

Research on data location method under the cloud computing environment

Zhu Qiang

Chongqing College of Electronic Engineering, 401331

Keywords: cloud computing environment; location; Calman filter

Abstract. in the research process of data location method under the cloud computing environment, with the current algorithm to locate data, complexity of correlation time is too high, results in large error for positioning. Therefore, a data location method under the cloud computing environment is proposed based on improved Calman filter algorithm. The method under cloud computing environment, first employs weighted least square method WLS to solve equation, so as to extract estimated value of data initial position, on this basis, the data fusion technology is adopted to eliminate the error existed in the initial position estimated data effectively, RSS value is viewed as weight to process optimal linear weighted, so as to get better data location estimation value, the estimated value is verified by the Calman filter, effectively complete the data location under cloud computing environment. Simulation results show that data positioning method based on improved Calman filter algorithm under cloud computing environment his high precision and strong adaptability.

Introduction

With the development and improvement of the level of computer technology, the scope of application of cloud computing platform in the industry has been more and more widely [1.2.3]. The data location under cloud computing environment is the key link for massive data centralized management and diversified information services in the cloud computing platform, causing a lot of experts and scholars to pay close attention [4.5.6]. Because the data location method under the cloud computing environment has profound significance of the development, it also becomes the focus of the research project [7~10].

At present, the mainstream of data location methods include the method based on least square algorithm, based on Markov algorithm and based on CPLS algorithm. Which are often used in data locating is the method based on the Markov algorithm, but the algorithm complexity is too high, and it has the problem of positioning error.

In view of the above problems, a kind of data positioning method under the cloud computing environment based on the improved Calman filter algorithm is put forward. The method is of high precision and strong adaptability.

The principle of data location under the cloud computing environment

In the data location process under cloud computing environment, assuming to divide n samples of m variables into t classes, x_{ikt} is the k -th variable of i -th sample of t -th class, n_t is the number of samples in t class, \bar{x}_{kt} is gravity center of k -th variable of t -th class, then the quadratic sum of estimation variance of t -th class, can be expressed by the next equation:

$$s_t = \sum_{i=1}^{n_t} \sum_{k=1}^m \left(\|y_{ikt} - \bar{y}_{kt}\| \right)^2 \quad (1)$$

According to formula, it can complete the data positioning under the cloud computing environment:

$$S = \sum_{t=1}^T S_t = \sum_{t=1}^T \sum_{i=1}^{n_t} \sum_{k=1}^m \left(\|y_{ikt} - \bar{y}_{kt}\| \right)^2 \quad (2)$$

Theory of data location optimization method under the cloud computing environment

3.1 extraction of the estimated value of data initial position

Using data fusion technique get rid of the error in the presence of estimated values in the data initial position, and the RSS value is the best linear weighted of weights, to obtain the better estimated values of the data location. The specific steps are as follows in detail:

In the data optimum location under cloud computing environment, assuming that the weighted least square method WLS equation is expressed by the formula:

$$y = cx \quad (3)$$

$$y = [y_1, y_2, y_3, \dots, y_n]^t \quad (4)$$

$$x = [x_1, x_2, x_3, \dots, x_m]^t \quad (5)$$

On the basis of the above formula, error vector in the data optimized positioning process under the cloud computing environment can be obtained with expression:

$$R = CX - Y \quad (6)$$

In the data optimum location under cloud computing environment, the error square and minimum is expressed as:

$$R(X) = (CX - Y)^2 = (CX - Y)^T (CX - Y) \quad (7)$$

In the data optimum location under cloud computing environment, the formula is made derivation, it is 0. It can be drawn the following formula:

$$\frac{dR(x)}{dx} = 2C^T CX - 2CY = 0 \quad (8)$$

In the formula, $C^T CX - CY = 0$, then $X = (C^T C)^{-1} C^T Y$, that is the least square solution of X .

In the data optimum location under cloud computing environment, the function of weighted sum-squared residual (WSSR) in data location, can be expressed by the next formula:

$$R_w(X) = (CX - Y)^t w(cx - y) \quad (9)$$

In the data optimum location under cloud computing environment, assuming, obtained data initial position estimation value are $POS1, POS2, POS3, \dots, POSm$, need to calculate the average to find errors. With the formula to express:

$$posAV = 1/m \times (pos1 + pos2 + pos3 + posm) \quad (10)$$

In the data optimum location under cloud computing environment, data initial position estimation value and the mean value are differenced, the biggest difference is data of obvious error, and the position of error data in the storage space is extracted, represented by the formula:

$$[A, B] = \max(POS_m - PISAV) \quad (11)$$

On the basis of the formula, A can be expressed as the maximum error estimation of data location value, B is the location of the error data.

In the data optimum location under cloud computing environment, POS_i and PR are processed with the optimal linear weighted data fusion and normalized, using the following equation to determine the final location estimation:

$$POS = \frac{1/M \left(\sum (PR(M) \times POS[M]) \right)}{PR_{av}} \quad (12)$$

3.2 The realization of data location under the cloud computing environment

In the data optimum location under cloud computing environment, the estimated value is verified by the Calman filter, effectively complete the data location under cloud computing environment. The specific steps are as follows in detail:

In the data optimum location under cloud computing environment, using the following formula to express the recursive equation of Calman filter:

$$\begin{cases} \chi_0(k-1) = \widehat{X}_2(k-1) \\ \chi_0(k-1) = \widehat{X}_2(k-1) + (\sqrt{(9+\gamma)}p_2(k-1)) \\ \chi_0(k-1) = \widehat{X}_2(k-1) - (\sqrt{(9+\gamma)}p_2(k-1)) \end{cases} \quad (13)$$

In the data optimum location under cloud computing environment, based on the structural data position in the above formula, validation matrix of estimated value and data position sum of squares of deviations can be expressed by the following formula:

$$\begin{cases} W_0 = \gamma / (9 + \gamma) \\ W_i = W_i = 1 / \{2 \times (9 + \gamma)\} \\ W_i = W_i = 1 / \{2 \times (9 + \gamma)\} \end{cases} \quad (14)$$

In this formula: γ is scale factor in the data optimum location under the cloud computing environment to adjust the filter performance.

In the data optimum location under cloud computing environment, based on the above formula, Calman filter gain equation can be expressed by the following equation:

$$k_2(k) = p_{xz} p_{zz}^{-1} \quad (15)$$

Then, in the data optimum location under cloud computing environment, based on the below formula it can complete the accurate location for data:

$$\widehat{x}_2(k) = \widehat{x}_2(k, k-1) + k_2(k) [z_2(k) - \widehat{Z}_2(k, k-1)] \quad (16)$$

The experiment and simulation

In order to prove the effectiveness of the data localization method based on improved Calman filter algorithm under the cloud computing environment, it needs an experiment. The experiment is taken in data positioning simulation platform under cloud computing environment built by *MATLAB*. 2000000 data are collected as experimental data under cloud computing environment. The improved algorithm and the traditional algorithm are used to locate the data under the cloud computing environment, and the positioning accuracy of 2 kinds of algorithms can be compared, the results are shown in figure 1.

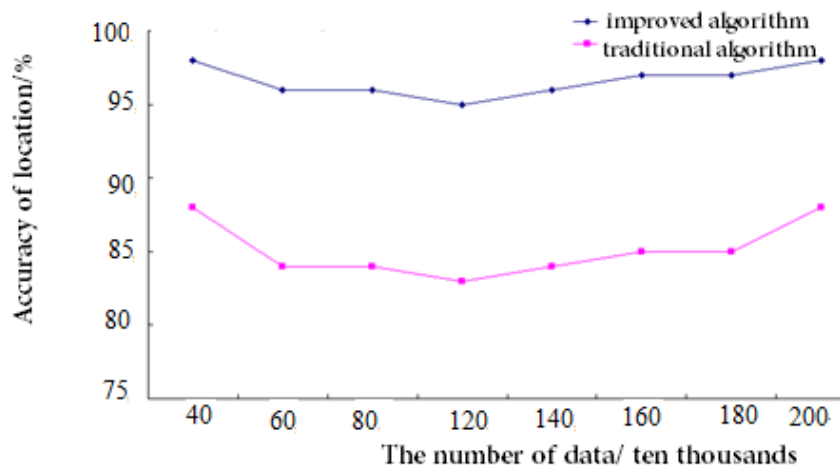


Figure 1 Comparison of data location accuracy of different algorithms

It can be seen from Figure 1, the accuracy of the data location using the improved algorithm under cloud computing environment is superior to the traditional algorithm, and the algorithm is guaranteed the accuracy of positioning.

Conclusion

With the current algorithm to locate data, complexity of correlation time is too high, results in large error for positioning. Therefore, a data location method under the cloud computing environment is proposed based on improved Calman filter algorithm. The method under cloud computing environment, first employs weighted least square method WLS to solve equation, so as to extract estimated value of data initial position, on this basis, the data fusion technology is adopted to eliminate the error existed in the initial position estimated data effectively, RSS value is viewed as weight to process optimal linear weighted, so as to get better data location estimation value, the estimated value is verified by the Calman filter, effectively complete the data location under cloud computing environment. Simulation results show that data positioning method based on improved Calman filter algorithm under cloud computing environment his high precision and strong adaptability.

References

- [1] Song-lin J, Bin Y, Xiao-bin Y. A Study on Data Security for Library under the Cloud Computing Environment[J]. Journal of Yichun University, 2014.
- [2] Shou-ming H, hong-li Z. Research on Apriori Algorithm under Cloud Computing Environment[J]. Journal of Tongling University, 2013.
- [3] Cai Chao, Zuo Xiaoqing, Chen Zhenting. A Cluster Method of Eliminating No-movement Data in Mobile Phone Location Data [J]. Traffic information and security, 2010, 28 (4): 60-63.
- [4] * X C S, Xue X F, Cheng Y P. Research on the Parallel Frequent Data Mining Strategy under the Cloud Computing Environment[J]. Materials & Engineering Technology, 2015.
- [5] Aravamudan N, Cheng K, Filepp R, et al. Method, System and Computer Programs to Assist Migration to a Cloud Computing Environment[J]. Sleepy Hollowny Usglen Ridgenj Usscarsdaleny Usridgefieldct Ussan Franciscoca Ustryorktown Heightsny Uswestportct Usporlandor Us, 2012.
- [6] Li Q, Lu Y, Gong X, et al. A Model of Data Push Service for Mobile Client Based on Fuzzy Synthetic Evaluation under Cloud Computing Environment[J]. Advances in Information Sciences & Service Sciences, 2013.
- [7] Ge J, Zhang B, Fang Y. Research on the Resource Monitoring Model Under Cloud Computing Environment[J]. Web Information Systems & Mining, 2010.
- [8] Chen Q, Lin W, Dou W, et al. CBF: A Packet Filtering Method for DDoS Attack Defense in Cloud Environment[C]. //IEEE International Symposium on Dependable, Autonomic & Secure Computing. IEEE, 2011:427-434.
- [9] Feng Chong, Zuo Xiaoqing, Cai Chao, et al. A Processing Method of Mobile Phone Location Data-based on Time and Spatial-Relation [J]. Journal of Guizhou University: Natural Science Edition, 2011, 28 (3): 78-82.
- [10] Corporation H P. Secure encapsulation and publication of biological services in the cloud computing environment.[J]. Biomed Research International, 2013, 2013(12):491-512.