Degradation Detection of Wireless IP Links
Based on Local Stationary Binomial Distribution Models

Kazunori Matsumoto Shigeki Muramatsu Naomi Inoue
KDDI R&D Laboratories INC.
2-1-15 Ohara, Kamifukuoka, Saitama 356-8502, JAPAN
{matsu, mura, inoue}@kddilabs.jp

Hideki Mori
Mitaka Access Network Research Center
Telecommunications Advancement Organization of Japan
9-11-7 Shimorenjaku, Mitaka, Tokyo 181-8505, Japan
wireless@central.schoolnet.gr.jp

Abstract

A degradation detection problem of link quality in a long-distance 2.4 GHz wireless system is discussed. The time series to be monitored is periodic and non-stationary. The decision algorithm for degradation is difficult to define, and methods based on conventional traffic theory are not useful for IP link quality. Thus we should introduce some kind of intelligent data analysis techniques. The authors propose to apply an AI-based method which solves a similar problem in a commercial switching telephone and ISDN network. The method partitions a target time-series into local stationary segments. Optimization of partitioning is based on the minimal Akaike's Information Criterion principle. And the technique called Sequential Probability Ratio Test is also applied to make efficient decisions about degradation. Thus experiments to apply our proposed method to this domain are conducted with wireless systems at a real field. The result shows the AI-based method is also effective for the degradation detection of wireless IP links.

1 Introduction

A long-distance 2.4 GHz wireless system is one of promising approaches to connect distant IP networks. The system does not require any special licenses to operate, and its operation cost is less expensive than other wireless systems such as FWA (Fixed Wireless Access). However, there exists the risk that the signal of a wireless link is interfered by other 2.4 GHz wireless systems nearby, as other persons may deploy such systems and happen to use the same radio channel. If the operator detects the long-term degradation of wireless link quality, he can change the radio channel to reduce the interference. Therefore the consideration about monitoring link quality and its on-line analysis is very important for long-distance 2.4 GHz wireless systems.

The decision algorithm for this degradation problem is difficult to define before the system operation, because we don’t know the wireless environment of the system at first. And more, methods based on conventional traffic theory are not useful for IP link quality which suddenly changes. Observation of re-transmission of packets in a physical layer on a link is very useful to measure the link quality, like the observation of Carrier-to-Interference Ratio (CIR). In this paper, we deal with the numbers of attempted packets to be transmitted and to be re-transmitted. Re-transmission rate is defined as the rate of these two values. The main reason to prefer such parameters rather than CIR is facility to use stochastic models to detect the degradation of the quality. There seems to be no practical models with a few parameters to express the time-series of observed CIR. On the other hand, the degradation detection method to use local stationary binomial models was applied successfully to a public switching telephone and ISDN networks[1][2]. And more, the above two numbers about re-transmission are easily obtained and are useful to decide network-management actions based on traffic volume on links.

For the proof of long-distance 2.4 GHz wireless system’s capability, IP networks connected with CFO-SS techniques[3] are deployed at several different regions. The main research interest of this paper is to investigate the formally developed detection method’s feasibility to this domain, i.e. the degradation detection in wireless links, by
using such real wireless systems.

In the next section, the authors show the problems to be solved for the above mentioned wireless system and propose the solution to apply the conventional method. The section 3 describes the conventional degradation detection techniques, especially how to partition a target time-series into local stationary segments. The section 4 describes how to apply Sequential Probability Ratio Test (SPRT)[6] for decision about degradation detection. The section 5 is the results of the investigation with real wireless systems. A optimized binomial distribution model for observed time-series of a link is compared with a normal (or Gaussian) distribution model. An experiment to partition observed time-series into local stationary segments is conducted. And degradation detection with SPRT is also evaluated. Finally our contributions are summarized in the section 6.

2 Problems and Proposal

Fig. 1 shows the re-transmission rate of packets in a physical layer on a certain wireless link whose two fixed end-points are about 4 kilometers distant. IP traffic on the link is generated by transmitting a 20 KB content hundred times at every five minutes. No other packets except this artificial traffic is transmitted on the link. In the figure, the averages of the rate at every hour are also shown. The averages are obtained from 12-days observation.

From this preliminary investigation, we found following problems to be solved.

- Although enough amount of packets are transferred at a constant pace, the re-transmission rate is not stable. It goes up and down rapidly in a day owing to background noises. In a general situation, the number of attempted packets to be transmitted changes all day. Thus the rate becomes more unstable in the period when the transmitted packets are few. To make the matter worse, lengths of transmitted packets are widely distributed. in a real situation The probability of fail to transmit a long packet is larger than that of a short packet. A simple degradation detection method based on only the re-transmission rate can’t dealt with these phenomena.

- The re-transmission rate is not a simple stochastic process because the average of the rate at one hour is changing. In the figure, we can observe that the rate of the day time is higher than that of night time.

- As known well, the nature of the IP traffic on links suddenly changes. Thus we need an efficient method to decide the degree of degradation with an on-line manner.

Similar problems are studied in the case of degradation detection for the switching networks[1]. The method used there also handle the real-time stream of established connections rates, which is a periodic and non-stationary time series. According to the experience of the switching networks, the solution for this problems of this paper are as follows:

- To make the better detection, we should handle the number of attempted packets to be transmitted and the that of packets to be re-transmitted at one time. By regarding these two values as the number of success occurrences and that of trials in Bernoulli trials respectively, we can introduce binomial distribution model with these two values.

- The entire time series is non-stationary, but by dividing a target time series into proper size segments, each segment can be regarded as stationary. Akaike’s Information Criterion[5] is useful for making partitions at proper positions in the time series. The value of AIC depends on both log-likelihood of a used model and the number of model parameters. We can select a good model by comparing AIC values, even when the numbers of model parameters are different.

- In Fig. 1, the average of the re-transmission rate depends on a time in a day. And the period of a repetition is 24 hours. Thus we can use the same parameter values for two different samples apart from 24 hours.

- The Sequential Probability Ratio Test developed by Wald[6] is suited for real-time decisions. It is popular in aerospace applications[7], but have not been used in this domain. The authors propose to apply this efficient stochastic test.

Takanami and Kitagawa[8] shows an application which generates proper models called locally stationary AR models based on the minimal AIC principle from a non-stationary time series. The method here also uses the same principle, but is extended to a periodic time-series. In our method, not only adjacent samples but also samples apart from repetition period are in the same stochastic model.

3 Optimization of Partitioning for Periodic Time Series

In this section, the authors introduce the formally developed method for the degradation detection in a commercial public switching telephone and ISDN network[1][2]. The formal description of a given observed time series and its partitioning of the method is described. The method
finds the most probable partitions which divides the original non-stationary periodic time-series into local stationary segments.

3.1 Definition of Partitioning

Observations are repeated with an equal time interval. Let

\[ S = (n_1, r_1), (n_2, r_2), \ldots, (n_M+1, r_{M+1}), (n_M+2, r_{M+2}), \ldots, \]
\[ (n_{(m-1)}+1, r_{(m-1)}+1), (n_{(m-1)}+2, r_{(m-1)}+2), \ldots, \]
\[ (n_m, r_m) \]

be a discrete time series, where \((n_t, r_t)\) denotes the pair that consists of the number of trials and that of success occurrences at time \(t\), \(M\) is the period of repetition, and \(m\) is the number of the periodic cycles.

Let \(\mathcal{P}^i = \{\pi_1^i, \pi_2^i, \ldots, \pi_{s(i)}^i\}\) be a set of partitions (called here partition model), where \(i\) is a unique identifier of the model and \(s(i)\) is the number of partitions in \(\mathcal{P}^i\).

Let \(S_1^i, S_2^i, \ldots, S_{s(i)}^i\) be the set of observed data divided by \(\mathcal{P}^i\) and \(F_j^i (0 < F_j^i < M)\) be the position of a partition \(\pi_j^i\), where \(S_j^i\) is given as follows.

\[ S_j^i = (n_{F_{j-1}^i}, r_{F_{j-1}^i}), \ldots, (n_{F_j^i-1}, r_{F_j^i-1}), \]
\[ (n_{M+F_{j-1}^i}, r_{M+F_{j-1}^i}), \ldots, (n_{M+F_j^i-1}, r_{M+F_j^i-1}) \]
\[ \ldots, \]
\[ (n_{(m-1)}+F_{j-1}^i, r_{(m-1)}+F_{j-1}^i), \ldots, (n_m, r_m) \]

\(S_j^i\) denotes the set of observed data which exists between \(\pi_{j-1}^i\) and \(\pi_j^i\). Note that \(S\) is non-stationary but the distribution of \(S_j^i\) is binomial under our assumption.

3.2 Partitions Selection Based on Minimal AIC Principle

Let \(P(r|n, p)\) be the probability that \(r\) successes occurs in \(n\) trials, each of which has probability \(p\) of success.

Log-likelihood of \(S_j^i\) is

\[ U(S_j^i|p) = \sum_{(n, r) \in S_j^i} \log P(r|n, p) \]
\[ = \sum_{(n, r) \in S_j^i} \log n C_r p^r (1-p)^{n-r} \]
\[ = \sum_{(n, r) \in S_j^i} \log n C_r \]
\[ + R_j^i \log p + (N_j^i - R_j^i) \log(1-p) \]

---

Figure 1. The rate of re-transmission on physical layer and the averages at every hour in a day.

Re-transmission Rate

![Graph showing re-transmission rate over time](image-url)

Observed Rate

1-Hour Average

Time of a Day

0:00 6:00 12:00 18:00

0.05 0.10 0.15
When $\hat{p}_j$ is the maximum likelihood estimation of $p$ for $S_j^p$, $\frac{\partial}{\partial p} u(S_j^p|\hat{p}_j) = 0$. Thus, $\hat{p}_j = \frac{R_j}{N_j}$. The maximum log-likelihood of $P^j$ is

$$MUL(P^j) = \sum_{j=1}^{n+1} u(S_j^p|\hat{p}_j)$$

$$= \sum_{j=1}^{n+1} \left\{ \sum_{(n,r) \in S_j^p} \log n_c C_r + R_j \log \frac{R_j}{N_j} + (N_j^p - R_j) \log \frac{N_j^p - R_j}{N_j} \right\}$$

$$= \sum_{l=1}^{mM} \log n_c C_{r_l} + \sum_{j=1}^{n+1} \left\{ R_j \log \frac{R_j}{N_j} + (N_j^p - R_j) \log \frac{N_j^p - R_j}{N_j} \right\}.$$ 

Log-likelihood is inappropriate for comparing models when the number of parameters of each model is different. Akaike information Criterion (AIC) [5] in the next equation is known to be a correct measure for a general model comparison. Let $\theta$, $LL(\theta)$, $|\theta|$ be the model to be compared, log-likelihood of the model $\theta$, and the number of free parameters in the model.

$$AIC(\theta) = -2 \times LL(\theta) + 2 \times |\theta|$$

Based on the minimum AIC principle, the model $\theta$ which has the minimum $AIC(\theta)$ is the best model.

Note that $\sum_{l=1}^{mM} \log n_c C_{r_l}$ in $MUL(P^j)$ appears in all partition models, and we can ignore it. The most probable partition model $P$ in \{P, P^2, \ldots \} is obtained by the following equations.

$$MUL'(P^j) = \sum_{j=1}^{n+1} \left\{ R_j \log \frac{R_j}{N_j} + (N_j^p - R_j) \log \frac{N_j^p - R_j}{N_j} \right\}$$

$$AIC(P^j) = -2 \times MUL'(P^j) + 2 \times \{n(i) + 1\}$$

$$P = \arg \min_\theta AIC'(P^j)$$

Cost calculation of each partitioning may be executed one by one. To make more efficient computation, the authors suggest the re-use of the value related to the segment whose log-likelihood has been calculated and stored. This is the reason why the same segment often appears in a different partitioning.

### 4 Degradation Detection by SPRT

As mentioned before, a technique called SPRT (Sequential Probability Ratio Test) is used for the sake of efficient decision. In this section, the conventional scheme of SPRT is explained first and application of SPRT to our degradation detection problem is described.

#### 4.1 Scheme of SPRT

SPRT is a well-known method for testing a hypothesis against an alternate one. Let $v_1, v_2, \ldots$ denotes the recent successive samples in a given time series vector. The basis for the SPRT is the recursive calculation of the logarithm of the likelihood ratio (LLR) function of the normal model and an alternate model with recent $q$ samples

$$LLR(q) = \log \frac{p_q(v_1, v_2, \ldots, v_q|H_1)}{p_q(v_1, v_2, \ldots, v_q|H_0)},$$

where $p_q(v_1, \ldots, v_q|H_1)$ is the probability density function when the process is degraded (hypothesis $H_1$ is true), and $p_q(v_1, \ldots, v_q|H_0)$ is also the probability density function when the process is normal ($H_0$ is true). After assuming that $v_i$ and $v_j$ are independent, the above formula becomes additive and the LLR function is computed recursively

$$LLR(q) = LLR(q-1) + \log \frac{p(v_q|H_1)}{p(v_q|H_0)}$$

where $P(v_q|H_1)$ and $P(v_q|H_0)$ are probability density function yielding $v_q$ when the process is degraded or normal, respectively. $LLR(q)$ is compared to two limits (degraded/normal), with a gray range between them. When it lies in the range, there is no decision. These two limits are derived from the allowable false alarm rate and the allowable missed alarm rate chosen by users.

### 4.2 SPRT for Degradation Detection in Binomial Models

SPRT requires two functions. One function evaluates the probability that the process is degraded and the other evaluates the probability that the process is normal. In the case of our degradation problem, LLR function is defined as below:

Let $\hat{p}$ be the maximum likelihood estimation of the segment in the most probable partition model to which the current observation belongs, let $\epsilon (1 < \epsilon)$ be a weighting parameter which is introduced to control the sensitivity of degradation, and let $Prob(n, r|H_1)$ and $Prob(n, r|H_0)$ be the probability when the process is degraded and normal respectively.

$$LLR(n, r) = \log \frac{Prob(n, r|H_1)}{Prob(n, r|H_0)}$$
\[
\begin{align*}
&= \log \frac{P(r|n, p_w)}{P(r|n, p_g)} \\
&= \log \frac{n \sum_{r} p_w^r (1 - p_w)^{n-r}}{n \sum_{r} p_g^r (1 - p_g)^{n-r}} \\
&= r \log \frac{p_w}{p_g} + (n - r) \log \frac{1 - p_w}{1 - p_g},
\end{align*}
\]

where \( p_w \) is an expected success probability under the degraded mode, and \( p_g \) is an expected success probability under the normal mode. \( p_w \) and \( p_g \) are defined as follows.

\[ p_w = \begin{cases} 
\frac{r}{n} & \text{if } \hat{p}/e > r/n \\
\hat{p}/e & \text{otherwise}
\end{cases} \quad (1) \]

\[ p_g = \begin{cases} 
\frac{r}{n} & \text{if } \hat{p} < r/n \\
\hat{p} & \text{otherwise}
\end{cases} \quad (2) \]

The rough meaning of \( \text{Prob}(n, r|H_1) \) is the probability that the process is working under the condition that the success probability is lower than \( \hat{p}/e \). The larger \( e \) is, the smaller the number of detections becomes.

A long sequences of successive normal states are often observed in high quality networks. Such sequences are harmful for SPRT in the sense that the accumulation of LLR becomes too large. In order to solve this problem, Chien and Adams[9] propose that LLR be reset to zero when the accumulation reaches either of the limits. Uosaki[10] proposes another solution, a backward evaluation of LLR continuing backward until a decision is obtained.

By the above definitions, we can apply SPRT to our problem and make an efficient decision when new observed data is obtained.

5 Empirical Results by Using Real Wireless Systems

The aim of this research is to know the feasibility of the formerly developed degradation detection method. In this section, we report the experiments to apply the method to the degradation detection problem in long-distant 2.4 GHz wireless IP links.

5.1 Experimental Environment

Real wireless IP links connected with long-distance 2.4 GHz systems called CFO-SS[3] are used for our experiments. The wireless links are a part of an educational wireless network[4] deployed by Telecommunications Advancement Organization of Japan (TAO). This wireless network consists of several sites, which include the city hall of Ohta and sixteen schools within the range of twenty kilometers.

A CFO-SS system has an original MIB (Management Information Base) and is capable of providing the next two parameters by SNMP protocol:

- Accumulated number of transmitted packet
- Accumulated number of re-transmitted packet

By examining above two values periodically with SNMP, we can obtain the number of attempted packets to be transmitted in a specified period and the number of attempted packets to be re-transmitted in a specified period. In the case of our experiments, SNMP data is transmitted not on CFOs’ links but on other IP links. Thus we can access the CFO’s MIB even when CFO link is down.

Fig. 2 shows the wireless links used for the experiments. On each link, the artificial traffic is transmitted at constant pace through the period of experiments. We obtained two types of 36-days time series data on each link:

- the number of packets to be transmitted in 5 minutes (denoted as \( n \))
- the number of packets to be re-transmitted in 5 minutes (denoted as \( r \))

The former part of the time series including 24-days data is used as a training data, and the rest part including 12-days data is used as a test data.

5.2 Evaluation on Binomial Distributions

A simple and basic way to estimate the number of re-transmitted packets is to evaluate a residual error. The residual error is the difference between an observed value of re-transmitted packets and an expected value of re-transmitted
packets. The expected value is calculated from an observed value of transmitted packets with the theory of linear correlation. Fig.3 (a) shows the distribution of such residual errors under the condition that observed values of transmitted packets in 5 minutes (n) on a link between the city hall and Ohta Technical H.S. are in a specified range (1200 < n < 1300). According to the figure, the distribution of the residual does not fit normal distribution well.

In our approach, the relation between the number of transmitted packets and that of re-transmitted packets is assumed to be binomial. Fig.3 (b) shows the distribution of observed values of re-transmitted packets in 5 minutes on the same wireless link when observed value of transmitted packets in 5 minutes (n) is in the specific range (1200 < n < 1300). By the comparing Fig.3 (a) and (b), the assumption of binomial distribution is more probable.

From the theoretical view, we also investigate the maximum log-likelihood of two models for every link in the Fig.2. First is a model based on normal distribution to assume the linear correlation and the other is a model based on binomial distribution. In a table, maximum log-likelihood of the binomial model is always larger than that of the normal distribution model at any links. Thus we can conclude that the assumption of binomial distribution is probable.

5.3 Evaluation on Partitioning

It is very important to find good partitions in order to detect the degradation well. In the table 2, the maximum log-likelihood and AIC on the wireless link between the city hall and Tech. H.S. for several types of partitioning. The best partitioning selected by the minimal AIC principle from the candidates in the table is the partitioning which divides time series of a day into 8 segments. And the log-likelihood of test data for this selected partitioning is also the best in the test data. Please note that the AIC-based order of the partitioning type in the training data is the same order of log-likelihood in the test data. From this table, the selection of partitioning based on AIC goes well for this link.

5.4 Evaluation on SPRT

At the final stage of the evaluation, we apply two versions of SPRT to observed time series, i.e. Chen’s SPRT and Uosaki’s SPRT. The decisions of both SPRT for all links are almost the same. From the technical view, we may use any of these two versions.

Table 3 shows the ratio of degraded period decided by Uosaki’s SPRT with an optimized partitioning for each wireless link. There are three links which tend to be decided as degraded, i.e. links of Minami E.S., Kyuhaku Jr. H.S. and Joto Jr. H.S. In Table 2 the maximum log-likelihood of binomial models of these three links are worse than that of the other three links in test data. This means that degree of model fitting with binomial distribution affects the rate of degradation. Anyway the rate of degradation can be controlled by a parameter e, which defines the re-transmission rate at the abnormal state.

Fig. 4 is a scatter graph of the number of transmitted packets and that of re-transmitted for the test data of Ohta Tec. H.S., where every sample is designated whether degraded or not. The designation is made by Uosaki’s SPRT.
<table>
<thead>
<tr>
<th>Training Data</th>
<th>Log-likelihood</th>
<th>Number of segments</th>
<th>AIC</th>
<th>Test Data</th>
<th>Log-likelihood</th>
</tr>
</thead>
<tbody>
<tr>
<td>Partition by every hour</td>
<td>-16,127.1</td>
<td>24</td>
<td>32,302.2</td>
<td>-23,154.5</td>
<td></td>
</tr>
<tr>
<td>Partition by every 2-hours</td>
<td>-16,133.8</td>
<td>12</td>
<td>32,291.5</td>
<td>-23,145.2</td>
<td></td>
</tr>
<tr>
<td>Partition by every 3-hours</td>
<td>-16,133.6</td>
<td>8</td>
<td>32,283.2</td>
<td>-23,099.3</td>
<td></td>
</tr>
<tr>
<td>Partition by every 4-hours</td>
<td>-16,144.0</td>
<td>6</td>
<td>32,300.1</td>
<td>-23,147.1</td>
<td></td>
</tr>
<tr>
<td>Partition by every 6-Hours</td>
<td>-16,154.2</td>
<td>4</td>
<td>32,316.4</td>
<td>-23,315.5</td>
<td></td>
</tr>
<tr>
<td>Partition by every 8-Hours</td>
<td>-16,162.1</td>
<td>3</td>
<td>32,330.2</td>
<td>-23,177.6</td>
<td></td>
</tr>
<tr>
<td>Partition by every 12-Hours</td>
<td>-16,211.0</td>
<td>2</td>
<td>32,426.0</td>
<td>-23,274.8</td>
<td></td>
</tr>
<tr>
<td>No partition</td>
<td>-16,216.9</td>
<td>1</td>
<td>32,435.9</td>
<td>-23,286.3</td>
<td></td>
</tr>
</tbody>
</table>

Table 2. Maximum log-likelihood and AIC of different partitioning type.

- The partitioning selected by AIC with a training data is probable in a test data.
- By applying SPRT to the test data with the best partitioning, it is possible to make an appropriate decision about degradation when a binomial distribution’s assumption is satisfied. Even if the satisfaction of the assumption is weak, we can control the ratio of degraded period and normal period with a parameter \( \epsilon \).

### Acknowledgment

The authors would like to express their gratitude to Mr. T. Asami, Dr. Y. Matsushima, Dr. Hashimoto, Dr. T. Mizuike, and Dr. Shinonaga to of KDDI R&D Laboratories, and researchers of Telecommunications Advancement Organization of Japan (TAO) for their encouragement during this work.

### References


Table 3. Period of normal states and abnormal states decided by Uosaki’s SPRT.

<table>
<thead>
<tr>
<th>School</th>
<th>( \epsilon = 1.3 )</th>
<th>( \epsilon = 1.5 )</th>
<th>( \epsilon = 1.7 )</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>normal</td>
<td>abnormal</td>
<td>normal</td>
</tr>
<tr>
<td>Sawano E.S.</td>
<td>99.1%</td>
<td>0.9%</td>
<td>99.5%</td>
</tr>
<tr>
<td>Minami E. H.S.</td>
<td>89.7%</td>
<td>10.3%</td>
<td>94.7%</td>
</tr>
<tr>
<td>Higashi Jr. H.S.</td>
<td>97.7%</td>
<td>2.3%</td>
<td>98.2%</td>
</tr>
<tr>
<td>Kyuhaku Jr. H.S.</td>
<td>75.9%</td>
<td>24.1%</td>
<td>93.0%</td>
</tr>
<tr>
<td>Ohta Technical H.S.</td>
<td>98.0%</td>
<td>2.0%</td>
<td>98.7%</td>
</tr>
<tr>
<td>Joto Jr. H.S.</td>
<td>38.8%</td>
<td>61.2%</td>
<td>73.8%</td>
</tr>
</tbody>
</table>


