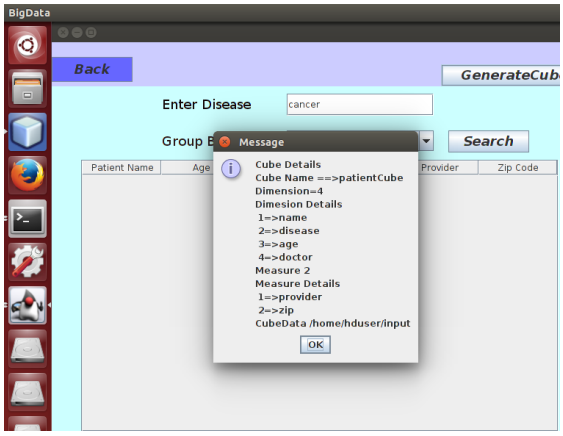

Figure 7: Cube Generation

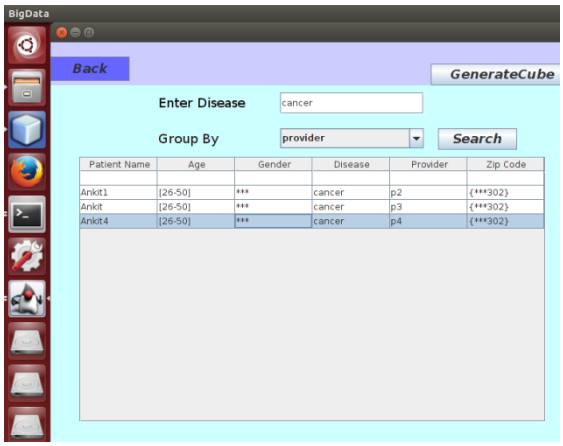

Figure 8: Search Result

6. Conclusion

Here the scalability problem for anonymizing the data on cloud for big data applications by using Bottom Up Generalization and proposes a scalable Bottom Up Generalization. The BUG approach performed as follows, first Data partitioning ,executing of driver that produce a intermediate result. After that, these results are merged into one and apply a generalization approach. This creates the anonymized data. The data anonymization is done using MR Framework on cloud. This indcate that scalability and efficiency are improved significantly over existing approaches.

References

[1] Xuyun Zhang, Laurence T. Yang, Chang Liu, and Jinjun Chen,"A Scalable Two-Phase Top-Down Specialization Approach for Data Anonymization Using MapReduce on Cloud", vol. 25, no. 2, february 2014.
[2] Ke Wang, Yu, P.S,Chakraborty, S, " Bottom-up generalization: a data mining solution to privacy protection".
[3] B.C.M. Fung, K. Wang, R. Chen and P.S. Yu, "Privacy-Preserving Data Publishing: A Survey of Recent Developments," ACMComput. Surv., vol. 42, no. 4, pp.1-53, 2010.
[4] K. LeFevre, D.J. DeWitt and R. Ramakrishnan, "Workload- Aware Anonymization Techniques for Large-Scale Datasets," ACMTrans. Database Syst., vol. 33, no. 3, pp. 1-47, 2008.
[5] B. Fung, K. Wang, L. Wang and P.C.K. Hung, "Privacy- Preserving Data Publishing for Cluster Analysis," Data Knowl.Eng., Vol.68,no.6, pp. 552-575, 2009.
[6] B.C.M. Fung, K. Wang, and P.S. Yu, "Anonymizing Classification Data for Privacy Preservation," IEEE Trans. Knowledge and Data Eng., vol. 19, no. 5, pp. 711-725, May 2007.
[7] Hassan Takabi, James B.D. Joshi and Gail-Joon Ahn, "Security and Privacy Challenges in Cloud Computing Environments".
[8] K. LeFevre, D.J. DeWitt, and R. Ramakrishnan, "Incognito: Efficient Full-Domain K-Anonymity," Proc. ACM SIGMOD Int'l Conf. Management of Data (SIGMOD '05), pp. 49-60, 2005.
[9] T. IwuchukwuandJ.F. Naughton, "K-Anonymization as Spatial Indexing: Toward Scalable and Incremental Anonymization," Proc. 33rdInt'lConf. VeryLarge DataBases (VLDB'07), pp.746-757, 2007
[10] J. Dean and S. Ghemawat, "Mapreduce: Simplified Data Processing on Large Clusters," Comm. ACM, vol. 51, no. 1, pp. 107-113,2008.
[11] Dean J, Ghemawat S. "Mapreduce: a flexible data processing tool," Communications of the ACM 2010;53(1):72–77. DOI:10.1145/1629175.1629198.
[12] Jiuyong Li, Jixue Liu , Muzammil Baig , Raymond Chi-Wing Wong, "Information based data anonymization for classification utility"
[13] X. Xiao and Y. Tao, "Anatomy: Simple and Effective Privacy Preservation," Proc. 32nd Int'l Conf. Very Large Data Bases(VLDB'06), pp. 139-150, 2006.
[14] Raj H, Nathuji R, Singh A, England P. "Resource management for isolation enhanced cloud services," In: Proceedings of the 2009ACM workshop on cloud computing security, Chicago, Illinois, USA, 2009, p.77–84.
[15] K.R.Pandilakshmi, G.Rashitha Banu. "An Advanced Bottom up Generalization Approach for Big Data on Cloud" , Volume: 03, June 2014, Pages: 1054-1059..
[16] Intel "Big Data in the Cloud: Converging Technologies".
[17] Jiuyong Li, Jixue Liu Muzammil Baig, Raymond Chi-Wing Wong, "Information based data anonymization for classification utility".