

# The Key Algorithm of the Sterilization Effectiveness of Pulsating Vacuum Sterilizer

Xicheng Fu<sup>1,\*</sup>, Mingcai Lin<sup>2</sup>, Xingdong Ma<sup>1</sup>

<sup>1</sup> School of Information Engineering, Hainan Institute of Science and Technology,  
China 571126  
davidxm99@126.com

<sup>2</sup> School of Information and Technology, Hainan Normal University, China 571158

**Abstract.** The relationship between the physical parameters of pulsating vacuum sterilizer and the piping system states is analyzed based on cluster analysis and interpolation approximation algorithm. The cluster analysis is performed first, followed by the interpolation approximation of the cluster data. The algorithm provides a comprehensive sterilization efficacy evaluation, as well as a feasible method to analyze the states of the system.

**Keywords:** data mining, cluster analysis, the average quadratic interpolation.

## 1 Introduction

Pulsating vacuum sterilizer is mainly used in hospitals, chemical industry, food industry, and scientific research. It is also widely used in the sterilization of the routine items in medicinal product testing, vaccination, and bioengineering industries. Having replaced the exhaust sterilizer, pulsating vacuum autoclaves are most widely used in hospitals. However, the solenoid valves, the check valves and other valves in the exhaust piping system are prone to failure and need repeated maintenance due to the complexity of the structure of the pulsating vacuum sterilizer piping system, The safety of the pulsating vacuum sterilizer also is a key problem.

## 2 B-D Test

B-D tests are performed to assess the residual air in the pulsating vacuum sterilizer and steam penetration (Figure 1); it is specially designed to test the air exhaust effect of the pulsating vacuum pressure steam sterilizer. The batch challenge test (Process Challenge Device, PCD) is used to determine whether the entire load steam sterilization cycle has achieved the sterilization conditions, often need to release early using the biological indicators or the fifth class of chemical sterilization indicator. The study of sterilization parameters is the prerequisite of a small probability of surviving microbes.

---

\* Corresponding author.

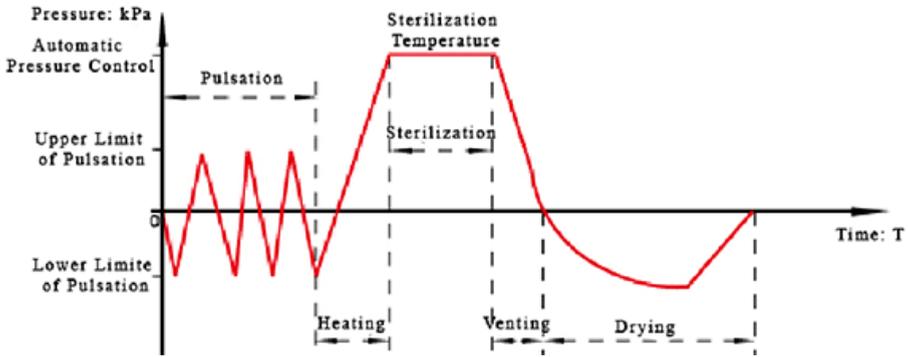


Fig. 1. B-D Test Program

It is often difficult to detect the microbes using the current sterility test method. Therefore, it is necessary to verify the reliability of the method of sterilization. The values of  $F$  and  $F_0$  can be used as the indicators to verify the reliability of the parameters of sterilization process.

Sterilization parameters are mainly:

1) Value of  $D$

At a certain temperature, the sterilization time required to kill 90% of the microorganism (or 10% residual):

Microorganism death rate:  $\lg N_0 - \lg N_t = kt / 2.303$

$$D = t = 2.303 / k(\lg 100 - \lg 10) \tag{2-1}$$

$D$  is the time required to reduce the number of microbes to the one-tenth of the original number, or to reduce to one logarithmic units or decrease ( $\lg 100$  reduce to  $\lg 10$ ). The value of  $D$  depends on the different types of sterilization methods and different types of microbes.

2) Value of  $Z$

$Z$  is the value of temperature increment that is required to lower the value of  $\lg D$ , i.e. sterilization time required to reduce to one tenth of the original number of microbes, or the temperature increment required to kill 99% of the microbes in the same time.

$$Z = T_2 - T_1 / \log D_2 - \log D_1 \tag{2-1}$$

3) Value of  $F$

The time required at the sterilization temperature ( $T$ ) to achieve the same sterilization effect as with the reference temperature ( $T_0$ ), under a certain value of  $Z$ . It is commonly used in dry heat sterilization with the unit of min.

$$F = \Delta t \sum 10 \tag{2-2}$$

4) Value of  $F_0$

The time that is required to sterilize at a temperature ( $T$ ) with  $Z = 10^\circ\text{C}$  to achieve the same sterilization effect at  $121^\circ\text{C}$  with  $Z = 10^\circ\text{C}$ .  $F_0$  value is limited to autoclaving

only. The biological F0 value is equivalent to the time needed to kill all the microbes at 121 °C during hot sterilization. F0 reflects the unity of sterilization temperature and time on the sterilizing effect. This value is more accurate and practical. A safety factor should be increased to ensure the sterilization effect, typically 50% more than the theoretical value in order to achieve the expected remaining number of microbes, i.e., the probability of contamination.

$$\text{The value of physical F0 is defined as: } F0 = \Delta t \sum 10^{T-121/Z} \quad (2-3)$$

$$\text{The value of biological F0 is defined as: } F0 = D121^{\circ C} \times (\lg N0 - \lg Nt) \quad (2-4)$$

How to correctly and timely assesses of the relationships between the operating conditions, such as solenoid valves, check valves in the piping system and the physical parameters that are associated with the sterilization effectiveness will be able to provide a feasible analysis algorithm for the equipment states. This paper constructs the correlation algorithm between the sterilization parameters and the state of the piping system based on the cluster analysis and numerical interpolation approximation.

### 3 Cluster Analysis

The concept of data mining (DM, Data Mining) was first proposed at the ACM annual meeting in 1995. It is the process of extracting implicit, undiscovered, and potentially valuable information from the database. Data mining is the product that IT has reached to a certain stage of development. It requires large-scale databases, efficient computing power, and effective calculation algorithms. Data mining is the process of extracting a useful knowledge from the large amount of data that stored in the database, data warehouse or other libraries.

According to the large number of historical data, cluster analysis is performed to determine the distribution of device parameters under different operating conditions. In this paper, the improved K-Means clustering algorithm is used.

The improved K-means clustering algorithm is discussed as the following:

Method of selecting the initial cluster centers:

- (1). Select the object that is the farthest from the mean point as the initial cluster centers first cluster (seed)  $O_1$ ;
- (2). Select object  $O_2$ , which is the farthest from object  $O_1$ , as the initial center of the second cluster;
- (3). Calculate of the minimum distance between of the remaining unallocated objects and the selected the seeds, and find the maximum value among the minimum distances. The object corresponds to this maximum value is taken as the initial seed of the next cluster;
- (4). Repeat the process K times to achieve K points. These K points are the initial the initial cluster seeds.

In this paper, K=3 is selected. The reason is that the status of the devices is always one of the following: qualified, sub-qualified and disqualified. The corresponding data will be presented in three different distributions. After digging out the three clus-

tering results, the quadratic interpolation is performed with respect to sterilization temperature, sterilization time, etc., to determine the distribution of the three clusters of data which are representative of the state apparatus. Results of data mining are closely related to the types of the data selected. The result of the analysis and the actual experiment may vary greatly if the sample data is not comprehensive. Then we need to select new sample data for data mining.

## 4 Lagrange Interpolation

**Problem 1.1:** Let  $y = f(x)$  has the values of  $y_0, y_1$  at the different nodes of  $x_0, x_1$ , construct a polynomial with the power of  $n$ , with  $n \leq 1$  :

$$p_1(x) = a_0 + a_1x \quad (4-1)$$

so that it satisfies  $p_1(x_0) = y_0; p_1(x_1) = y_1$ . This interpolation is a linear interpolation, also known as Lagrange interpolation.

Linear interpolation calculation is easy; it replaces the curve by a straight line. Therefore it generally requires small interpolation interval  $[x_0, x_1]$ , and changes over this interval should be stable, and otherwise the errors may be significant. To overcome this drawback, we consider using a simple curve to approximate the complex curve. [2]

**Problem 1.2:** Let  $y = f(x)$  has the values of  $y_0, y_1, y_2$  at the different nodes of  $x_0, x_1, x_2$ , construct a polynomial with the power of  $n$ , with  $n \leq 2$  :

$$p_2(x) = a_0 + a_1x + a_2x^2 \quad (4-2)$$

So that it satisfies  $p_2(x_0) = y_0; p_2(x_1) = y_1; p_2(x_2) = y_2$

This interpolation problem is defined by parabolic interpolation, also known as Lagrange quadratic interpolation [3]. Its geometric meaning is to seek a parabolic curve of  $p_2(x)$  to approximate  $f(x)$  via the three nodes  $(x_0, y_0), (x_1, y_1), (x_2, y_2)$  that are on  $f(x)$ .

## 5 The Construction of the Key Algorithm of Equipment Status Analysis by Piecewise Average Quadratic Interpolation Based on Cluster Analysis

### 5.1 Improved K-Means Algorithm

K-Means clustering algorithm refers to the grouping the collection of physical or abstract objects into different clusters according to the similarities the objects.

Input: A database with N objects, the number of clusters K, with  $\beta \in (0,1)$

Output: Clustering result, i.e., K clusters.

Method:

- (1). Select K initial cluster seeds according to the initial cluster centers shown previously;
- (2). Repeat;
- (3). Assign each object the most similar class according to the principle of the most similarity within the class (the closest distance between the clusters and the cluster seeds);
- (4). Calculate the minimum similarity  $MinSim_{t(k-1)}$  between with the cluster seed  $O_{t(k-1)}$  and the cluster data, with the threshold value of  $1 - \beta * (1 - MinSim_{t(k-1)})$ ;
- (5). Select the data with the similarity greater than the threshold in the cluster  $C_{t(k-1)}$  and the cluster seeds  $O_{t(k-1)}$ , to get the set of  $CN_{t(k-1)}$ ;
- (6). Calculate the mean value of  $CN_{t(k-1)}$ , thus being the cluster seed;
- (7). Repeat until there is no further change between the two adjacent cluster seeds.

### 5.2 Piecewise Average Quadratic Interpolation

The piecewise average quadratic interpolation will be constructed in this paper. The procedure of building the average quadratic interpolation is shown as the following:

- (1) Partition the interval  $[a, b]$  as:  $\Delta : a = x_0 < x_1 < x_2 < \dots < x_k = b$
- (2) Construct the Lagrange quadratic interpolation polynomial  $p_{i2}(x)$  over the interval of  $[x_{i-1}, x_i, x_{i+1}]$  ;

$$p_{i2}(x) = y_{i-1} \frac{(x-x_i)(x-x_{i+1})}{(x_{i-1}-x_i)(x_{i-1}-x_{i+1})} + y_i \frac{(x-x_{i-1})(x-x_{i+1})}{(x_i-x_{i-1})(x_i-x_{i+1})} + y_{i+1} \frac{(x-x_{i-1})(x-x_i)}{(x_{i+1}-x_{i-1})(x_{i+1}-x_i)}$$

(5-1)

Likewise, construct the Lagrange quadratic interpolation polynomial  $p_{(i+1)2}(x)$  over the interval  $[x_i, x_{i+1}, x_{i+2}]$ :

$$p_{(i+1)2}(x) = y_i \frac{(x-x_{i+1})(x-x_{i+2})}{(x_i-x_{i+1})(x_i-x_{i+2})} + y_{i+1} \frac{(x-x_i)(x-x_{i+2})}{(x_{i+1}-x_i)(x_{i+1}-x_{i+2})} + y_{i+2} \frac{(x-x_i)(x-x_{i+1})}{(x_{i+2}-x_i)(x_{i+2}-x_{i+1})}$$

(5-2)

- (3) Construct the interpolation polynomial over every interval of  $[x_i, x_{i+1}]$  :

$$p_i(x) = \frac{p_{i2}(x) + p_{(i+1)2}(x)}{2} \tag{5-3}$$

- (4) Stitch together all the  $p_i(x)$  as  $p(x)$ , take  $p(x)$  as the interpolation function of  $f(x)$  over  $[a, b]$ , i.e.:

$$p(x) = p_i(x) \quad x \in [x_i, x_{i+1}] \tag{5-4}$$

## 6 Example of Relationship between the Sterilization Parameters and the Pipeline Equipment States

The data of 1000 pot experiments are collected using the XG1-DWED0.8 pulsating vacuum sterilizer from Shandong Xinhua Medical Instrument Corporation. The data is show as Figure 2. It is shown that the clustering analysis of the equipment operating status with respect to the F0 value using the improved K-means algorithm. The values of F0 for the three clusters, qualified  $F0_M$ , sub-qualified  $F0_N$ , and disqualified  $F0_L$ , are calculated by:

$$F0 = D121^{\circ}C \times (\lg N0 - \lg Nt) \tag{6-1}$$

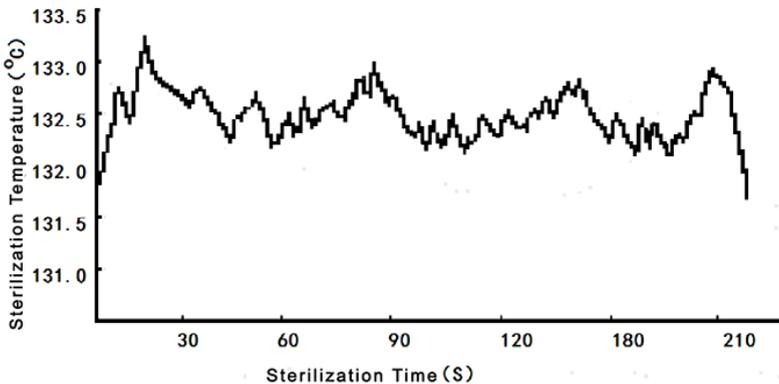


Fig. 2. The curve of sterilization temperature vs. sterilization time

The Lagrange quadratic interpolation function can be constructed using multiple discrete data points according to equation (5-4). The values of F0 for the three clusters are taken as:  $F0_M$  for qualified state,  $F0_N$  for sub-qualified state, and  $F0_L$  disqualified state. Different values of F0 correspond to different operating states. The interpolation curve of disqualified  $F0_L$  according to the above algorithm is shown in

Figure 3. It shows the interpolation of  $F0_L$  when the sterilization steam trap valve malfunctions. The temperature drops below  $132^{\circ}\text{C}$  because of the malfunction. The interpolation curve reflects the equipment state very well. In the mean time, the interpolation function is able to eliminate the interferences of different operating states so that the accuracy of the analysis is improved.

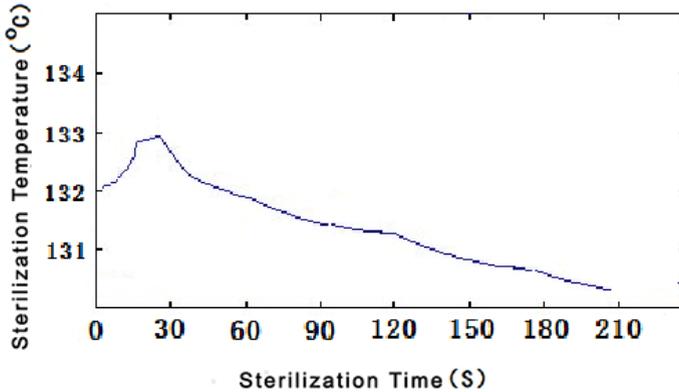


Fig. 3. The Interpolation Curve of  $F0_L$

The working states of the sterilization device can be automatically informed according to the clustering interpolation data curve, with pre-set threshold values of the piping system operating states.

In this example, the equipment operating status is analyzed using the parametric statistical method using the collected data, and good results are achieved. The automatic recording and the automatic notification before reaching the specified number of times are realized.

## 7 Conclusion

In summary, the operating states of the sterilization system and the status of the units can be analyzed using the algorithm proposed in this paper. Thus improve the quality of sterilization. It has ensured the supply of the sterile items in clinical and hospital operating rooms. It can also provide a practical algorithm to analyze system states for other equipments.

**Acknowledgements.** The work of this paper is supported by the Hainan Natural Science Foundation under grant number 614249, and the Scientific Research Project of Hainan Education Department under grant number HNKY2014-111.

## References

1. Association for the Advancement of Medical Instrumentation. ANSI/AAMI ST 79:2006 Comprehensive guide to steam sterilization and sterility assurance in health care facilities (2006)
2. Xu, Z.: The application of data mining in fault diagnoses for gear box. MSc Thesis, North University of China (February 2006)
3. Zhang, Y., et al.: Theory and applications of dynamic fault diagnoses system. National University of Defense Technology Press (1997)
4. Fu, X., Guo, X.: Lagrange interpolation method. Natural Science Journal of Hainan University (6) (2007)
5. Mo, T.: Fault diagnose of Xinhua XG1.DME pulsating vacuum sterilization device. China Medical Equipment (03) (2010)
6. Liu, Y., Zhu, L., Liu, Z.: Two improvements for pulsating vacuum sterilization device. China Instrumentation (04) (1996)
7. Burden, R.L., Faires, J.D.: Numerical Analysis, 3rd edn., pp. 98–86. Prindle, Weber&Schmidt, Boston (1985)
8. Lin, L.: A parallel algorithm of Lagrange interpolation polynomial. Journal of Xiamen University (Natural Science) 43(5), 592–595 (2004)
9. Wu, Z.: Calculation Method, pp. 61–84. Tshinghua University Press (2004)
10. Wu, Q., Wu, D., Fu, X.: Analysis of medical treatment data based on decision tree Computer CD Software and Applications (January 2014)