Journal of Theoretical and Applied Information Technology
15th December 2012. Vol. 46 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PROBABILITY DISTRIBUTION MODELING OF THE INTERFERENCE OF THE TRACTION CURRENT IN TRACK CIRCUITS

LINHAI ZHAO*, MAOJIAO LI

School of Electronic and Information Engineering, Beijing Jiaotong University, Beijing 100044, China

ABSTRACT

As a main interference to track circuit signal, the interference of the traction current(ITC) will directly influence the transmission quality of track circuit signal and even lead to signal distortion which would endanger the traffic safety. The ITC has a complicated formation mechanism and it is affected by factors of train compositions, load and running speed, so it is hardly possible to establish accurate mathematical analytic model of ITC. This paper has analyzed the probability distribution characteristics of the ITC extracted from real train operating data collected by the cab signal system. A method using Kolmogorov–Smirnov (K-S) test and least square approximation has been proposed to establish the probability distribution model of the ITC. Feasibility and applicability have been confirmed by examples. The probability distribution model of the ITC has a significant importance for analyzing track circuit signals especially for interference signal simulation.

Keywords: *Track Circuit, Interference of Traction Current, Probability Distribution, K-S Test*

1. INTRODUCTION

The operating efficiency and traffic safety of the train is guaranteed by the train control system (TCS). As a fundamental and key part of TCS, track circuits undertake the tasks of train detection and ground-to-train information transmission. In electric railway track circuit, regardless of the specific traction supply system, the two steel tracks are both transmission line of the track signal and return line for electric locomotive traction current back to electric substation. When the traction current in two steel tracks are unbalanced, an unignorable interference caused by the traction current would directly influence the transmission quality of track circuit signal and even lead to distortion of the signal read by the track circuit reader (TCR), thus the traffic safety is endangered.

As the quick development of the high-speed and heavy-haul railways, the high-power locomotives have been widely used and this would bring a bigger interference due to the higher traction current. Yet no thorough research has been done to analyze the interference of the traction current (ITC). Present studies mainly focused on the distribution model of the rail currents, mainly consisting of track circuit current and traction return current. A model of the distribution of the traction return current in AC and DC electric railway systems has been established by Andrea Mariscotti[1], based on which the traction return current in the rails could be simulated. Another model of the Auto Transformer (AT) electric traction systems for high speed railways has been established and the return currents at different points of the rail have been simulated [2]. A further model of the conducted interference prediction has been established based on the distribution model traction return current [3]. These researches have made great contributions to estimate the interference to the track circuit signals. However, even these models have considered a lot of factors, none of them has studied the characteristics of the interference caused by the traction current. Besides, it is hard to use these models directly to simulate the ITC in a specific railway.

This paper takes another way to analyze the ITC based on the train operating data. Firstly, the harmonics of ITC were extracted from a plenty of real train operating data collected by the rolling stock in railway, “Jing-Hu” high speed railway here in this paper has been taken as an example. Then Kolmogorov-Smirnov test and least square approximation algorithm have been used to establish the probability distribution models of the harmonics of the ITC. With these models, simulation signal of the ITC could be achieved.

Journal of Theoretical and Applied Information Technology
15th December 2012. Vol. 46 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

2. TRACK CIRCUIT AND THE INTERFERENCE OF THE TRACTION CURRENT

2.1 The Basic Structure and Operating Principle of the Track Circuit

As shown in Figure 1, for an idle track circuit with health state, the track circuit signal generated by the transmitter flows through the tuning area at sending end, rails and compensation capacitors, the tuning area at receiving end and finally into the receiver under the electrical isolation effect of tuning areas. According to the received train control information, the transmitter of the rear track circuit section in the running direction generates corresponding ‘idle’ signal to express the permission to enter this idle track circuit. When a train enters a track circuit, the track circuit signal will be shunted by the first wheels and axle of the train locomotive and flow back into the transmitter. As a result, the received signal in the receiver becomes lower than the given threshold and then the rear track circuit gives ‘occupied’ signal to express no entry. The train detection is realized in this way. Meanwhile the vehicle-mounted TCR equipment receives induced voltage through the electro-magnetic induction between TCR antenna and the rails so that the train control information is delivered from ground to train. After signal processing of such induced voltage in TCR, the train control information could be obtained and transmitted to Vital Computer (VC) for train speed control.

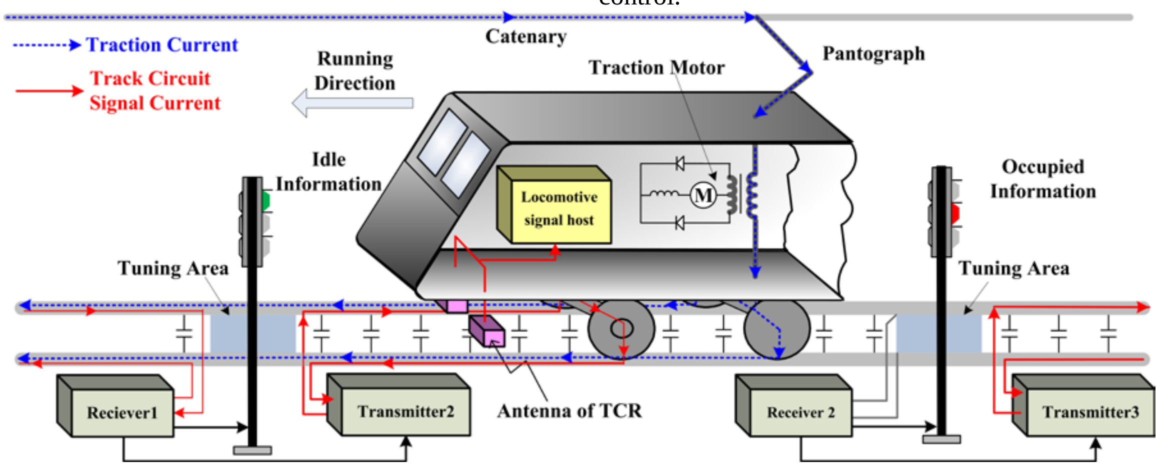

Figure 1: Diagram Of The Operating Principal Of Track Circuits And Traction Currents

2.2 The Formation Mechanism of the ITC

As shown in Figure 1, In electric railway track circuit, the power of the locomotive is supplied by the traction current which is transmitted from the electric substation through the overhead line system and the steel tracks. So the change of the impedance of the tracks and the resistance of the ballast may cause the unbalance transmission of the traction current in two steel tracks. As shown in Figure 2, suppose the traction currents in two steel tracks are I_a、I_b, then the total traction current and the unbalanced current would be $I_g = I_a + I_b$ and $I_n = I_a - I_b$, therefore the traction currents could be expressed as:

$$I_a = (I_g + I_n)/2$$
$$I_b = (I_g - I_n)/2 \quad (1)$$

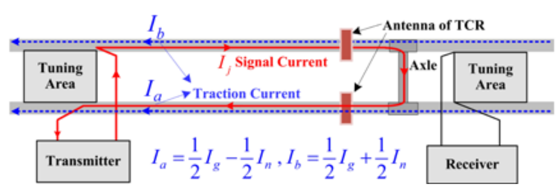

Figure 2: Signal And Traction Currents In Two Tracks

It can be seen that traction current I_a consists of $I_g/2$ and $I_n/2$, and traction current I_b consists of $I_g/2$ and $-I_n/2$. As shown in Figure 3, the induced voltages in the two receiving coils of TCR caused by component $I_g/2$ of I_a and that of I_b through electro-magnetic induction would then be cancelled by each other. Similarly, the induced voltages in the two receiving coils, caused by the $I_n/2$ of I_a and the $-I_n/2$ of I_b, would be added to each other forming the interference against induced voltage respectively due to the track circuit signal.

Journal of Theoretical and Applied Information Technology
15th December 2012. Vol. 46 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

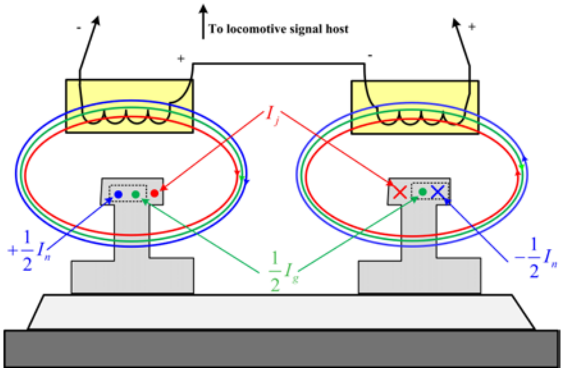

Figure 3: Ground-Train Information Transmission And ITC Formation Mechanism

3. DATA BASED PROBABILITY DISTRIBUTION MODELING OF THE ITC

3.1 Generation of the Statistical Samples of ITC

Assume that the train operating data set is described as:

$$D = \{D_1, D_2, ..., D_N\} \quad (2)$$

Where, D_i represents the data of induced voltage collected by the train running in a single track circuit, N is the capacity of the sample. FFT (Fast Fourier Transform) with a frequency resolution of 0.1Hz is done to D_i and its frequency spectrum is denoted as S_i .The frequency of the fundamental wave is $f_0 = 50\text{Hz}$ in this paper and a fluctuation range $\eta = 1\%$ is supposed for the frequency of traction voltage. Then the h-th harmonic of D_i is extracted from S_i by:

$$e_i(h) = \max_{j=j_h-5h}^{j_h+5h}\left(|S_i(j)|^2\right) \quad (3)$$

Where, j is the index of the spectrum and j_h stands for the index of the spectrum at the frequency of the h-th harmonic.

In order to reflect the proportionate relationships between harmonics and fundamental wave, a database containing the ratios of harmonics to the fundamental wave is established based on Equation (4), which could be denoted as:

$$E_0(h)=\begin{cases}\left\{X_i^{(h)} \mid X_i^{(h)} = e_i(h), i = 1,2,...,N\right\} & ,h = 1\\ \left\{X_i^{(h)} \mid X_i^{(h)} = \dfrac{e_i(h)}{e_i(1)}, i = 1,2,...,N\right\} & ,h > 1\end{cases} \quad (4)$$

The Grubbs Criterion is then used to process the sample of $E_0(h)$ so as to eliminate the influence of maximum value or minimum value. Without loss of generality, same symbols in the Equation (4) are used to describe the final database:

$$E(h)=\begin{cases}\left\{X_i^{(h)} \mid X_i^{(h)} = e_i(h), i = 1,2,...,N_1\right\} & ,h = 1\\ \left\{X_i^{(h)} \mid X_i^{(h)} = \dfrac{e_i(h)}{e_i(1)}, i = 1,2,...,N_h\right\} & ,h > 1\end{cases} \quad (5)$$

3.2 Probability Distribution of the Harmonics Based On the Kolmogorov-Smirnov Test

The Kolmogorov–Smirnov (K–S) test[4] is a nonparametric test for the equality of continuous, one-dimensional probability distributions that can be used to compare a sample with a reference probability distribution. The Kolmogorov–Smirnov statistic quantifies a distance between the empirical distribution function of the sample and the cumulative distribution function (CDF) of the reference distribution. The null distribution of this statistic is calculated under the null hypothesis that the sample is drawn from the reference distribution. This method has considered both local and whole which could achieve a good accuracy.

Following steps are taken to establish the probability distribution of the harmonics.

1) Calculate the empirical distribution function $F_e^{(h)}(x)$ for the $E(h)$:

$$F_e^{(h)}(x) = \frac{1}{N_h}\sum_{i=1}^{N_h} f_{in}(X_i \le x) \quad (6)$$

Where, $f_{in}(X_i \le x)$ is the indicator function, equal to 1 if $X_i \le x$ and equal to 0 otherwise.

2) The following hypothesis is then made:

$$H_0 : F_e^{(h)}(x) \equiv F_0(x) \leftrightarrow H_1 : F_e^{(h)}(x) \ne F_0(x) \quad (7)$$

Here the H_0 is the null hypothesis that $E(h)$ is drawn from the distribution $F_0(x)$ and H_1 is the alternative hypothesis.

3) Evaluate the distribution parameters P_a of the $F_0(x)$. Here in this paper Maximum likelihood (ML) [5] method is used to evaluate the optimum parameters $P_a^{(h)}$ of P_a for the sample $E(h)$, thus the hypothesized distribution could be determined as $F_0^{(h)}(x)$.

4) Calculate the Kolmogorov–Smirnov statistic[6] for the distribution $F_0^{(h)}(x)$:

$$T^{(h)} = \max_{i=1}^{N_h}\left\{\left|F_e^{(h)}(X_{(i)}^{(h)}) - F_0^{(h)}(X_{(i)}^{(h)})\right|\right\} \quad (8)$$

5) Test the hypothesis. A critical value $T_{dist}^{(h)}(N_h, \alpha)$ of the Kolmogorov–Smirnov statistic， which

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

which does not depend on $F_0^{(h)}(x)$, could be looked-up from the table of Kolmogorov distribution[7] that was published by Nikolai Vasilyevich Smirnov[8]. Here α is the significance level meaning the maximum probability to reject the null hypothesis while it is actually right and N_h is the capacity of $E(h)$.The null hypothesis should be accepted at a probability of $1-\alpha$ if $T^{(h)} < T_{dist}^{(h)}(N_h, \alpha)$ is achieved.

3.3 Probability Distribution of the Harmonics Based On Least Square Approximation

According to the least square approximation algorithm, it is feasible to use an appropriate polynomial to fit the probability density function (PDF) of the harmonics of the ITC.

Maximum value $X_{\max}^{(h)}$ and minimum value $X_{\min}^{(h)}$ of the sample $E(h)$ is firstly searched in order to draw the histogram. Least-square polynomial fitting method is adopted to get a simple expression of the frequency density. The fitted polynomial is expressed as $p_0^{(h)}(x)$ which could be regarded as the PDF of the sample $E(h)$.Furthermore, the CDF $F_0^{(h)}(x)$ could be established in terms of the PDF $p_0^{(h)}(x)$ as following equation:

$$F_0^{(h)}(x) = \int_{-\infty}^{x} p_0^{(h)}(t)dt \tag{9}$$

There would be a constant term in the result of the Equation (9) which is determined by the following equation.

$$F_0^{(h)}(X_{\max}^{(h)}) = F_e^{(h)}(X_{\max}^{(h)}) \tag{10}$$

Where $F_e^{(h)}(x)$ is the empirical CDF of the sample $E(h)$.

3.4 Error Analysis of the Probability Model of the Harmonics of the ITC

The root mean square error (RMSE) of the fitted distribution function and the empirical distribution function, defined as Equation (11), is used to evaluate the accuracy of fitting. The closer to zero the RMSE is, the better the probability distribution model is.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N_h}\left(F_0^{(h)}(X_i^{(h)}) - F_e^{(h)}(X_i^{(h)})\right)^2}{N_h - 1}} \tag{11}$$

4. EXAMPLE ANALYSIS

Taking advantage of 476 segments of induced voltage data collected by the train running along the 'Jing-Hu' high-speed railway, the harmonic samples $\{E(h), h = 1, 2, ..., 20\}$ are extracted according to Equation (2)-(5) and then the probability distribution model is established using the K-S test and the least square approximation method.

4.1 Probability Distribution of the ITC Based On K-S Test

For each harmonic sample $E(h)$, four hypotheses that the sample is drawn from Gamma distribution [9,10], Generalized Extreme Value (GEV) distribution [11], 2-parameter Log-logistic(LL2) distribution[12] and Weibull distribution [13] are made. Then the hypotheses are tested at a significance level $\alpha = 0.05$. The results of the K-S statistic $T^{(h)}$ and the critical value $T_{dist}^{(h)}(N_h, \alpha)$ are shown in table 1.

The hypothesized distribution that has the least K-S statistic is chosen as the final distribution model for the samples that accepted more than one hypothesis. And for samples rejected all the hypotheses, none distribution is chosen in table 1 and the least square approximation method should be used to establish the probability distribution model.

Journal of Theoretical and Applied Information Technology
15th December 2012. Vol. 46 No.1

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 1: Results of the K-S test for the hypothesis

Sample	Name and K-S Statistic of hypothesized distribution				Critical Value	Distribution Model
	Gamma	GEV	LL2	Weibull	$T_{dist}^{(h)}(N_h, \alpha)$	
$E(1)$	0.0819	0.1102	0.097	0.0818	0.0655	None
$E(2)$	0.0719	0.0644	0.0501	0.0477	0.0671	Weibull
$E(3)$	0.0261	0.0394	0.0447	0.0340	0.0667	Gamma
$E(4)$	0.1265	0.0539	0.0689	0.1002	0.0672	GEV
$E(5)$	0.1455	0.0643	0.0742	0.123	0.0669	GEV
$E(6)$	0.1763	0.0728	0.1098	0.1426	0.0676	None
$E(7)$	0.0739	0.0596	0.0448	0.0644	0.0675	LL2
$E(8)$	0.0753	0.1272	0.0889	0.0722	0.0670	Weibull
$E(9)$	0.149	0.0755	0.0759	0.1231	0.0676	None
$E(10)$	0.1398	0.0639	0.0768	0.1057	0.0671	GEV
$E(11)$	0.1179	0.0618	0.0601	0.0899	0.0668	LL2
$E(12)$	0.1257	0.0835	0.054	0.0837	0.0667	LL2
$E(13)$	0.0537	0.0785	0.0837	0.0544	0.0671	Gamma
$E(14)$	0.1524	0.0671	0.0737	0.1162	0.0069	None
$E(15)$	0.1337	0.0523	0.0582	0.0976	0.0672	GEV
$E(16)$	0.1016	0.1458	0.1049	0.1023	0.0669	None
$E(17)$	0.1306	0.0726	0.0494	0.0884	0.0667	LL2
$E(18)$	0.1484	0.0516	0.0609	0.0996	0.0672	GEV
$E(19)$	0.1346	0.0580	0.0613	0.0848	0.0662	GEV
$E(20)$	0.0733	0.0562	0.0348	0.0499	0.0669	LL2

4.2 Probability Distribution of the ITC Based On Least Square Approximation

Sample $E(1)$ is taken as an example to illustrate procedure to establish the probability distribution model using least square approximation. Firstly, the frequency density of the sample is calculated and the histogram of the frequency density is shown in Figure 4. Least square polynomial fitting method is then used to fit the density curve.

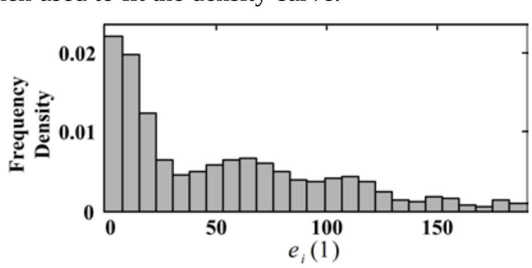

Figure 4: Histogram Of $E(1)$

In order to higher the accuracy of the fitting, polynomials of degree 5 to 14 are used to fit the curve and RMSEs are calculated according to Equation (11). Then the polynomial that has the least RMSE is chosen as the PDF of the sample. As for sample $E(1)$, a polynomial of degree 14 is achieved by this strategy and the CDF is achieved by integral of the polynomial. The RMSE is 0.0161, which indicates a good accuracy of the model. The curve of the PDF is shown in Figure 5.

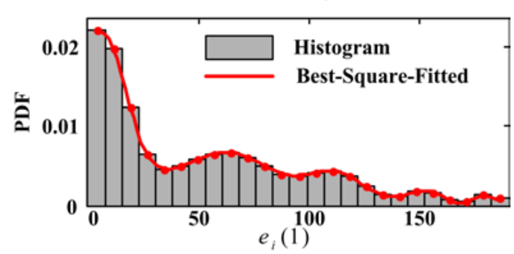

Figure 5: Probability Density Function Of $E(1)$

The same procedure has been taken to establish the probability distributions of the rest of samples that have rejected all four hypothesized distributions. The RMSEs of the fitted CDF and the empirical CDF are 0.016, 0.026, 0.018, 0.026 and 0.013 for the samples $E(1)$, $E(6)$, $E(9)$, $E(14)$, $E(16)$, respectively. The curves of the probabilities

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

calculated based on the PDFs are shown in Figure 6-Figure 9.

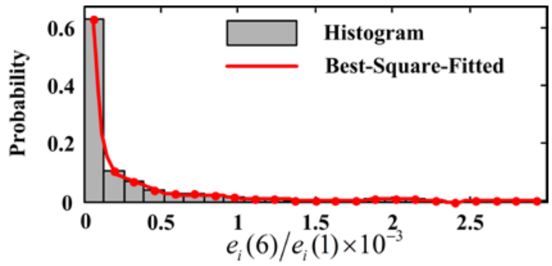

Figure.6: Probability Distribution Of The $E(6)$

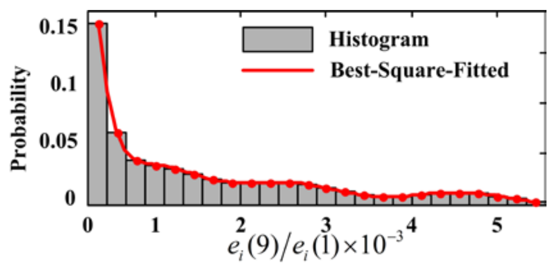

Figure 7: Probability Distribution Of The $E(9)$

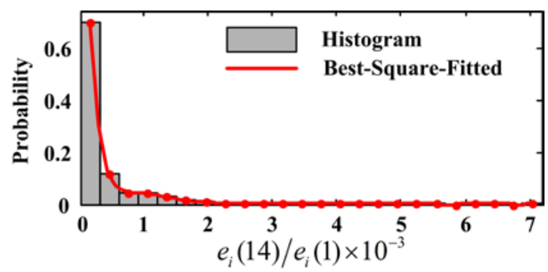

Figure 8: Probability Distribution Of The $E(14)$

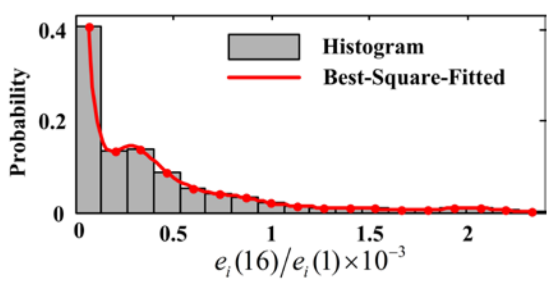

Figure 9: Probability Distribution Of The $E(16)$

4.3 Simulation of the ITC Based On the Probability Distribution Model

First of all, random number $x_e^{(h)}$ could be generated from the distribution $F_0^{(h)}(x)$ by solving the equation:

$$F_0^{(h)}(x) = Y_{rnd} \tag{12}$$

Where the Y_{rnd} is a random-uniform number within interval $\left[F_0^{(h)}(X_{\min}^{(h)}), F_0^{(h)}(X_{\max}^{(h)})\right]$.

Then according to Equation (5),the amplitude of the harmonic could be calculated by:

$$A_h = \begin{cases} \sqrt{x_e^{(h)}} & ,h=1 \\ \sqrt{x_e^{(h)}\cdot x_e^{(1)}} & ,h>1 \end{cases} \tag{13}$$

Finally, as shown in Equation (14),a simulation signal of ITC could be achieved by superposition of series of sinusoidal signals.

$$S_{tci}(t) = \sum_{h=1} A_h \cdot \sin(2\pi h f_0 t) \tag{14}$$

A simulation ITC signal is achieved by this strategy and its frequency spectrogram is shown in Figure 10.

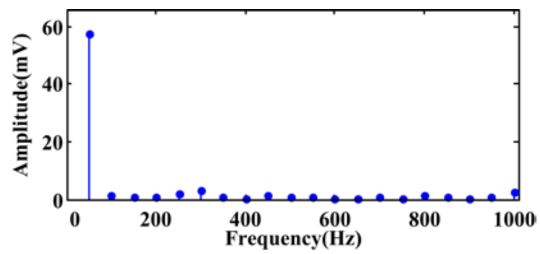

Figure 10 Spectrogram Of A Simulating ITC Signal

A piece of real train operating data, collected in 'Jing-Hu' high speed railway, is randomly chosen as a comparison. The wave and frequency spectrogram of the ITC from the operating data is shown in Figure 11 and Figure12.

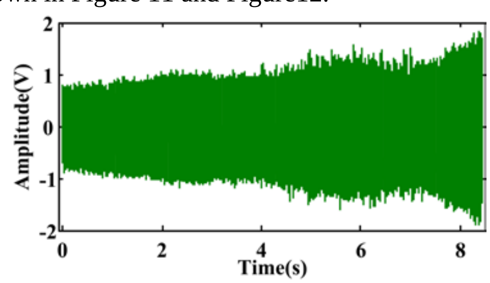

Figure 11: Time Domain Wave Of The Comparison Data

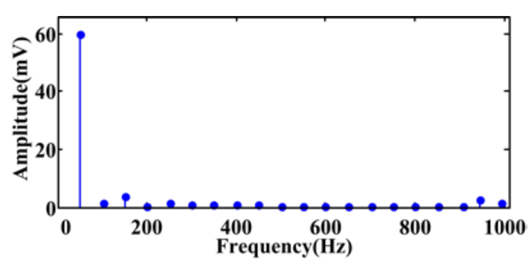

Figure 12: Spectrogram Of ITC Extracted From The Comparison Data

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5. CONCLUSION

A strategy of using K-S test and least square approximation method is firstly designed to establish the probability distribution model of the interference of traction current in railways. By using this strategy, the probability distribution model has been established by taking advantages of the real train operating data collected in ‘Jing-Hu’ high speed railway and a good accuracy of the model has been achieved. By this distribution model, ITC signal could be achieved by simulation instead of field collection which is practically difficult. Feasibility and applicability of the strategy have been confirmed by these experiments.

ACKNOWLEDGEMENTS

This work was supported by the key project of China Ministry of Railway: “The research of basic theory of high-speed railway -- Research on basic theory of reliable transmission information of track circuit and balise in high-speed railway” (2011X021-C).

REFERENCES:

[1] Mariscotti A, “Distribution of the Traction Return Current in AC and DC Electric Railway Systems”, *IEEE Transactions on Power Delivery*, Vol.18, No.4, 2003, pp.1422-1432.

[2] Mariscotti A, Pozzobon Paolo, “Distribution of the Traction Return Current in AT Electric Railway Systems”, *IEEE Transactions on Power Delivery* , Vol. 2, No.3, 2005, pp.2119-2128.

[3] Mariscotti A , Ruscelli M, Vanti M, “Modeling of Audiofrequency Track Circuits for Validation, Tuning, and Conducted Interference Prediction”, *IEEE Transactions on Intelligent Transportation Systems*,Vol. 11, No. 1, 2010, pp.52-60.

[4] Ledwina, Teresa, “Nonparametric tests for stochastic ordering”, *TEST*. Vol.21, No.4, 2012, pp730-756.

[5] Olteanu, Denisa, Freeman, Laura, “The Evaluation of Median-Rank Regression and Maximum Likelihood Estimation Techniques for a Two-Parameter Weibull Distribution Comments”, *Quality Engineering*,Vol 22, No.4, 2010, pp.256-272.

[6] Kolmogorov A, “Sulla determinazione empirica di una legge di distribuzione”, *G. Inst. Ital. Attuari*, Vol.4, 1933, pp. 83.

[7] Marsaglia G, Tsang WW, Wang J, “Evaluating Kolmogorov’s Distribution”, *Journal of Statistical Software*,Vol.8, No.18,2003, pp. 1-4.

[8] Smirnov NV. “Tables for estimating the goodness of fit of empirical distributions”. *Annals of Mathematical Statistics*,Vol.19, 1948, pp. 279.

[9] Magnus Röding, Diana Bernin, Jenny Jonasson, Aila Särkkä, Daniel Topgaard, Mats Rudemo, Magnus Nydén, “The gamma distribution model for pulsed-field gradient NMR studies of molecular-weight distributions of polymers”. *Journal of Magnetic Resonance*, Vol.222, 2012, pp.105-111.

[10] Keating J, Glaser R, Ketchum N, “Testing hypotheses about the sharp parameter of a gamma distribution”, *Technometrics*, Vol.32, 1990, pp.67-82.

[11] Hasan Husna, Fadhilah N, Radi A, Kassim S, “Modeling the Distribution of Extreme Share Return in Malaysia Using Generalized Extreme Value (GEV) Distribution”, *The 5th International Conference on Research and Education in Mathematics (ICREM)*, Inst Teknologi Bandung (ITB), Bandung, October 22-24, 2011,pp. 82-89.

[12] Shoukri. M. M, Mian. I. U. H, Tracy. D. S, “Sampling Properties of Estimators of the Log-Logistic Distribution with Application to Canadian Precipitation Data”, *The Canadian Journal of Statistics*, Vol. 16, No. 3, 1988, pp.223-236.

[13] Ali A. Ismail, “Estimating the parameters of Weibull distribution and the acceleration factor from hybrid partially accelerated life test”, *Applied Mathematical Modelling*, Vol. 36, No. 7, 2012, pp. 2920-2925.