Ranked k-NN Crowdsourced Model for Cloud Internet of Things (CIoT)

Dr. M. Duraipandian,

Head of Department, Department of Computer Science and Engineering,

Vivekanandha College of Technology for Women,

Namakal, India.

Email: svsduraipandian@gmail.com

Abstract: - Internet of Things (IoT) has gained more attention in recent years and its influence over future internet is projected to be more as a promising technology. IoT enables sensors to merge with smart devices to monitor, observe and analyse the real time data. These features make the IoT, a suitable technology, for smart applications. On the other hand, cloud offers a better computing paradigm to store and analyse the data. Cloud reduces the complexities in day today life with its novel applications and services, in an efficient manner. However, present IoT and Cloud solutions are focused towards centralized solutions, which limits the user capacity. To enrich the Cloud integrated IoT benefits, a flexible large-scale data collection and analysis is introduced as crowdsourcing, which provides a new dimension in data mining applications. This research work presents a cloud computing crowdsourced data analysis model implemented over IoT, to obtain better computation speed with improved sensitivity, specificity and accuracy.

Keywords: Cloud Computing, Crowdsourcing, Internet of Things

1. Introduction

Cloud computing is an inevitable technology which offers services through internet. The enhanced features of Cloud attract the user to convert their business from conventional environment to Cloud. It reduces the infrastructure investment cost and allows the user to pay as per the usage. Reduced operating cost and risk towards infrastructure, allows the user to shift their service infrastructure to Cloud, as an outsourcing service. Also, Cloud offers guaranteed benefits such as energy efficiency, resource utilization and flexibility. The characteristics of Cloud and IoT are independent of each other. Generally, IoT and cloud computing are preferred to reduce the storage, processing and communication constraints and obtain the resources virtually which has unlimited capabilities. Cloud provides its extended support to IoT as an intermediate layer and handles the real time distributed and dynamic environment data through its application module. Due to this, the complexity and unnecessary functionalities which impact the data processing is reduced in the applications.

Cloud IoT is used to deliver personalized application to collect and distribute the data at low cost. The communication in cloud environment effectively observe, collect and manage anything irrespective of place and time, through its built-in applications and customized portals. This high-speed network service effectively monitors the remote things and provides a real time access to the data. Cloud significantly improves the IoT communication, but still has few limitations, while handling large amount of data. Figure 1 gives an illustration of cloud IoT.
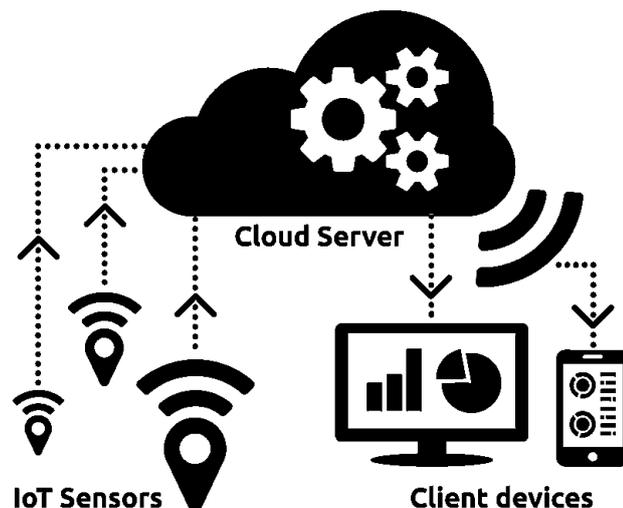
I-SMAC

Fig.1 Cloud IoT architecture

IoT handles large amount of data which is semi structured and non-structured. The three important characteristics of big data are represented as volume which defines the data size, velocity which defines the data generation frequency and variety which represents the types of data. Cloud virtual storage is considered as solution to handle the prolonged life of data at minimum cost based on user demand. Cloud IoT offers better opportunities to data aggregation, data integration as a cost-effective solution to IoT with high level security, so that, data could be accessed virtually from any place. Similarly, the computation in IoT devices does not allow complex data processing due to the limitation of energy resources. Without suitable infrastructure processing, the collected information is a challenging task in IoT. Cloud offers this infrastructure virtually with unlimited processing facilities based on demand which satisfies the real time data analysis in IoT. The adaptation of cloud IoT offers smart services to deal real life service paradigm.

Crowdsourcing is a general-purpose problem-solving process which performs tasks by a group of participants to obtain solution for the given problem. The application of crowdsourcing is not limited into medicine, machine learning, linguistics, software development etc., Figure 2 gives an illustration of crowdsource system for cloud infrastructure. Crowdsensing has one-way communication between the issue owner and the participants. The personal attributes and characteristics of the participants are hidden in nature similarly as black box model, until they complete the given task. In some cases, the communication is processed as two way, since the issue owner needs to obtain the feedback to understand the ground truth. In case of complex tasks, the communication becomes two way to maintain the relationship between the owner and the participants. Crowdsourcing is classified into two categories as collaborative and cooperative. The terms look like similar but the real fact is, collaboration is a coordinated process and the synchronous activity attempt to construct and maintain the origin of the issue. On the other hand, cooperation is an individual process, in which each person is responsible to solve the given issue.
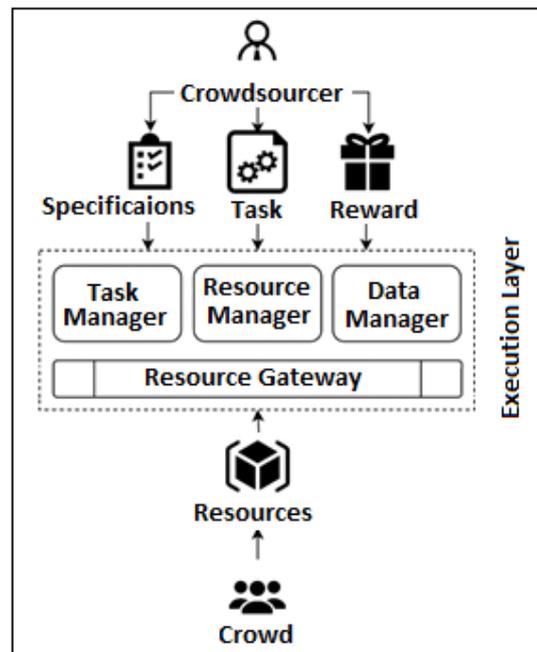
Fig.2 Crowdsource for cloud infrastructure

In collaborative crowdsourcing, the participants have knowledge about the process. The frequent communication of other participants helps them to realize to complete the given task. In cooperative crowd sensing, the process is controlled by computation algorithms, so that the participant has to solve the issue individually. This is because that the other participant contributions are hidden in nature. Based on the given workflow, the members are accomplished to complete the entire issue or a part of the issue to obtain useful results. Generally, crowdsource provider defines the responsibilities as workflow through production manager as planning, control and coordination of process. In order to obtain reliable results, the tasks are divided into subtasks and execute them parallelly or sequentially. Workflow must handle one or more tasks to combine the results and perform quality activities before or after the execution stages and this maintains the quality of the process. The workflow process in crowdsourcing motivated us to implement the crowdsource to improve the performance of cloud IoT module.

2. Related Works

The research towards crowdsourcing and its applications are familiar in recent times. Cloud based IoT applications are introduced in recent times to support smart applications. Duraipandian et al [1] presented a cloud based IoT application model for smart connected objects. Research article provides deep insight about cloud IoT advantages and its impact over recent applications. Similarly Various research works are progressed based on crowdsourcing. Christoforou et al [2] defines crowdsourcing as group of tasks performed by employees with less cost function. To achieve a defined result for the given issue the abilities of peoples are used. Based on the obtained data the decision is taken in crowdsourcing. Qingchen Zhang et al [3] observations indicate that crowdsourcing is used in many applications as a problem solving, classification, research, development, information propagation and decision support model. The advantage of crowdsourcing is in its individual results due to multiple crowd users. Another advantage is its quality of results due to its active participants. The advantages of crowdsourcing using online based applications, improves the human perception and is used as an indicator to perform the tasks.

Fei Chen et al [4] reports the advantages of crowdsourcing in media generation and distribution. Conventional streaming methods includes large number of data channels to minimize the latency. In case of live

streaming the demand towards source needs high computation and better streaming model. The proposed model is developed through cloud computing platform which provides support to live streaming of crowdsourced data in a cost-effective manner and improves the quality of streaming. Chongwu *et al* [5] presented the similar live streaming model in crowdsourcing through multi-CDN-assisted CSLS architecture to improve the quality of service by reducing the operational cost. Yuanhuan *et al* [6] reported the issues in Crowdsourced Live Game Video Streaming in mobile applications. To ensure the quality of service, cost-effective adaptive live streaming model is presented in the research work. The heterogeneity issues and quality affective parameters are formulated as a stochastic optimization issue and by using Lyapunov optimization process the solutions are obtained. This optimization model increases the quality of service with minimum cost compared to other models.

Sabita *et al* [7] presents the advantages of multimedia crowdsourcing and its challenges in data acquisition. Traditional incentive methods particularly designed based on the application and user needs to bring the full potential of crowdsourcing. To reduce the limitation of traditional methods, optimal incentive method is proposed which establishes connection to smartphone users and reward them for their active participation. The results obtained in optimal model is much better than traditional incentive method. The limitation of this optimal model is its increased cost function. Dejun Yang *et al* [8] proposed an incentive method to attract participants in crowdsourcing. As a crowd source centric model, the source provides rewards the participants and also the participants have more control over the received payment. Using a game, the incentive mechanism for crowdsource centric is proposed in the research work to maximize the quality. This method is computationally efficient and trustworthy compared to other incentive models. Yanru Zhang *et al* [9] analysed the issues in incentive methods in mobile crowdsourcing to encourage the user participation in crowdsourcing. Initially one-dimensional performance analysis is performed in order to increase the utility a multidimensional analysis is proposed in the research model which increases the active participants in crowdsourcing.

Crowdsourcing is widely used in various applications. To ensure the security and trustworthiness various security mechanism are evolved in crowdsourcing applications. Kyuwoong Hwang *et al* [10] presented an audio recognition model using crowdsourcing which combines the event, user location, background and phone context. It collects the information from various peoples and generated a model which is accurate. Based on Gaussian Histogram and K- Nearest classifier, the data is prepared and presented to the user with high accuracy. Gaoqiang Zhuo *et al* [11] reports the issues in mobile crowdsourcing in terms of data collection, aggregation and analysis. The large volume data needs suitable filtering and pre-processing strategies to process the raw data. This leads into shortage of resources also leads into privacy issues in an untrusted cloud. To reduce the issues, the proposed model used cloud to preserve the identities and by using batch verification and updating process the computation cost is reduced. Arijit Karati *et al* [12] reported the privacy issues in crowdsourced data through significant cryptographic method. To meet the authenticity and confidentiality requirement, identity based signcryption is introduced using bilinear pairing. The proposed model is suitable for low bandwidth communication and lags in performance over high bandwidth.

Task matching in crowdsourcing is reported in Suma *et al* [13] has discussed a research model to address the privacy issues. Conventional task matching models lags to meet the privacy requirements. The proposed model is developed using blockchain based matching to provide secure matching in cloud environment. The reliability and transparent matching are the advantages of the proposed model. Qingchen Zhang *et al* [14] presented a double projection deep computation model to process big data. Conventional deep computation model is modified to learn the features of bigdata to improve the training efficiency. To improve the data privacy, BGV encryption is used in the proposed model. Compared to conventional DCM models, the performance of proposed model is much better in terms of classification accuracy and privacy.

3. Proposed Model

From the above intense survey, the application of crowdsourcing is analysed. The impact of crowdsourcing in data management and privacy are notable. The existing methods are focused towards a single point objective and not over the quality of services. Based on the observations, proposed model is presented to obtain better QoS in crowdsourced cloud IoT environment. The overview of proposed cloud computing IoT

crowdsourcing model uses IoT clusters to access the activities of the participants in the environment. Figure 3 gives an illustration of proposed model. The data collected through IoT clusters are summarized and processed by a data analyst. To validate the collected information, a two phased process is performed in the proposed model. Initially, it is given into crowd source as a questionnaire to obtain the desired solution. Based on the results, such as candidate explanation and source information the resultant data is processed.
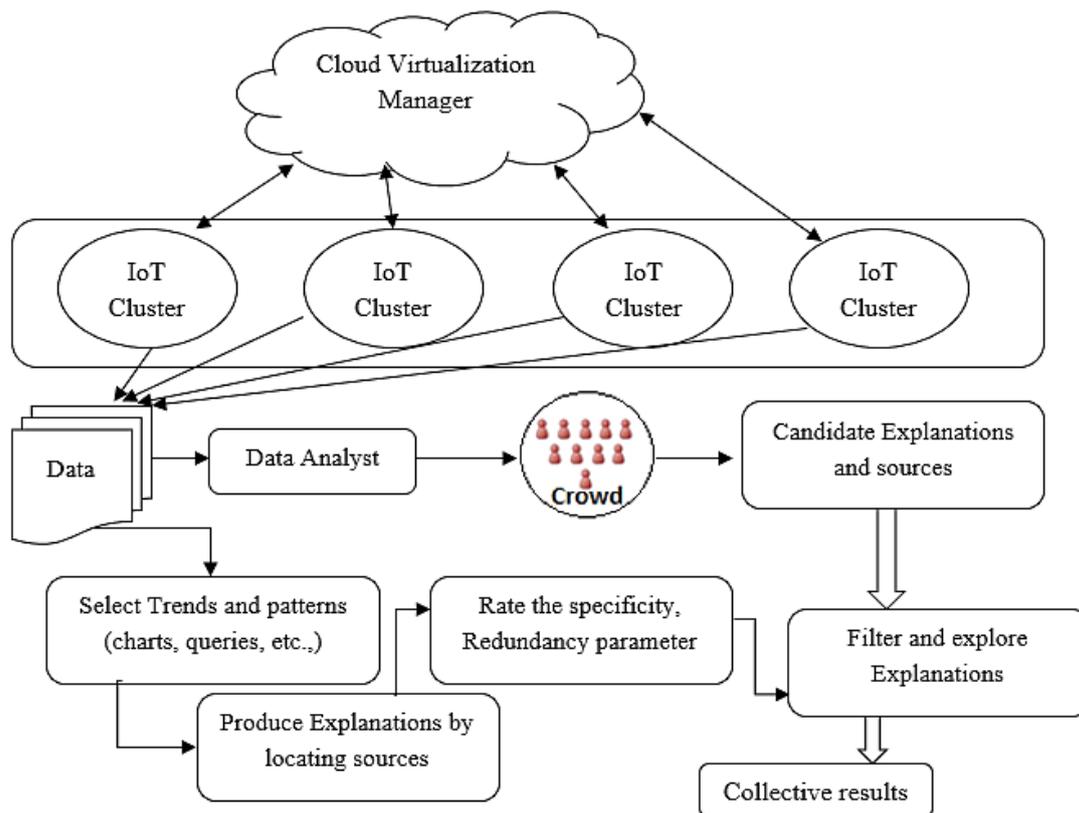

Fig.3 Proposed crowdsourcing model for Cloud IoT

In the second phase, parallel operation is performed to select the trends and patterns similar to query models. Upon providing suitable explanation and source information the details are analysed in terms of specificity and redundancy. This process eliminates the redundant data in the collected information. In the last stage, the analysed information and the processed information are further filtered to obtain the desired explained results. The resultant collected information is used as necessary information in data mining process. To validate the data analysis proposed model is utilized with rank-based model and to manage the cluster progression k-NN clustering scheme is used. Compared to other clustering schemes, k-NN clustering is more efficient and simpler to implement.

4. Results and Discussion

The proposed model is experimented in Cloudsim using ParticipAct data set, which contains campaigns 173 students in Italy on a period of 15 months. To validate the performance metrics of the proposed model, deep convolution and conventional k-NN model are compared in terms of specificity, sensitivity, efficiency and accuracy. The parameters used in simulation is given in table 1.

Table 1 Simulation Parameters

S. No	Parameters	Range
1	Simulation Area	500×500m
2	Signal Range	15m
3	Network Bandwidth	10-100Mbps
4	Mobility	0.5-2m/s
5	Number of service providers	2
6	Number of crowdsourcing environments	1
7	Number of participants	10-20 per cluster
8	Number of clusters	4

The efficiency comparison of the proposed model is depicted in figure 4. It is observed that proposed model attains better efficiency due to its improved rank-based clustering in crowdsourcing compared to conventional machine learning models.
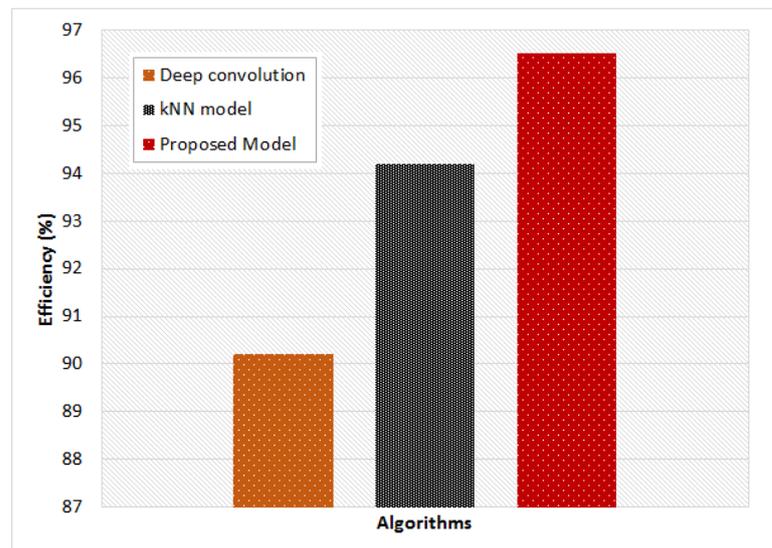

Fig. 4 Efficiency Comparison

The performance metrics of proposed model is validated in terms of specificity, sensitivity to improve the quality of service in cloud IoT environment. Figure 5 gives the details as a comparison of all the three models. It is observed from the figure that, the performance of proposed model is superior to deep convolution model and conventional k-NN model.
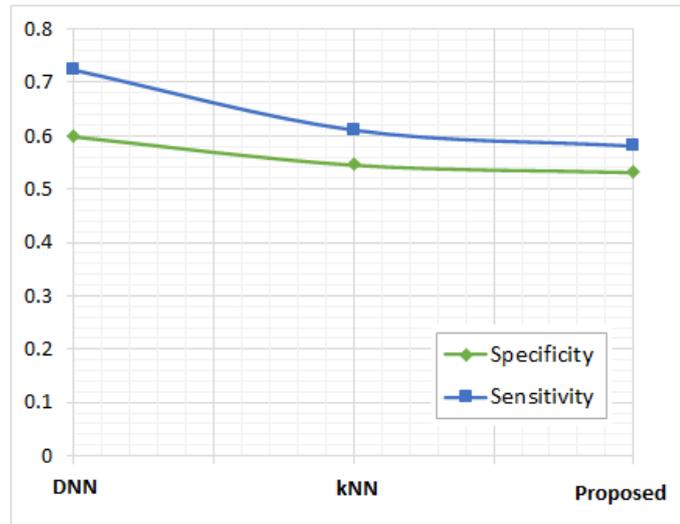
178

I-SMAC

Fig.5 Specificity and sensitivity comparison

The accuracy percentage of proposed model is much better than conventional models. The overall acquired accuracy of all the three models is 82.44%,83.86% and 89.92% respectively. From these observations, it is clear that proposed model achieved better accuracy in data handling process compared to other models.

5. Conclusion

Crowdsourcing is an emerging technology and its impact in cloud IoT environment provides a new perceptive in data handling approaches. Crowd offers solutions to various issues in a simple way. This proposed research work is focused on implementation of crowdsourcing over cloud IoT environment to improve the quality of services. The experimental result highlights the superior performance of proposed model. To measure the performance metrics proposed model is compared to deep convolution model and k-NN model. With accuracy of 89.92% the proposed model attains better improvement in quality of service compared to conventional models.

References

1. Duraipandian, M, Vinothkanna,R (2019). Cloud Based Internet of Things for Smart Connected Objects. *Journal of ISMAC*, 1(02), 111-119.
2. Christoforou,E, Fernández Anta,A., Georgiou, C., Mosteiro, M., and Sánchez, A. 2013. Crowd Computing as a Cooperation Problem: An Evolutionary Approach. Journal of Statistical Physics 151(3), 654-672.
3. Qingchen Zhang, Laurence Tianruo Yang, Zhikui Chen, Peng Li, Fanyu Bu (2019). An Adaptive Dropout Deep Computation Model for Industrial IoT Big Data Learning with Crowdsourcing to Cloud Computing. *IEEE Transactions on Industrial Informatics*, 15(4), 2330-2337.
4. Fei Chen, Cong Zhang, Feng Wang, Jiangchuan Liu, Xiaofeng Wang, Yuan Liu (2015). Cloud-Assisted Live Streaming for Crowdsourced Multimedia Content. *IEEE Transactions on Multimedia*, 17(9), 1471-1483
5. Chongwu Dong, Yin Jia, Hua Peng, Xiaoxing Yang, Wushao Wen (2018). A Novel Distribution Service Policy for Crowdsourced Live Streaming in Cloud Platform. *IEEE Transactions on Network and Service Management*, 15(2), 679-692.

I-SMAC

6. Yuanhuan Zheng, Di Wu, Yihao Ke, Can Yang, Min Chen, Guoqing Zhang (2017). Online Cloud Transcoding and Distribution for Crowdsourced Live Game Video Streaming. *IEEE Transactions on Circuits and Systems for Video Technology*, 27(8),1777-1789.

7. Sabita Maharjan, Yan Zhang, Stein Gjessing (2016). Optimal Incentive Design for Cloud-Enabled Multimedia Crowdsourcing. *IEEE Transactions on Multimedia*, 18(12), 2470-2481.

8. Dejun Yang, Guoliang Xue, Xi Fang, Jian Tang (2016). Incentive Mechanisms for Crowdsensing: Crowdsourcing With Smartphones. *IEEE/ACM Transactions on Networking*, 24(3), 1732-1744.

9. Yanru Zhang, Yunan Gu, Miao Pan, Nguyen H. Tran, Zaher Dawy, Zhu Han (2018). Multi-Dimensional Incentive Mechanism in Mobile Crowdsourcing with Moral Hazard. *IEEE Transactions on Mobile Computing*, 17(3), 604-616.

10. Kyuwoong Hwang, Soo-Young Lee (2012). Environmental audio scene and activity recognition through mobile-based crowdsourcing. *IEEE Transactions on Consumer Electronics*, 58(2), 700-705.

11. Gaoqiang Zhuo, Qi Jia, Linke Guo, Ming Li, Pan Li (2017). Privacy-Preserving Verifiable Set Operation in Big Data for Cloud-Assisted Mobile Crowdsourcing. *IEEE Internet of Things Journal*, 4(2), 572-582.

12. Arijit Karati, SK Hafizul Islam, G. P. Biswas, Md Zakirul Alam Bhuiyan, Pandi Vijayakumar, Marimuthu Karuppiah (2018). Provably Secure Identity-Based Signcryption Scheme for Crowdsourced Industrial Internet of Things Environments. *IEEE Internet of Things Journal*, 5(4), 2904-2914.

13. Suma, V. (2019). Security and Privacy Mechanism Using Blockchain. *Journal of Ubiquitous Computing and Communication Technologies (UCCT)*, 1(01), 45-54.

14. Qingchen Zhang, Laurence T. Yang, Zhikui Chen, Peng Li, M. Jamal Deen (2018). Privacy-Preserving Double-Projection Deep Computation Model With Crowdsourcing on Cloud for Big Data Feature Learning. *IEEE Internet of Things Journal*, 5(4), 2896-2903.