International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:5, 2016

A Prediction Method for Large-Size Event Occurrences in the Sandpile Model

S. Channgam, A. Sae-Tang, T. Termsaithong

Abstract**—**In this research, the occurrences of large size events in various system sizes of the Bak-Tang-Wiesenfeld sandpile model are considered. The system sizes (square lattice) of model considered here are 25×25, 50×50, 75×75 and 100×100. The cross-correlation between the ratio of sites containing 3 grain time series and the large size event time series for these 4 system sizes are also analyzed. Moreover, a prediction method of the large-size event for the 50×50 system size is also introduced. Lastly, it can be shown that this prediction method provides a slightly higher efficiency than random predictions.

Keywords**—**Bak-Tang-Wiesenfeld sandpile model, avalanches, cross-correlation, prediction method.

I. Introduction

WHEN natural disasters occurred, they caused serious damages to environment and nearby area. These disasters cannot be accurately predicted and immediately prevented. Many researchers tried so hard to study and develop such prediction methods. Bak et al. introduced a concept of the self-organize criticality (SOC) [1]. They proposed a sandpile model to describe the mechanism of self-organized criticality on a square lattice. They have shown that some dynamical systems with extended spatial degrees of freedom can naturally evolve into self-organized critical state, which is barely stable. This SOC system can be found in several forms of nature such as earthquakes, forest fires and landslides [2]. Several years later, the Bak-Tang-Wiesenfeld sandpile model (BTW sandpile model) became widely well-known as one of a typical model of self-organized criticality. The sandpile model has been further developed by many researchers, called the Olami-Feder-Christensen model and the rice-pile model [3]-[7]. Ramos et al. [8] made some experiments using a pile of beads to represent a self-organized criticality system and a dropped bead to represent a disturbance. They used an image processing to find an internal structure of pile which is the shape factor. A shape factor is a measure of the disorder in pile. They used the cross-correlation analysis between the shape factor and the occurrence of a large-size event of the pile of beads and showed that the shape factor is correlated to the occurrence of large size event. For prediction, they developed an alarm for a large-size event occurrence. The alarm is on when the difference of the spatial average of the shape factor between the current time and 50 time steps before is larger than zero.

Channgam, et al. [9] proposed a factor used to predict the large-size event occurrence in the BTW sandpile model as *the ratio of sites containing 3 grains*. The value of cross-correlation at lag k ; $k=-80,...,-1$ were considered. The value of cross-correlation begins to increase at lag (-80). It was shown that the increase of the ratio of sites containing 3 grains indicate the occurrence of large-size events. This approach can be used to develop a better prediction method of the large-size event occurrence.

In this study, we also presented a prediction method by considering the ratio of sites containing 3 grains. The structure of our study is as follows: In introduction, we described the self-organized criticality system. In Section II, we describe the procedure of the BTW sandpile model. In Section III, we described two related time series used to analyze the cross-correlation. In Section IV, the comparison of cross-correlation among these 4 different system was described. In Section V, we presented our new prediction method. In Section VI, we showed the result of such prediction which is divided into 4 cases and the ability of prediction method for each case. In Section VII, we provided some concluding remarks.

II. Methods and Materials

A. The BTW Sandpile Model

Bak, Tang and Wiesenfeld [1] proposed a model which shows the characteristic of self-organized criticality. The mechanism of model was driven by dropping a grain into a randomly selected site on square lattice. When a site accumulated 4 grains, the grains are redistributed to the 4 nearest sites. At boundary sites, the grains in these sites fall off the square lattice. The redistribution of grains can lead to further instabilities of other sites on a square lattice. In this work, the BTW model on a square lattice of a size $L \times L$ will be studied. The process of adding a grain to the sites of a square lattice by adding one grain to a randomly chosen site i with z_i is according to:

$$z_i \to z_i + 1 \qquad (1)$$

where z_i is the number of grain in site i, $z_i \geq 0$, and $i = 1, \ldots, L^2$. Here, the threshold value z_{th} of this sandpile model is z_{th} = 4. A site is called an unstable site when z_i reaches the threshold value or $z_{th} = 4$. The 4 grains in the site

S. Channgam and A. Sae-Tang are with the Department of Mathematics, Faculty of Science, King Mongkut's University of Technology Thonburi (KMUTT) 126 Pracha Uthit Road, Bangmod, Thung Khru, Bangkok 10140, Thailand (e-mail: siroruth.c@gmail.com, anuwat.sae@kmutt.ac.th).
T. Termsaithong is with the Learning Institute, King Mongkut's University of Technology Thonburi (KMUTT) 126 Pracha Uthit Road, Bangmod, Thung Khru, Bangkok 10140, Thailand (e-mail: teerasit.ter@ kmutt.ac.th).

z_i will move to the neighboring site. Then, each of the 4 grains in the z_i will move upward, rightward, downward and leftward to its neighboring sites (z_{n_i}), respectively. Therefore, the number of grains in the site z_i is decreased by 4 and the number of grains in each of the 4 neighboring site will be increased by 1. (For the case of boundary sites, some of the grains may fall off the lattices.) The process is according to

$$z_i \to z_i - 4 \ , \ z_{n_i} \to z_{n_i} + 1 \tag{2}$$

In this way, neighboring site may be activated, and an avalanche of redistribution events may take place. The neighboring sites keep being updated until all sites are stable. Then we can start adding a grain at the next time step. The avalanche size is defined as the number of grains moving between two consecutive dropping time steps (the grains that fall off the lattice are also included here). The distribution of avalanche sizes follow power-law distribution displayed in Fig. 1.

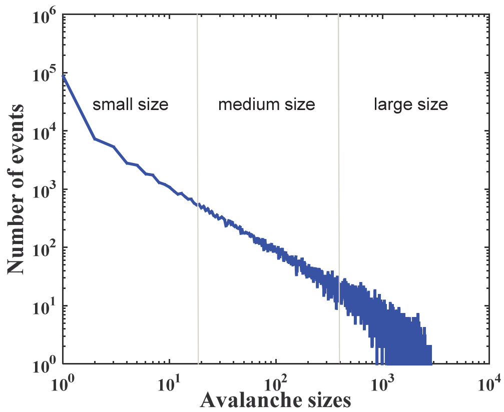

Fig. 1 The distribution of avalanche sizes

III. Related Time Series

A. The Ratio of Sites Containing 3 Grains

This study is interested in a state that is closest to the threshold. We consider the BTW sandpile model with the threshold value $(z_{th} = 4)$. After adding the j^{th} grain and waiting until the end of redistribution, we count the number of sites containing 3 grains(m_j). Then, the ratio of sites containing 3 grains time series can be defined as

$$\tau_j = \frac{m_j}{L \times L} \tag{3}$$

where m_j is the number of sites containing 3 grains at time step $j; j = 1,...,10000$ and $L \times L$ is the system size. Here, we considered the system size $L \times L = 50 \times 50$. The example of the ratio of sites containing 3 grains time series is displayed in Fig. 2.

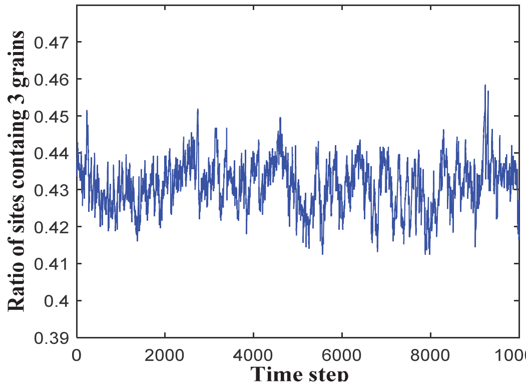

Fig. 2 The example of the ratio of sites containing 3 grains time series for the BTW sandpile model on the lattice size 50×50

B. The Large-Size

We divided the avalanche size into 3 different sizes as in Fig. 1. The 3 sizes were described below.

The first size of the avalanche size is defined as the small size event. The avalanche size is called small size if $10^0 <$ avalanche size $\leq 10^{4/3} \approx 22$.

The second size of the avalanche size is defined as the medium size event. The avalanche size is called medium size if $20 <$ avalanche size $\leq 10^{8/3} \approx 464$.

The third size of the avalanche size is defined as the large size event. The avalanche size is called large size if $398 <$ avalanche size $\leq 10^{12/3} \approx 10000$.

We are only interested the large-size event (avalanche size > 398).

The time series of large-size event can be defined as

$$s_j = \begin{cases} 1 & ;\text{if the avalanche is large} \\ 0 & ;\text{otherwise} \end{cases} \tag{4}$$

IV. The Cross-Correlation

The cross-correlation analysis is a statistical method that measures the relationship between two time series at a given time shift k.

We calculated the cross-correlation between two time series which are the large-size event occurrence time series and the ratio of sites containing 3 grains time series for 10,000 time steps. The cross-correlation $\rho_{s\tau}(k)$ at a lag k is defined as

$$\rho_{s\tau}(k) = \frac{\sum \left(s_j \tau_{j+k} - \mu_s \mu_\tau\right)}{\sqrt{\sum \left(s_j - \mu_s\right)^2 \sum \left(\tau_{j+k} - \mu_\tau\right)^2}} \tag{5}$$

where s_j is the large-size event time series and τ_j is the ratio time series at the time step j; $j = 1,...,10000$. The average of each time series is μ_s and μ_τ, respectively. Here we calculate cross-correlation at lag k for $k = 0, \pm 1, \ldots, \pm 250$. The lag (-$k$) is k time steps before the large-size event occurrence. Lag 0 is

International Journal of Engineering, Mathematical and Physical Sciences
ISSN: 2517-9934
Vol:10, No:5, 2016

time step at the large-size event occurrence. Lag (k) is k time step after the large-size event occurrence.

In [9], the cross-correlation between two time series on sandpile model with system size 50×50 were provided. Here we applied the same procedure to the different system sizes which are $25\times 25, 75\times 75$ and 100×100. The graph of cross-correlation for these 3 system sizes together with the system size 50×50 are shown in Fig. 3. By comparing the cross-correlation of these 3 system sizes with the system size 50×50, it is found that the smaller system size gives the higher cross-correlation at lag 0 than the 50×50 system size with L=50. In addition, the larger system size (75×75 and 100×100) provides the lower cross-correlation than the 50×50 system size.

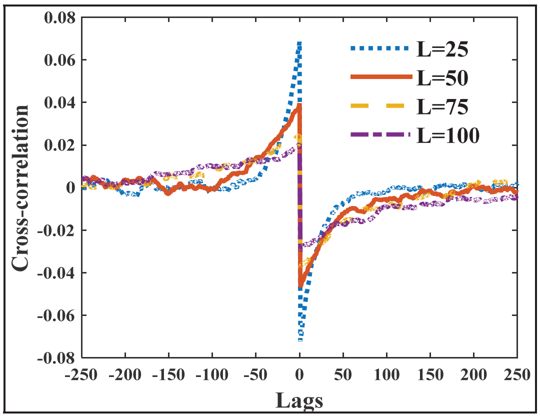

Fig. 3 The value of cross-correlation between the ratio and the large-size event occurrence. Compute cross-correlation on sand pile model with system size L=25, 50, 75 and 100

V. A Prediction Method

Here, we proposed the simple prediction method. Consider the 50×50 system size. From Fig. 3, the value of cross-correlation begins to increase at lag (-80) and increase continuously near lag (-30). Then we compare the ratio of site of containing 3 grains between at the current time step and the previous 30 time steps. Next, we consider the average M_{dL} of the difference $\tau_j - \tau_{j-30}$ for all time step j before the large size event occurrence $(M_{dL} \approx 0.0047)$. Then our simple prediction method can be constructed by predicting that the large size event will occur at time step $j+1$ if the difference $\tau_j - \tau_{j-30}$ is above the average M_{dL} $(\tau_j - \tau_{j-30} \geq M_{dL})$. Otherwise, it predicts that the large size event will not occur at time step $j+1$. Therefore, the simple prediction method is defined as

$$P_{j+1} = \begin{cases} 1 & ;\tau_j - \tau_{j-30} \geq M_{dL} \\ 0 & ;\tau_j - \tau_{j-30} < M_{dL} \end{cases} \tag{6}$$

For all $j = 31, 32, 33, \ldots$.

VI. Results

In this study, we simulated 30 datasets with 10,000 time step per one dataset. We used our simple prediction method to predict the large-size event occurrence.

The results are divided into 4 cases as follows:

Case A: When it predicts that the large-size event will occurs, and the large-size event really occurs.

Case B: When it predicts that the large-size event occurs, but the large-size event not really occurs.

Case C: When it predicts that the large-size event does not occurs, but the large-size event really occurs.

Case D: When it predicts that the large-size event does not occurs, and the large-size event not really occurs.

The meaning of each case is:

- Case A means that the large-size event really occurs which is true positive, and we can prevent damage immediately.
- Case C means that the large-size event really occurs but we cannot prevent damage immediately because the prediction is false positive.
- Case B and Case D means that the large-size event not really occurs.

Here, we focus on Case A. The results of the simple prediction method are shown in Table I.

TABLE I
The Result of the Simple Prediction Method

Results of the simple prediction method			
Case A	Case B	Case C	Case D
6.61%	93.39%	5.91%	94.09%

From Table I, in Case a, the ability of the simple prediction method is 6.61%. To check the ability of this prediction method whether it works well or not, we use probability of a large-size event occurrence $(p = 0.0608)$ which obtained from the simulated avalanche 30 datasets on sandpile model. We compared between the mean of Case A and the mean probability of a large-size event occurrence. The simple prediction method works well if and only if the mean of Case A greater than 0.0608. We concluded that the simple prediction method works better than random guessing (t-test = 5.991 and p-value < 0.05).

VII. Discussion and Conclusion

In this study, we proposed the simple prediction method. We used the difference of the ratio of sites containing 3 grains at current time step and 30 previous steps which greater than or equal to the mean of the different of $\tau_j - \tau_{j-30}$ at time step before the large size occurrence. The cross-correlation is shown in Fig. 3. The value of cross-correlation begins to increase at lag (-30). Then we can use the time window smaller than 30 time steps before the large-size event occurrence.

The maximum cross-correlation decreases when the system size increases. Consequently, different sizes of a system give different values of cross-correlation. According to (5), this difference is caused by the difference between the product of the large-size event and the ratio of sites containing 3 grains time series $(s\tau)$ and the product of the averages of these time series $(\mu_s\mu_\tau)$, which is higher in a system of small size. The other cause is that the variance of τ in the system of small size is less than the variance of the larger system. Furthermore, the fact that the graph of the cross-correlation of small systems increasing more highly than that of the larger system implies that the rise of the ratio of sites containing 3 grains time series in a small system is more obvious than that in a larger system. On the other hand, when the system size approaches infinity, we probably cannot see the rise of the ratio before a large avalanche. Moreover, from the increasing of cross-correlation before the occurrence of large-size event (Fig. 3), it is noticed that rise of the ratio in the larger systems occurs earlier than that in small systems. Thus, it can be implied that we can get the warning sign (rise of the ratio) earlier in the larger system.

To measure the ability of prediction method, here we used only Case A which may be not the actual ability of prediction method. We can use all Cases (Case A, Case B, Case C and Case D) by applying with the Receiver Operating Characteristic Curve (ROC Curve) [10] to measure the ability of prediction method more precisely.

In conclusion, the ratio of sites containing 3 grains time series can be used for predicting the occurrence large-size event. We used the cross-correlation method, which is a method for indicating the pattern of relationship between two time series data sets, and calculated cross-correlation between the ratio of sites containing 3 grain time series and the time series of large-size event occurrence on the system size are $25\times 25, 50\times 50, 75\times 75$ and 100×100. Our result has shown that the value of cross-correlation at lag 0 given the maximum value (0.07) in the small system size 25×25. For the larger system size $50\times 50, 75\times 75$ and 100×100, it is found that the value of cross-correlation decreased. For the case of 50×50 system size, the simple prediction method

$$P_{j+1} = \begin{cases} 1 & ;\tau_j - \tau_{j-30} \geq M_{dL} \\ 0 & ;\tau_j - \tau_{j-30} < M_{dL} \end{cases}$$

is introduced. This prediction method provides a slightly more efficiency than random guessing.

Acknowledgement

This study is partially supported by Dhurakij Pundit University, Thailand.

References

[1] P. Bak, C. Tang and K. Wiesenfeld, Self-Organized Criticality. *Physical review A*, Vol 38, No. 1, 1988 pp. 364-375.

[2] D. L. Turcotte and B. D. Malamud, Landslides, forest fires, and earthquakes: examples of self-organized critical behavior. *Physica A: Statistical Mechanics and its Applications*, Vol 340, No. 4, 2004, pp. 580-589.

[3] B. K. Urnanc, S. Havlin and H. E. Stanley, Temporal correlations in a one-dimensional sandpile model, *Physical review E*, Vol 54, No. 6, 1966, pp. 6109-6113.

[4] E. Altshuler, O. Ramos, C. Martínez, L. E. Flores, and C. Noda, Avalanches in One-Dimensional Pile with Different Types of Bases, *Physical review letters*, Vol 86, No. 24, 2001, pp. 5490-5493.

[5] W. J. Reed and K. S. McKelvey, Power-law behavior and parametric models for the size-distribution of forrest fires, *Ecological Modelling*, Vol 150, No. 3, 2001, pp. 239-254.

[6] Z. Olami, H. J. S. Feder, and K. Christensen, Self-organized criticality in a continuous, nonconservative cellular automaton modeling earthquakes, *Physical review letters*, Vol 68, No. 8, 1992, pp. 1244.

[7] L. A. Núñez-Amaral and K. B. Lauritsen, Self-organized criticality in a rice-pile model, *Physical review E*, Vol 54, 1996, pp. 6109-6113.

[8] O. Ramos E. Altshuler and K. J. Måløy, Avalanche prediction in a self-organized pile of beads. *Physical review letters*, Vol 102, No. 7, 2009, pp. 078701(1)- 078701(4).

[9] S. Channgam, A. Sae-Tang and T. Termsaithong, in *Proc. 6th International Science, Social Science, Engineering and Energy Conf.,* Udonthani Rajabhat University, Thailand, 2014, pp. 26-31.

[10] A. Deluca, N. R. Moloney, and A. Corral, Data-driven prediction of thresholded time series of rainfall and self-organized criticality models, *Physical review E*, Vol 91, No. 5, 2015, pp. 6109-6113.