages, then we show numerical solution.

The function $ET(0, N)$ is defined by recursive expressions (4) and (5), then $ET(0, N)$ can be computed by recursively. In the process of this computation, $P(\overline{S}_{n+1} \mid S^n)$ can be calculated as follows.

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 $E_{cn}$ happens under the condition that failing has already occurred for continuing $n$ times. Based on the model in Fig.1-3,

$$P(\bar{S}_{n+1} \mid S^n) = (1-R)(1-P(E_{cn} \mid S^n)) + (1-R_c)P(E_{cn} \mid S^n) \quad (6)$$

Subsequently, we show the Lemma 1 and Lemma 2.

**Lemma 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 $E_{cn}$ happens under the condition that failing ("Out") occurs in continuing $n$ times. $P(E_{cn} \mid S^n)$ is an increase function for $n$.

Proof: Let $E$ be the event that there is no structural change. According to the Bayes' theorem, we have

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

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

$$= \frac{1}{1 + \dfrac{(1-\lambda)^n R^n}{\sum_{i=0}^{n-1} (1-\lambda)^i R^i \lambda R_c^{n-i}}} = \frac{1}{1 + D(n)} \quad (7)$$

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 follows.

$$D(n) = \frac{(1-\lambda)^n \left(\frac{R}{R_c}\right)^n}{\lambda \sum_{i=0}^{n-1} (1-\lambda)^i \left(\frac{R}{R_c}\right)^i} = \frac{\mathbf{X}^n}{\lambda \sum_{i=0}^{n-1} \mathbf{X}^i} \quad (8)$$

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

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 $P(E_{cn} \mid S^n)$ is a monotonous increase function. This means that, if the number of the failing times $n$ increases, the probability that the structural change has occurred increases. This meets our intuition clearly.

**Lemma 2**: The conditional probability $P(\bar{S}_{n+1} \mid S^n)$ is a decrease function for $n$.

Proof: Notice the equation (6).

$$P(\bar{S}_{n+1} \mid S^n) = (1-R)(1-P(E_{cn} \mid S^n)) + (1-R_c)P(E_{cn} \mid S^n)$$

The first term in the RHS of (6) shows the probability that the hitting ("In") occurs for the ($n$+1)-th time observed data when the structure is unchanged. The second term shows the probability that the hitting occurs for the ($n$+1)-th time observed data when the structure changed.

From the equation (6), we have

$$P(\bar{S}_{n+1} \mid S^n) = 1 - R + P(E_{cn} \mid S^n)(R - R_c) \quad (9)$$

By the aforementioned Lemma 1, $P(E_{cn} \mid S^n)$ is an increase function, and $R < R_c$, therefore, $P(\bar{S}_{n+1} \mid S^n)$ is a decrease function for $n$.

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

By using the above Lemma 1 and 2, and the reduction to absurdity, the following theorem holds [6], that gives the $n$ that minimizes the expectation value $ET(0,N)$.

**Theorem [6].**

The $N$ that minimizes $ET(0,N)$ is given as the largest number $n$ that satisfies the following Inequality (10).

$$a < (A+a) \cdot P(\bar{S}_n \mid S^{n-1}) \quad (10)$$

where the number $N$+1 can also be the optimum one that minimizes $ET(0,N)$, i.e., $ET(0,N) = ET(0,N+1)$, only if

$$a = (A+a) \cdot P(\bar{S}_{N+1} \mid S^N)$$

## III. VERIFICATION AND EXTENSION

### 1. Verification of the Theorem by numerical computation

By numerical computation, we evaluate the $ET(0,N)$ and the Inequality(10) under the same conditions. Fig.4 (a) shows that the relation between the expectation $ET(0,N)$ and $N$, and Fig.4(b) shows the value $a-(A+a)\cdot P(\bar{S}_{N+1} \mid S^N)$. We can easily verify that the number 3 minimize $ET(0,N)$, and at the same time, is the largest number satisfying Inequality (10). Fig.5 also shows that the relation between the expectation $ET(0,N)$ and $N$ by varying the $A/a$ and fixed $\boldsymbol{\lambda}$. It implies that as the action cost *A* is greater, the *N* that minimizes *ET*(*0,N*) becomes greater.

**2. Extension of the change point notion**

We separate the notion of CPD into two that one is the time point when the change has been detected so far at the observing time and another is the estimated change point just the time when the change has occurred.

Then we define that, if the aforementioned detected change point is tc, then estimated change point exits within a section [tc-*N*, tc]. We have also verified by experimentation for ongoing real time series data, that the extended definition works very well.

## IV. CONCLUSION

We have verified and shown by numerical computation, that the Theorem surely holds. And we also have proposed the extended notion of change point detection. We consider that the optimal stopping DP method and the extension will be promising.

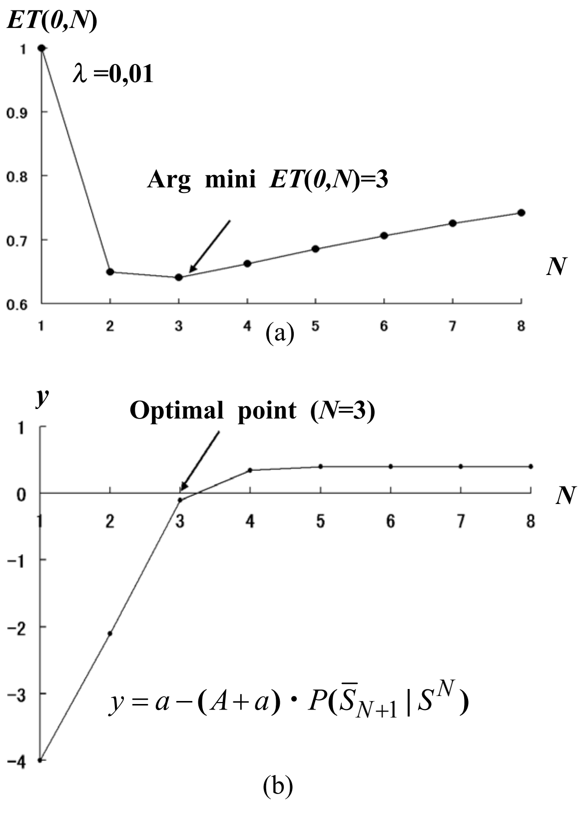



Fig.4. Evaluation of the expectation ***ET*(*0,N*)** and Inequality $a<(A+a)\cdot P(\bar{S}_n|S^{n-1})$ appeared in the Theorem [6].

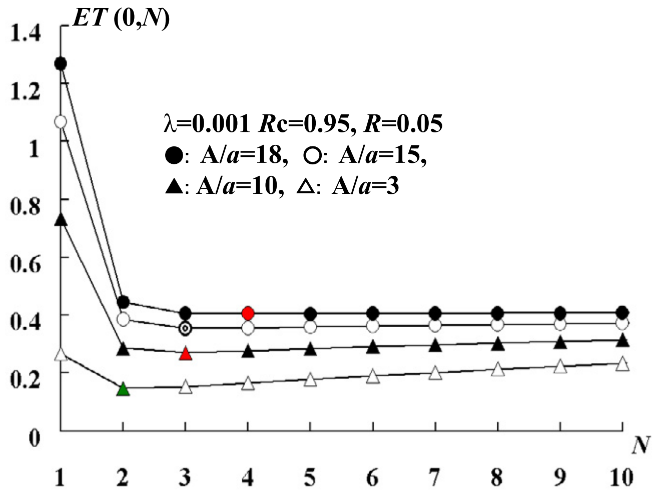



Fig.5. Relation between the expectation ***ET*(*0,N*)** and ***N*** with $R_c$ and λ fixed and varying *A*/*a*(18, 15, 10, 3). Optimal *N* is 4,3,3,2, respectively, depending the order of the above ratio of A/*a*.

## REFERENCES

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

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

[3] 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.

[4]E.S.Page (1954), Continuous inspection schemes, Biometrika, Vol.41, pp.100-115.

[5] Tetsuo Hattori, Katsunori Takeda, Izumi Tetsuya, Hiromichi Kawano (2010): "Early Structural Change Detection as an Optimal Stopping Problem (I) --- Formulation Using Dynamic Programming with Action Cost ----", Proc. of the 15th International Symposium on Artificial Life and Robotics (AROB15th'10), ISBN 978-4-9902880-4-4, pp.763-766.

[6] Hiromichi Kawano, Tetsuo Hattori, Katsunori Takeda, Izumi Tetsuya (2010): "Early Structural Change Detection as an Optimal Stopping Problem (II) --- Solution Theorem and its Proof Using Reduction to Absurdity ----", Proc. of the 15th International Symposium on Artificial Life and Robotics (AROB15th'10), ISBN 978-4-9902880-4-4, pp.767-772.