Research on Power Harmonic Noise Suppression Method Based on Genetic Algorithm for Urban TEM Data

Sun Yishu1,2, Huang Sihe1,2, Chen Jian1,2, Zhang Yang1,2

1 College of Instrumentation and Electrical Engineering, Jilin University, Changchun 130061, PRC
2 Key Laboratory of Geophysics Exploration Equipment, Ministry of Education of PRC, Jilin University, Changchun 130061, PRC
E-mail: sunyishu@yeah.net

Abstract. In recent years, ground accidents are constantly occurring in the process of urban underground space development. TEM provides effective technical support for the rational development and safe operation of urban underground space as a flexible and low-cost geophysical method. However, the interference of human electromagnetic noise is serious in the urban environment, especially the power frequency harmonic noise has wide frequency band distribution, as it is difficult to suppress it quickly and effectively, resulting in low accuracy of the extracted secondary field signal and large inversion error. Aiming at the problem of power frequency harmonic noise interference, this paper proposes a noise suppression method based on genetic algorithm. Through the optimization algorithm, the feature matching segment in the reference noise is quickly searched, and it is self-adaptive offset with the power frequency harmonic noise in the signal. The simulation results show that the method can obtain higher signal-to-noise ratio when collecting less sample data. Furthermore, it is proved that this method of this paper has the characteristics of fast noise suppression and no signal loss.

1. Introduction
With the development of social economy, the development of urban underground space has become an effective way to alleviate the shortage of land resources [4,10]. Therefore, applying advanced geophysical prospecting methods to conduct general survey or exploration early warning will provide strong technical support for disaster prevention. Transient electromagnetic method has been widely applied in the field geological structure and geological resources exploration as an efficient means. It has a broad prospect for urban underground space exploration. However, the signal is often submerged in the noise due to the complexity and serious interference of human noise in urban environment, especially the periodic harmonic noise which called power harmonic noise. Therefore, the key to the reliable application of TEM in urban is to suppress the noise.

A lot of research has been done to improve the signal-to-noise ratio of TEM in the world [5,11]. In view of the power frequency harmonic noise, the relevant scholars have adopted the regional subtraction and modeling method to suppress. Butler proposed the method of Block Subtraction to remove harmonic noise from seismic signal [1,2]. This method has the advantages of simple processing, small calculation and no loss of signal. It is difficult to control the accurate and fast phase matching, and the frequency bandwidth is only 10 Hz-1000 Hz, which brings a lot of troubles and certain limitations to the urban noise suppression effect as the accuracy of the method depends on the accuracy of the integer period of the sample area, and it is difficult to achieve the best noise matching, so the it impedes the improvement effect of the signal-to-noise ratio. The above modeling methods are all based on the ideal situation of single noise source and fixed frequency. But in the urban environment, there are many kinds of power line, underground power grid, all kinds of electrical equipment and other power frequency harmonic noise sources, and the frequency is difficult to be fixed. To solve this problem, Thoma et al. proposed a 2-D grid-search approach and segmentation strategy, which effectively removed the harmonic noise under these difficult conditions [9]. However, the frequency range of SNMR signal is only 800-2800 Hz. This method can model and evaluate the amplitude and phase of each power frequency harmonic noise, which has a good effect after suppression. However,
the frequency band of TEM signal is in the range of 0-100 kHz, meanwhile the number of power frequency harmonic noise is large. It is extremely time-consuming and unrealistic to adopt this method. Therefore, a fast and effective method to suppress power frequency harmonic noise is urgently needed to provide technical support for the efficient extraction of accurate secondary field signals.

In view of the above problems, considering the periodicity of power frequency harmonic noise and the fact that it will not mutate in a short period of time, this paper proposes power harmonic noise suppression method based on genetic algorithm for urban TEM data (HNSGAU) on the basis of Butler's research. Firstly, the pure noise data and secondary field signal are collected. Secondly, the HNSGAU is applied to search the optimal coordinates of noise suppression in the feature area. Thirdly, the signal of secondary field is subtracted from the locked noise feature area. Finally, the traditional noise suppression method of TEM system "digital multi-point average" is applied to suppress random noise and residual power frequency noise. The output signal studied in this paper has been obtained. It is proved that this method of this paper has the characteristics of fast noise suppression and no signal loss.

2. Analysis of the influence of urban underground space detection signal

Transient electromagnetic method is a non-contact geophysical exploration method based on electromagnetic induction law. The amplitude of transient electromagnetic signal response has a large dynamic range, and its early signal has high intensity and fast attenuation speed. However, the late signal, which represents the depth information of detection has a much slower attenuation speed and weak signal so as to be submerged in noise. While the late signal includes low-frequency and deep geophysical information. Therefore, the effective data length of the acquired signal is the main factor determining the TEM detection depth. Due to the weak late signal, it is easy to submerge in noise, especially under the interference of power frequency harmonic noise in urban environment. This is also the main bottleneck of TEM in the effective detection of urban underground space.

In the transient electromagnetic system, the noise mainly comes from the external electromagnetic interference and the internal background noise of the instrument exists in the situation of random noise. The noise source of urban environment is mainly the noise of human cultural facilities. The frequency of industrial power in China is 50 Hz, so the interference noise is 50 Hz and its harmonic level. The frequency stability is 0.5%~1%. The collected pure noise is the coupling noise of various types of noise, such as power frequency noise, random noise, etc. The data of measuring points are expressed as follows:

\[ c(t) = r(t) + n(t) \]  

Where \( c(t) \) is the acquisition signal, \( r(t) \) is the real effective signal and \( n(t) \) is the coupling pure noise signal (hereinafter referred to as pure noise). The expression of pure noise is:

\[ n(t) = \sum_{i=1}^{k} a_i \cdot n_i(t) \]  

Where \( k \) is the number of noise types, \( n_i(t) \) is the time domain expression of all kinds of noise, and \( a_i \) is the coupling weight coefficient.

Let \( s_i \) collect a set of signal data and \( n_i \) collect a set of noise data. At a certain time, the signal energy is \( s \) and the noise energy is \( n \). The SNR after \( N_i \) sampling and superposition is as follows:

\[ \text{SNR} = \frac{\sum_{i=1}^{N_i} S_i}{\sqrt{\sum_{i=1}^{N_i} n_i^2}} = \sqrt{\frac{N_i \cdot s}{n}} \]  

It can be seen that the signal-to-noise ratio of the signal is increased by \( \sqrt{N} \) times of the original after \( N \) times digital multi-point average. It can be analyzed that SNR is significantly improved when stacking times are small. But with the increase of stacking times, the improvement of SNR becomes slower. The new method proposed in this paper makes use of this characteristic to achieve the purpose of efficient noise suppression. In the case of a large number of reduction of stacking times, it can achieve the same amount of approximate noise suppression effect and obtain effective geophysical information, i.e. noise suppression quickly is realized.
3. Urban electromagnetic noise reduction method based on HNSGAU

3.1 Noise reduction strategy

How to search the pure noise data, which has the highest correlation between the noise component of the secondary field signal and the pure noise data becomes the key of the research. For the secondary field signal with noise component, the region with strong noise correlation in a group of secondary field signal information features is obtained by fast searching in pure noise data based on the genetic algorithm, and the data length is the same as the secondary field signal, which is called the Noise Feature Domain (NFD). As shown in Figure 1 (a) shows a group of pure noise and single secondary field signal after being polluted. It can be seen that the secondary field signal is completely submerged in power frequency noise and its overall trend is similar to pure noise. Figure 1 (b) shows the result of processing the polluted single secondary field signal with the method in this paper. After the power frequency noise is suppressed, only the random noise pollution is left. In particular, since the subtracted component is only noise, the signal component has no attenuation.

![Figure 1. Method illustration of HNSGAU](image)

The suppression strategy is shown in Figure 2. The process can be divided into four steps. The whole process not only achieves the accurate noise suppression of the secondary field signal, but also completes the noise suppression work efficiently.

![Figure 2. Suppression theoretical flow chart of NFD](image)

3.2 NFD optimal search by genetic algorithm

Genetic algorithm was proposed by Holland, a professor of the University of Michigan in 1975. In the 1990s, genetic algorithm ushered in a period of prosperity and development.
The thought origin of genetic algorithm is Darwin's evolutionism. Simulating the natural selection principle of "survival of the fittest". At the same time, through "crossover" and "mutation" operations, fresh blood injection can ensure the population's moderate evolutionary diversity [6,7,8]. In this way, we can finally choose the solution that best suits our goal. The purpose of using genetic algorithm is to search the target quickly [3], and the essence of optimal search is to search the starting point coordinate of optimal in the NFD.

The genetic algorithm in NFD suppression firstly encodes the initial generation starting point coordinates in binary system to generate the first generation random population for searching. The objective fitness evaluation function is:

\[ f = \frac{\sum_{i=1}^{N} (n(t_{i}+x) - c(t_{i}))^{2}}{N_{i}} \]  

where \( N_{i} \) is the number of effective data points for collecting a set of secondary field signals and \( t_{i}+x \) is the starting point coordinate of NFD search. After decoding, the objective function is brought in and the fitness function value of the initial generation \( \text{eval}(U_{k}) \) is calculated, and the sum of the evaluation function values is calculated:

\[ F = \sum_{i=1}^{400} \text{eval}(U_{i}) \]  

Where \( U_{k} \) is the binary search coordinate without decoding and the probability of each chromosome being copied is:

\[ P_{i} = \frac{\text{eval}(U_{i})}{F} \]  

And then calculate the cumulative probability, which is expressed as:

\[ Q_{k} = \sum_{i=1}^{k} P_{i} \]

In order to select the chromosomes that are more conducive to the search starting point coordinates of optimal noise suppression and generate the next generation of chromosomes, "roulette idea" is applied to screen the crossed and mutated chromosomes by genetic algorithm according to the cumulative probability; the objective evaluation function value is calculated and the current optimal independent variables and function values are recorded until all iterations are completed and the optimal solution is searched.

**4. Simulation experiment and analysis**

In order to verify the effectiveness of HNSGAU, the SNR before and after processing is compared. This paper focuses on the late signal of the deep underground information because of the large amplitude of the early signal of the secondary field to represent the shallow underground information. So the definition only considers that the later 90% of the late signal's signal amplitude of the early signal

\[ \text{SNR}_{\text{early}} = 20 \times \log_{10} \left( \frac{\sum_{i=1}^{400} \text{signal}^{2}(t)}{\sum_{i=1}^{400} \text{noise}^{2}(t)} \right) \]

Where \( N_{e} \) is the total number of acquisition points, signal(t) represents the voltage of discrete time-domain data points of theoretical model and noise(t) represents the voltage of discrete time-domain data points. The quantitative evaluation of improving SNR adopts the method of relative quantity, while the evaluation of absolute quantity adopts the defined "effective absolute residual value" (EAVR), which can also be used to eliminate the gross error in the extreme cases of post-processing. The definition is expressed as:

\[ \text{EAVR} = \frac{1}{N} \sum_{i=1}^{N} (s_{i} - n_{i})^{2} \]

The performance of HNSGAU is verified by simulating the secondary field signal and simulating pure noise. The forward modeling data is used as the theoretical secondary field signal. The simulation model of power frequency component in pure noise is:

\[ n_{s}(t) = g_{1} \cdot \sin(2\pi f_{t}t + \phi_{1}) + g_{2} \cdot \sin(2\pi f_{t}t + \phi_{2}) + g_{3} \cdot \sin(2\pi f_{t}t + \phi_{3}) + g_{4} \cdot \sin(2\pi f_{t}t + \phi_{4}) \]

Where \( g_{1}, g_{2}, g_{3}, g_{4} \) are harmonic weights. The Gaussian white noise model in pure noise is given by the simulation software. Then a random segment of the noise sample is added by the simulation
software to generate the simulation secondary field signal. The simulation results are shown in Figure 3 and 4, which are pure noise simulation images and secondary field signal simulation images. It can be concluded from the figure that a group of pure noise data contains three similar periodic components, and the theoretical real signal of the secondary field is completely submerged in the noise in the later stage. The simulation results show that the input SNR of the polluted secondary field simulation signal is -78.42 dB, and the EARV of the polluted secondary field simulation signal is 7.00×10^{-4}.

![Figure 3. Pure noise simulation image](image)

![Figure 4. Single period secondary field simulation signal](image)

As shown in Figure 5, the optimal solution can be obtained by HNSGAU, and the secondary field signal can be subtracted from the optimal NFD to achieve the purpose of maximizing the suppression of power frequency harmonic noise. The simulation results show that the $\text{SNR}_{\text{out}}$ is -38.62 dB, and the signal noise improvement ratio ($\text{SNIR}$) is:

$$\text{SNIR} = \frac{\text{SNR}_{\text{out}}}{\text{SNR}_{\text{in}}} = \frac{-38.62}{-78.42} = 0.49$$

(11)

The EARV is $4.27\times10^{-4}$ after applying the HNSGAU. It should be noted that the result is not significantly reduced in the order of magnitude as the HNSGAU can only achieve good suppression effect on power frequency harmonic noise and the random noise is not in the scope of HNSGAU. So HNSGAU is only applied as a means of data preprocessing. From the simulation results, we can get the expected results.

![Figure 5. Comparison of polluted secondary field signal, theoretical secondary field signal and signal processed by HNSGAU](image)

5 Conclusion and prospect

The signal of late transient electromagnetic is completely submerged in noise in the scientific research work of urban underground space detection. The existing research foundation is difficult to meet the needs of urban underground space detection. Considering the periodicity of power frequency harmonic noise, which will not mutate in a short time, through the numerical simulation experiment, the qualitative and quantitative reliability analysis of the new method is carried out. The results show that the HNSGAU can effectively preprocess the urban noise dominated by power
frequency harmonic noise, that lay a good foundation for more accurate and efficient extraction of real information and verify the accuracy and efficiency of this method Effectiveness. It is proved that the HNSGAU has the characteristics of fast noise suppression and no signal loss indeed.

The next research direction of this study can be applied to the noise suppression work in the field environment. In addition, the research work of this paper still needs to be studied: first, to improve the core of genetic algorithm; second, to improve the collection method of feature-related noise to obtain the feature domain with stronger feature-related.

Acknowledgements
This research was supported by the National Foundation of China (Grant Nos. 41827803 and 61903151).

References
[1] Bulter K E and Russell R D 1993 Subtraction of powerline harmonics from geophysical records Geophysics 58 898-903
[2] Bulter K E and Russell R D 2003 Cancellation of multiple harmonic noise series in geophysical records Geophysics 68 1083-1090
[3] Bajaj A and Sangwan O P 2019 A Systematic Literature Review of Test Case Prioritization Using Genetic Algorithms IEEE Access 7 126355-126375
[4] Cao L, Li X and Wu W, et al. 2011 Geological modeling research of Suzhou City based on the identification of urban underground resources Proceedings of the 2011 International Conference on Remote Sensing, Environment and Transportation Engineering 3171-3174
[5] СИДОРОВ В А 1985 Импульсная индуктивная электроразведка (Недра)
[6] Davis L 1991 Handbook of genetic algorithms
[7] Holland J H 1992 Genetic algorithms Scientific American 267 66-73
[8] Karr C and Freeman L M 1998 Industrial applications of genetic algorithms (CRC press)
[9] Kremer T, Juul Larsen J and Nguyen F 2019 Processing harmonic EM noise with multiple or unstable frequency content in surface NMR surveys Geophysical Journal International 219 753-775
[10] Liu H, Xie Y and Liu Y, et al. 2019 Mapping the Knowledge Structure and Research Evolution of Urban Rail Transit Safety Studies IEEE Access 7 186437-186455
[11] Nabighian M N 1979 Quasi-static transient response of a conducting half-space—An approximate representation Geophysics 44 1700-1705