

## Research Article

# Cloud Platform Based on Mobile Internet Service Opportunistic Drive and Application Aware Data Mining

**Ge Zhou**

*School of Information Engineering, Chongqing Youth Vocational & Technical College, Chongqing 400712, China*

Correspondence should be addressed to Ge Zhou; [gezhou\\_cy@163.com](mailto:gezhou_cy@163.com)

Received 7 July 2015; Accepted 1 September 2015

Academic Editor: James Nightingale

Copyright © 2015 Ge Zhou. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Because the static cloud platform cannot satisfy the diversity of mobile Internet service and inefficient data mining problems, we presented a reliable and efficient data mining cloud platform construction scheme based on the mobile Internet service opportunistic driving and application perception. In this scheme, first of all data selection mechanism was established based on mobile Internet service opportunistic drive. Secondly, through the cloud platform different cloud and channel aware, nonlinear mapping from the service to a data set of proposed perceptual model is applied. Finally, on the basis of the driving characteristics and extraction of perceptual features, the cloud platform would be constructed through the service opportunities of mobile Internet applications, which could provide robust and efficient data mining services. The experimental results show that the proposed mechanism, compared to the cloud platform based on distributed data mining, has obvious advantages in system running time, memory usage, and data clustering required time, as well as average clustering quality.

## 1. Introduction

The cloud platform has high performance of data management and efficient data mining [1], which has been widely applied in various fields, such as knowledge discovery and Leaf Spot dynamics [2], and ERP applications [3]. So, the cloud platform [4, 5] has played an important role for quality of service guarantee and data mining of mobile Internet.

In data mining, Rakocevic et al. [6] embedded a distributed data mining algorithm into a sensor network, which used local predictors on each sensor node to make a local prediction and offers several original voting schemes. The architecture of the huge amount 3D video data was proposed by [7], which designed a model of Key Encryption Model for protecting the privacy video data. Ronowicz et al. [8] researched the cause-effect relationships between pellet formulation characteristics, as well as the selected quality attribute.

About the service driver, Ghosh et al. [9] developed the cost analysis and optimization framework by using stochastic availability and performance models of an IaaS cloud. The economic model as a Bayesian network was proposed by Ye et al. [10] for selecting and composing cloud services. The

importance and challenges in designing event-driven mobile services were discussed [11], which could detect conditions of interest to users and notify them accordingly.

In addition, the fast scalable video coding- (SVC-) based channel-recommendation system for IPTV on a cloud and peer-to-peer (P2P) hybrid platform was studied in [12]. The mechanisms for orchestrating cloud-enabled hardware and software resources were proposed by Petcu et al. [13], which should be supported by a recently developed open-source platform as a service. The novel monitoring architecture was proposed by Alcaraz Calero and Aguado [14], which should be addressed to the cloud provider and the cloud consumers.

However, these research results ignored the relationship between data set of cloud platform and service requirements of mobile Internet and lack the in-depth study of influence of Internet application on cloud platform. Hence, we proposed the cloud platform with mobile Internet service opportunistic drive and application aware data mining.

The rest of the paper is organized as follows. Section 2 describes the mobile Internet services driven data selection mechanism. In Section 3, we design the application perception model. In Section 4, we proposed the efficient

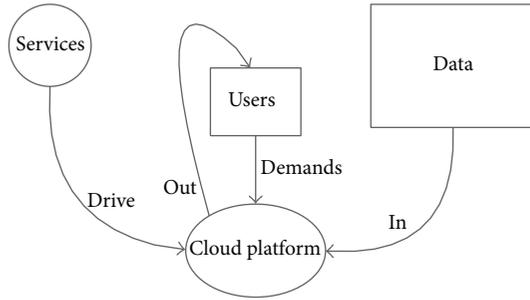


FIGURE 1: Services architecture in mobile Internet.

data mining cloud platform. Simulation results are given in Section 5. Finally, we conclude the paper in Section 6.

## 2. Data Selection Mechanism Driven by Mobile Internet Services

In mobile Internet, we assume a service data set function, which was used to optimize data selection for satisfying the service demands and provided the optimal data set for service. This is the role scheme which was used to solve the combination issue of mobile Internet and cloud computing. In general, the mobile Internet service architecture model is shown in Figure 1. Data input was driven by service set and stored in the cloud platform. When the service request was received successfully, the cloud platform would select the data and sent output feedback to the user according to the request.

There are some problems that should be considered in the progress of data selection, which are as follows:

- (1) whether the data is corresponding to the service and whether it can meet the service needs,
- (2) lack of user information feedback mechanism which is used to judge whether the user is satisfied with the data,
- (3) how to coordinate multiple user needs,
- (4) how to coordinate the input and output of data without increasing the load of the cloud platform.

The mapping relationship of user set and data set is shown in Figure 2. Here, the relationship was constructed by the services, which include the following cases:

- (1) single user to intersection of multiple data sets,
- (2) single data set to multiple users,
- (3) single user to the union of data sets.

According to the above problems, we give the following service driven definition:

- (1)  $SE_{set}(n)$ : service set and its length which is  $n$ ,
- (2)  $DA_{set}(m)$ : data set and its length which is  $m$ ,

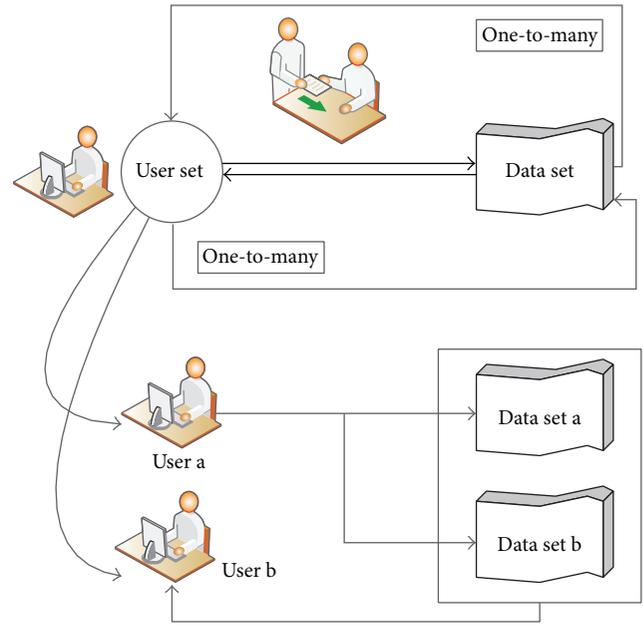


FIGURE 2: Mapping relationship and correlation between users and data set.

- (3)  $US_{set}(p)$ : user set and its length which is  $p$ ,
- (4)  $QE(i)$ : service satisfaction of the  $i$  data,
- (5)  $L_{CP}(j)$ : the number of data services in the  $i$  unit time,
- (6)  $T(i)$ : the time consumed by the  $k$  data service,
- (7)  $L_{CP}(a, b)$ : the number of service sessions between  $a$  and  $b$ .

According to the above stations, there is the relationship shown as formula (1) in mobile Internet application:

$$\begin{aligned}
 N &= L_{CP}(US_{set}(1), DA_{set}(u) \mid 1 < u \leq m) \\
 &\quad + L_{CP}(US_{set}(v) \mid 1 < v \leq n, DA_{set}(1)), \\
 T &= \sum_{i=1}^N QE(i) T(i), \\
 T_h &= \sum_{i=1}^N SE_{set}(i) L_P L_{CP}(i).
 \end{aligned} \tag{1}$$

Here, let  $N$  denote the number of sessions. Let  $L_P$  denote the length of data packet. Let  $T$  denote the time of completing the  $N$  services. Let  $T_h$  denote the throughput.

So, formula (2) gives the data selection model based on service opportunistic driven. Here, service request launched the service opportunistic driven in accordance with the probability. The data sets to meet the service request through the cloud platform. For the request of hybrid multiple data sets of service, the data set could be computed with XOR operation

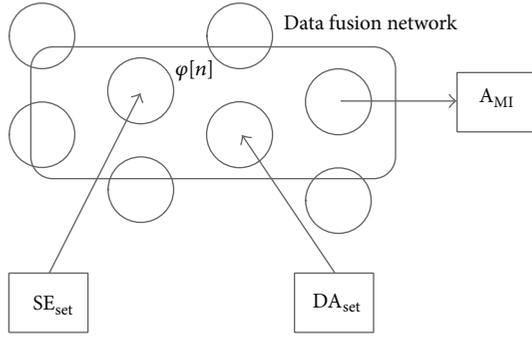


FIGURE 3: Nonlinear mapping relationship.

to form a single data set, for selecting the data without interference and avoiding the performance degradation:

$$DA_{set}(SE) = \begin{cases} T_h, & \text{dataType} = 0, \\ \sum_{type=1}^{T_y} DA_{type}(T_h) \oplus DA_{type+1}(T_h), & \text{dataType} = 1, \end{cases} \quad (2)$$

$$SE_{set} = [SE_1, SE_2, \dots, SE_n],$$

$$D_{rv} = SE_{set} \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}.$$

Here, let  $D_{rv}$  denote the service opportunistic request driven factor. Vector  $[a_1, \dots, a_n]$  denotes the service opportunistic weight. The value of dataType could be set according to the ability of satisfying the service requests. If dataType is 0, the current data set can be independent of the service request. If dataType is 1, more than 2 data sets could satisfy the current service request. Let  $T_y$  denote the number of data sets.  $DA_{set}(SE)$  denote the data set which gives feedback to users.

### 3. Application Perception Model

In the processing progress of mobile Internet applications, cloud platforms can be achieved through nonlinear mapping between different clouds and channel from the service to the data set, which could meet the needs of different data sets from different cloud data fusion. Here, it is important to coordinate transmission and fusion of multicloud applications, as shown in Figure 3. Here, the cycle symbol of Figure 3 denotes the cloud of platform.

Based on the same service request and being located in the same area, data transmission coordination network composed of multicloud needs to be aware of the application and keep the time synchronization. When the application requests are met in different regions, in order to guarantee the space synchronization, the time must keep asynchronous, as shown in formula (3). Based on the above time and space of synchronous request and asynchronous operation, formula

(4) illustrated the application of perception matrix and data output matrix and relationship:

$$\begin{aligned} T(i) &= T(j), \\ S_{pace}(DA_{set}(i)) &= S_{pace}(DA_{set}(j)), \\ T(i) &\neq T(j), \\ S_{pace}(DA_{set}(i)) &\neq S_{pace}(DA_{set}(j)). \end{aligned} \quad (3)$$

Here,  $i$  and  $j$  represent the different application request. Let function  $S_{pace}(DA_{set}(i))$  denote the data set location of the application mapping:

$$A_{MI}[n] = T \otimes (\nabla A_{MI}[n-i] \oplus N(i)), \quad i = 1, 2, \dots, \sqrt{n}, \quad (4)$$

$$\varphi[n] = \ln\left(\frac{\nabla A_{MI}[n-i]}{\sqrt{n}}\right),$$

$$O_{MI} = e^{A_{MI}[n]} DA_{set}(SE)^{\varphi[n]}.$$

Here, let  $A_{MI}[n]$  denote application perception matrix. Let  $\varphi[n]$  denote the multicloud coordination flux. Let  $O_{MI}$  denote the output matrix of data set.

Hence, the data output flow of mobile Internet application perception is as follows:

- (1) initializing the vector and parameters according to the principles and models,
- (2) computing the matrix  $A_{MI}$  according to formula (4),
- (3)  $O_{MI}$  is set to 1, which means service request is active and waiting for data output,
- (4) data selection and integration after the activation of cloud devices,
- (5) the consistency and time of the region are judged according to formula (3),
- (6) if the space is consistent, the time synchronization is obtained by formula (4),
- (7) If the space is not the same, the time is asynchronous. The cloud platform is in a stable state. In an asynchronous cycle, the multiple cloud segmentation data are activated. The end of the independent data set is obtained after data integration.

### 4. Efficient Data Mining Cloud Platform

According to the service opportunities driving characteristics and extraction of perceptual features, how to build cloud platform for mobile Internet applications to provide robust and efficient data mining service has become a key issue of QoS guarantee in mobile Internet. In cloud platform, the customer service satisfaction  $QE(DA_{set})$  could be computed by adding each data service satisfaction as shown in formula (5):

$$QE(DA_{set}) = \sum_{i=1}^n QE(i). \quad (5)$$

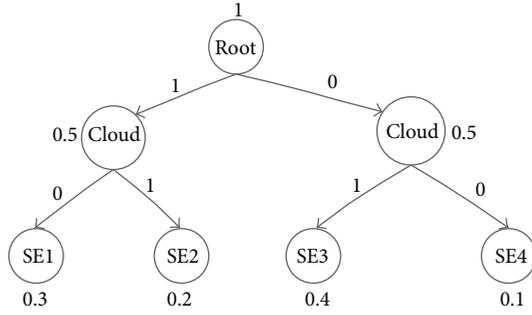


FIGURE 4: Service driven tree.

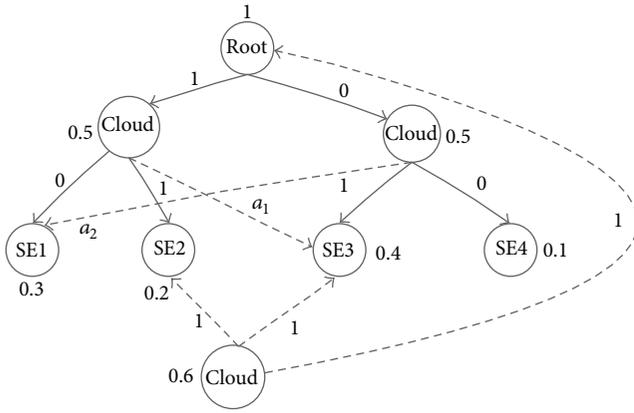


FIGURE 5: Service opportunistic driven tree.

If customer service demand satisfaction of data sets is larger than the user specified minimum value, the data sets are robust. Otherwise, the data sets should be reconstructed through data mining.

In order to facilitate the service opportunity to play the driving characteristics, we defined the service drive tree. The tree consists of  $K$  leaf nodes. Each leaf node represents a service identifier and the opportunity probability. The intermediate node indicates the subtree ability which is denoted by the summary of service opportunities driven weight of leaf nodes. The service driven tree is shown in Figure 4. If the current node and the child nodes are time synchronization, the link will be recorded as 1; otherwise, it is denoted by 0. From the root node to any leaf node, a set of binary bit strings could be obtained. Each leaf node can access the service based on adaptive driving vector, which could reduce data mining complexity. The service opportunities driven tree is shown in Figure 5.

Here, we set 4 leaf nodes and service initiation probability of each leaf node is given. The sum of 4 leaf nodes probability is equal to 1. The service opportunity driven weight of root node is equal to 1. The intermediate nodes are composed of 2 clouds, which are shown as in Table 1.

As shown in Table 1, although the SE4 opportunity weight is low, the service driven path vector is all asynchronous, the data mining complexity is low, and the data clustering service can be obtained with higher real-time. Although the SE3 has the best chance, but the service driven path is heterogeneous,

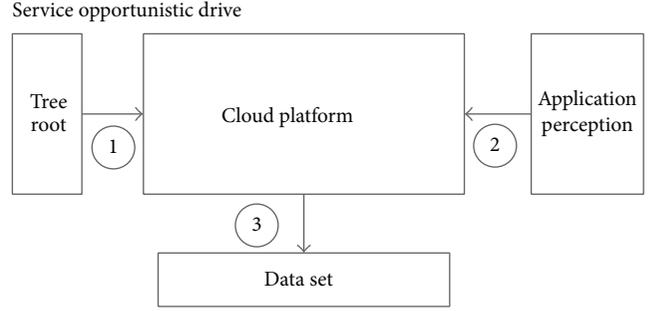


FIGURE 6: Cloud platform architecture for efficient data mining.

TABLE 1: Service driven.

Service ID	Weight	Service vector	Data mining complexity
SE1	0.3	10	High
SE2	0.2	11	Low
SE3	0.4	01	High
SE4	0.1	00	Low

TABLE 2: Service opportunistic driven.

Service ID	Weight	Service vector	Cloud weight	Data mining complexity
SE1	$a_2/0.3$	10/00	0.5/0.6	High/low
SE2	0.2	11	0.5	Low
SE3	$a_1/0.4$	01/11	0.5/0.6	High/low
SE4	0.1	00	0.5	Low

the data mining is complex, so the data clustering is poor, and the real-time performance is poor. In addition, the value of SE1, SE2, SE3, and SE4 could be computed based on the service requests and service driven tree as shown in Figure 4.

As shown in Table 2, with opportunistic driven scheme, the data mining complexity of 4 leaf nodes has been reduced. For SE2 and SE3, we can establish the path of multiple data transmission through the opportunity weighting, so as to ensure the real-time and accuracy of data clustering. In addition, the cloud platform could be constructed by choosing the same area of the current leaf nodes and service opportunities driven. Based on the above service opportunity driving features and application of data clustering mining, the architecture of the cloud platform is as shown in Figure 6. Here, ① expresses the analysis of service opportunistic drive characteristics; ② denotes the feature extraction of application perception; ③ denotes the clustering data mining. The construction algorithm is described as follows.

*Algorithm 1.* Construction of cloud platform is as follows.

*Input.* It includes  $SE_{set}(n)$ ,  $DA_{set}(m)$ ,  $US_{set}(p)$ ,  $QE(i)$ , and data mining complexity threshold  $DM_{th}$ .

*Output.* It includes data set  $DA_{set}(C.P)$  and cloud platform  $C_{P_{DM}}$ .

*Begin*

```

While (1)
Initial parameters:  $L_{CP}(j), L_{CP}(a, b)$ ;
Obtain the value of  $N$ 
While  $i \geq N$ 
If  $QE(i) = 1$  or  $QE(i) = 3$ 
Count time;
 $i++$ ;
End while
Computing the value of  $T$  and  $T_h$  based on  $QE(i)$  and  $T(i)$ ;
Obtain the service opportunity request drive factor  $D_{rv}$ ;
Constructing the service opportunity driven tree;
Analyzing the data mining complexity  $DM$ ;
If  $DM$  is larger than  $DM_{th}$ 
Updating the tree;
Else
If  $S_{pace}(DA_{set}(i)) = S_{pace}(DA_{set}(j))$ 
Time synchronization
Else
Time asynchronous
End if
End if
Computing multicloud coordination flux  $\varphi[n]$ 
Select the optimal clouds to construct the cloud platform  $C\_P_{DM}$ 
Activate the multicloud segmentation data and obtain the final independent data set  $DA_{set}(C\_P)$  after the fusion data
End while

```

*End*

## 5. Performance Evaluation

In order to verify the cloud platform performance and data mining performance of the proposed cloud platform based on opportunities in the service of mobile Internet driver and data mining of application perception (CP-SAD), we designed the experiment with 1 GB storage space and 100 data sets randomly distributed in 10 pieces of cloud equipment and evaluated with the transaction data base server. Verification metrics include the system running time and memory occupancy rate, the time required for data clustering, and the average clustering quality, which compared with the cloud computing mechanism with distributed data mining (CP-DDM); the results are as shown in Figure 7.

From Figure 7(a), we found that CP-DDM mechanism system running time increases linearly with increasing data

set, while the operating time of the proposed CP-SAD mechanism remained stable, because the use of service opportunities driven model and real-time analysis of the characteristics of the driver can accurately grasp the service request, optimize the system initialization and service driven and application aspects of perception in the delay, not only shorten the operation time but also smooth the delay jitter.

Figure 7(b) proved that the proposed mechanism could optimize the memory capacity, because of the application of perceptual features extraction mechanism, which can not only be aware of the applications required data scale and distribution, but also optimize the composition of the cloud platform, by combining with the opportunity to serve driver tree. The proposed mechanism can increase the utilization ratio of memory, avoiding the cloud platform system performance degradation caused by memory overflow.

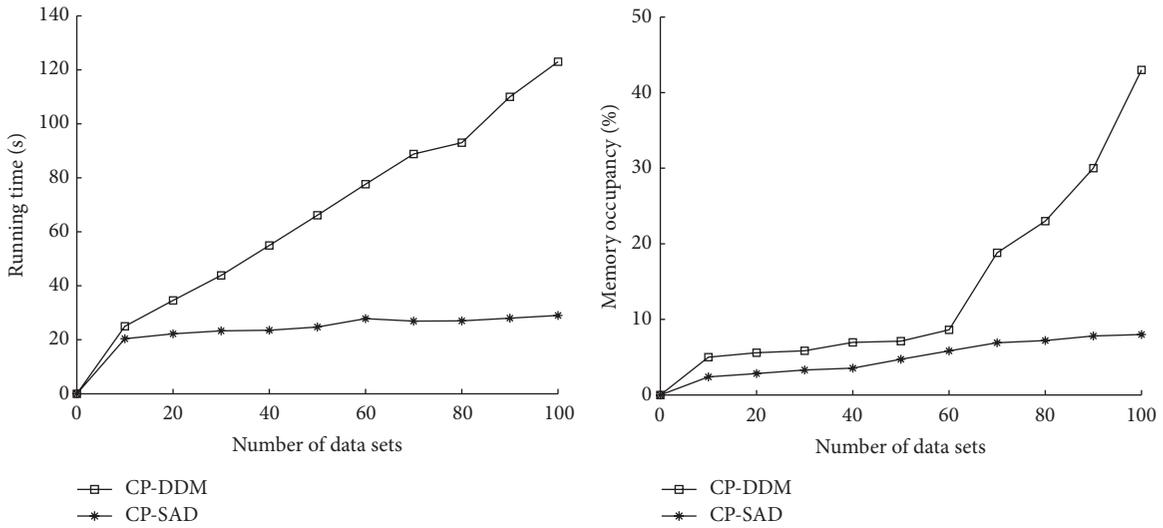
From Figure 7(c), we can see that the required time of data mining clustering increases first and then decreases with the increase of cloud scale activation. This is because the large computation result is more time consuming in initialization phase. However, the proposed CP-SAD mechanism can complete cloud equipment initialization and reduce data mining time faster and earlier. The CP-SAD mechanism can complete initialization after the first cloud device activation, and CP-DDM mechanism complete initialization when three-cloud device is activated, which wasted a lot of time for data mining clustering stage. This is because the distributed data mining is composed of a cloud platform data distributed storage and application features and service requests are not built for nonlinear mapping, resulting in data mining delay being large.

Clustering quality average is shown in Figure 7(d) with the increase of the size of the cloud. In the activation of the cloud number from 0 to 4, the average clustering quality of two mechanisms changes in the same law, and the gap is about 10 percentage points. With a further increase in the cloud, the average clustering quality of CP-DDM mechanism began to decline, until 10 clouds are activated to rise to 60%. However, the average clustering quality of CP-SAD mechanism has stalled when there are four clouds and then increases linearly. This is because the CP-SAD mechanism can effectively grasp the opportunity to drive characteristics and extraction of perceptual features and construct, for mobile Internet applications to provide robust and efficient data mining service cloud platform, mobile Internet applications for improving the efficient and reliable quality of service.

The iteration number analysis results of time complexity of CP-DDM and CP-SAD were shown in Figure 7(e). The proposed mechanism could compute the data efficiently and mine the data accurately with less iteration number than one of CP-DDM, which has the very low time complexity.

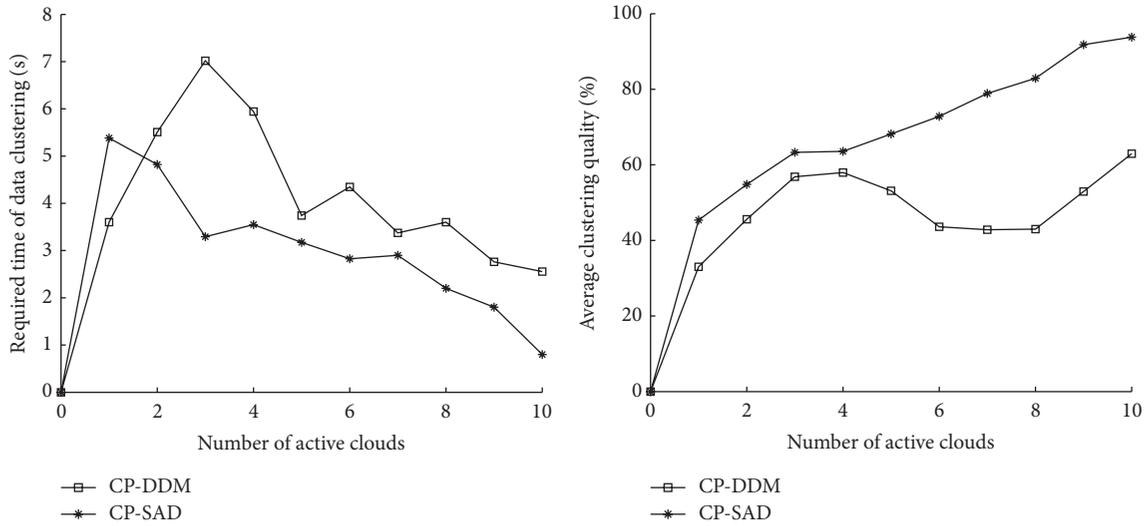
## 6. Conclusions

In order to improve the cloud platform of the mobile Internet services with quality assurance and data mining efficiency, we researched the opportunities in the service of mobile Internet driver, as well as the reliable and efficient application aware



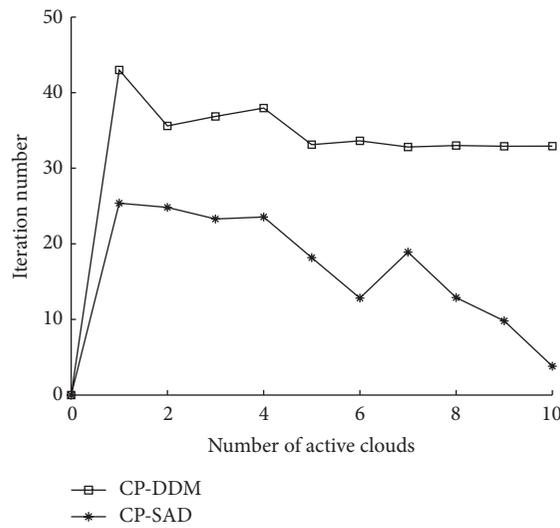
(a) Running time

(b) Memory occupancy



(c) Required time of data clustering

(d) Average clustering quality



(e) Iteration number

FIGURE 7: Performance evaluation.

data mining. Firstly, a data selection mechanism is designed based on the analysis of the characteristics of the mobile Internet service opportunity driving. Second, based on the relationship between the Internet service and data set, the nonlinear mapping between cloud and the Internet channel was established, and it established an Internet application perception model. Third, based on the characteristics of service opportunity driven and application perception, a cloud platform is proposed to provide efficient and robust data mining services for mobile Internet applications. The experimental results show that the proposed mechanism in the runtime and memory occupancy rate, data clustering, and the time required for the average clustering quality are better than the cloud platform based on distributed data mining.

### Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

### References

- [1] A. Cuzzocrea, "Models and algorithms for high-performance data management and mining on computational grids and clouds," *Journal of Grid Computing*, vol. 12, no. 3, pp. 443–445, 2014.
- [2] A. K. Tripathy, J. Adinarayana, K. Vijayalakshmi et al., "Knowledge discovery and Leaf Spot dynamics of groundnut crop through wireless sensor network and data mining techniques," *Computers and Electronics in Agriculture*, vol. 107, pp. 104–114, 2014.
- [3] C.-S. Chen, W.-Y. Liang, and H.-Y. Hsu, "A cloud computing platform for ERP applications," *Applied Soft Computing*, vol. 27, pp. 127–136, 2015.
- [4] J. Carretero and J. G. Blas, "Introduction to cloud computing: platforms and solutions," *Cluster Computing*, vol. 17, no. 4, pp. 1225–1229, 2014.
- [5] H. Qi, M. Shiraz, J.-Y. Liu, A. Gani, Z. Abdul Rahman, and T. A. Altameem, "Data center network architecture in cloud computing: review, taxonomy, and open research issues," *Journal of Zhejiang University SCIENCE C*, vol. 15, no. 9, pp. 776–793, 2014.
- [6] G. Rakocevic, Z. Tafa, and V. Milutinovic, "A novel approach to data mining in wireless sensor networks," *Ad-Hoc & Sensor Wireless Networks*, vol. 22, no. 1-2, pp. 21–40, 2014.
- [7] T. Xu, W. Xiang, Q. Guo, and L. Mo, "Mining cloud 3D video data for interactive video services," *Mobile Networks and Applications*, vol. 20, no. 3, pp. 320–327, 2015.
- [8] J. Ronowicz, M. Thommes, P. Kleinebudde, and J. Krysiński, "A data mining approach to optimize pellets manufacturing process based on a decision tree algorithm," *European Journal of Pharmaceutical Sciences*, vol. 73, pp. 44–48, 2015.
- [9] R. Ghosh, F. Longo, R. Xia, V. K. Naik, and K. S. Trivedi, "Stochastic model driven capacity planning for an infrastructure-as-a-service cloud," *IEEE Transactions on Services Computing*, vol. 7, no. 4, pp. 667–680, 2014.
- [10] Z. Ye, A. Bouguettaya, and X. Zhou, "Economic model-driven cloud service composition," *ACM Transactions on Internet Technology*, vol. 14, no. 2-3, Article ID 2651420, pp. 255–273, 2014.
- [11] A. Boukerche, A. A. F. Loureiro, E. F. Nakamura, H. A. B. F. Oliveira, H. S. Ramos, and L. A. Villas, "Cloud-assisted computing for event-driven mobile services," *Mobile Networks & Applications*, vol. 19, no. 2, pp. 161–170, 2014.
- [12] H.-Y. Chang, C.-C. Lai, and Y.-W. Lin, "A fast SVC-based channel-recommendation system for an IPTV on a cloud and P2P hybrid platform," *Computer Journal*, vol. 57, no. 12, pp. 1776–1789, 2013.
- [13] D. Petcu, S. Panica, C. Crăciun, M. Neagul, and C. Şandru, "Cloud resource orchestration within an open-source component-based platform as a service," *Concurrency and Computation: Practice and Experience*, vol. 27, no. 9, pp. 2443–2469, 2015.
- [14] J. M. Alcaraz Calero and J. G. Aguado, "MonPaaS: an adaptive monitoring platform as a service for cloud computing infrastructures and services," *IEEE Transactions on Services Computing*, vol. 8, no. 1, pp. 65–78, 2015.

