Volume: 10 Issue: 03 Pages: 3893-3897 (2018) ISSN: 0975-0290

Storage of Mobile Sensor Data in Clouds using Information Classification Algorithms

Prashant Sangulagi
Department of Electronics and Communication Engineering, BKIT Bhalki, INDIA
Email: psangulgi@gmail.com
Ashok V Sutagundar
Department of Electronics and Communication Engineering, BEC Bagalkot, INDIA
Email: sutagundar@gmail.com
Stelvarani S
Department of Electronics and Communication Engineering, BKIT Bhalki, INDIA
Email:stelvaranis@gmail.com

----------**ABSTRACT**----------

Mobiles are equipped with sensors like accelerometer, magnetic subject, and air strain meter, which assist within the system of extracting context of the person like area, scenario and so on. But, processing the extracted sensor facts is generally an aid intensive assignment, which can be offloaded to the general public cloud from mobiles. Mobile devices have become an essential part of our day to day life by which the user is able to access, create and share information at any location. This design especially objectives at extracting beneficial statistics from the accelerometer sensor records. The design proposes the utilization of parallel computing to the use of Map Reduce at the cloud for spotting human behavior primarily based on classifiers and ultimately calculating its accuracy. The sensor facts is extracted from the cellular, sent to the cloud and processed using threepopular classifier algorithms namely, Kernel Naïve Bayes, Naive Byes Classifier and K-Nearest-Neighbors. The results are verified at different scenarios of human activity and finally the accuracy is calculated using the classification algorithms.

Keywords-**Sensor nodes, Cloud Computing, Information Classification, Sensor Data, K-NN Classifier, Naïve Bayes Classifier**

Date of Submission: Aug 30, 2018 Date of Acceptance: Sep 28, 2018

1. INTRODUCTION

The current upgrades in cell technology are fostering the emergence of ubiquitous environments wherein the user is able to get admission to, create and proportion information at any place with large ease. Furthermore, mobile gadgets have grow to be an important a part of allotted architectures that can be used for monitoring the context of the person (e.g. location, state of affairs, and so on.), and thus reacting in a proactive way. Alternatively, smart telephones are equipped with a diffusion of sensors (GPS, magnetic subject, etc.) that increase the mobile packages with region recognition and sensing talents. These advances permit becoming contextual requirements for enhancing the quality of service (QoS) inside the programs because it lets in adapting the interplay between the handset and the user in actual-time.

For example, a sensor along with the accelerometer is used for sensing how the user is conserving the handset, and as a consequence changing the role of the display as a end result. The accelerometer is a sensing detail that measures the acceleration related to the site of a weight in which it has been embedded (e.g. a cellular tool). Relying on the range of axes, it can collect the acceleration records throughout multiple dimensions. Within the case of a tri-axial accelerometer[1] that is the maximum not unusual in mobiles from companies along with HTC, Samsung, Nokia etc., the acceleration can be sensed on three axes (x, y and z). The axes are related with the actions forward/backward, left/right and up/down, respectively. for example: in the case of a runner, up/down is measured because the crouching when he/she is warming up before beginning to run, ahead/backward is related with dashing up and slowing down, and left/proper includes making turns while he/she is jogging. Even as maximum of the actual-global interest popularity structures can depend upon classifiers (Pastime reputation offerings) which might be pre-skilled with big records sets, they require long periods of schooling. Consequently, the more records to be had for education, the more correct the classifier will become, in terms of reputation, detection and prediction. Mobile offerings for hobby recognition are confined to the processing skills of the handset[2], and consequently the cell user cannot enjoy the hobby popularity richness (variety of activities that can be diagnosed) of a pre-educated classifier for automating cell functionality (e.g. calendar alarms, and so on.). To address this trouble, the gift work proposes 1) to persist the each day data amassed from the accelerometer within the cloud, 2) the utilization of parallel computing framework (Map Reduce[3]) at the cloud for training a classifier based on which can effortlessly scale in performance, time and accuracy, 3) to show those alleviate the handset from the time eating invocation of a cloud carrier which will reveal the applicability of our idea, we have evolved an Android application which gathers accelerometer data and offloads it to the cloud[4]. Later, by using imposing parallel variations of acknowledged algorithms (bays, selection

Volume: 10 Issue: 03 Pages: 3893-3897 (2018) ISSN: 0975-0290

timber, etc.), the classifiers are pre-educated, published as a carrier at the cloud and consumed for cellular interest recognition.

Paper is organised into different sections given by, Section 2 gives, history and review of sensor classifiers on mobile technologies. Section 3 depicts methodology of the previous work and Section 4 describes proposed work, implementation using MATLAB and simulation results in section 5. Finally Section 6 concludes the paper.

2. LITERATURE OVERVIEW

Previously work has been done related to sensor cloud and human activity recognition with respect to cloud. Some of them are as follows. The most extensive work, where they used 5 biaxial accelerometers on multiple parts of the frame for the popularity of activities (inclusive of eating, sitting, studying and so on.) the use of a set of facts accrued from users, beneath real-world occasions and without researcher supervision. In their paintings, hobby reputation become executed the use of selection tables, instance-based totally learning (IBL) or nearest neighbor, and Naive Bayes strategies [2]. It defines the evaluation of low decision accelerometer data for continuous human interest popularity as a type trouble, considering support vector machines. They advocate a technique for incorporating smoothing class temporally which can be easily coupled with any classifier at a minimal attempt; they declare this approach increases the classification accuracy by using 3% [5].

It additionally tackles the human hobby popularity as a type hassle thinking about several algorithms. In addition, they introduce a mechanism for amassing the statistics the usage of clever phones and outside accelerometers which aren't embedded in the device. Shuangquan W, additionally utilize type algorithm however introduce a mechanism to collect the statistics the usage of Wi-Fi sensor networks [6] & [7]. This communicate will first introduce the improvement of cell cloud computing and describe a few packages concerning multimedia, imaginative and popularity, snap shots, gaming, textual content processing. Next, it's going to gift the transmission, computation, and sensing demanding situations of green computing in cell cloud [8]. In this design, they observe previous paintings on the naïve Bayesian classifier and evaluate its limitations, which include sensitivity to correlated capabilities. They respond to this hassle via embedding the naive Bayesian induction scheme inside an algorithm that contains out a grasping seek thru the space of capabilities [10].

3. METHODOLOGY

Large studies has been performed approximately category algorithms and how they improve the accuracy and the performance within the recognition method. Throughout the schooling phase a set of rules has to method huge quantities of labelled facts to analyze the patterns and rules which in particular are hidden. Once the set of rules is skilled, it is ready for predicting and classifying unknown label records. Most often the training segment is time and resource disturbing and the accuracy, overall performance and effectiveness of the algorithm are at once dependent on the dimensions of the training facts set. The greater label facts is used for education, the faster and better predictions are finished, but also the extra time and assets are required all through the schooling segment. In programs whose aim is the prediction and recognition of human activity patterns, the facts used for training functions calls for to be accumulated in situations as close as possible to actual situations. In fashionable, statistics collected in laboratories does no longer incorporate the pastime styles that human beings utilize of their day-to-day life. Moreover, collecting training statistics for such human interest pattern discovery is an intricate system, particularly while it entails embedded sensors inclusive of the accelerometer in cellular devices. To address these troubles, we utilize the embedded accelerometer sensor in clever telephones for accumulating the statistics in real eventualities and cloud offerings for storing training repositories which are further categorised using parallel computing additionally in the cloud. The person helps in gathering real accelerometer records that is further used for training and prediction functions through the classifiers within the cloud. The classifiers have been defined in this proposed work. The data will be sensed by the sensor and sensed information initially sent to sink node and the unaltered data will be sent to classifier. The classifier will use one of the classification method to remove the redundant data and accept only related information for further processing hence reduces the latency problem and saves the bandwidth. The processed information is stored in the cloud with priority based approach. The General architecture of Wireless sensor network is shown in fig 1 where the human activity is sensed and sensed data is store in the cloud server.

1.1. Cloud-based Sensor Classifier for the Recognition of Human Activities:

Even as the accelerometer can music information for the popularity of a couple of human sports (walking, going for and walks, so forth.) for a right away reaction, each pastime is differentiated in keeping with its behavior and entropy.

Figure 1. Wireless Sensor Networks

As already cited, the accelerometer provides information across time related with acceleration alongside x, y, z axes. This records can be used for the identity and category of certain styles that are described in the standards

Volume: 10 Issue: 03 Pages: 3893-3897 (2018) ISSN: 0975-0290

as; status, walking and strolling. However, for appearing such evaluation three class algorithms primarily based on Hadoop are used in this observe. Hadoop is a framework [9] that gives assist for the analysis of data-in depth dispensed programs over hundreds of nodes, inside the cloud. Algorithms are carried out for education the classifier based totally on choice bushes, Naive Bays approach and candidate neighbors. Each algorithm is implemented using MATLAB, Hadoop 0.20 and Python because the programming language [10]. The three classifiers are used they are KNN, Naïve Bays and Kernel Naïve Bays.

1.1.1Naive Bays classifier

The Naive Bays classifier [11] offers simplicity by means of the usage of Bays theorem for probabilistic [12] class. This set of rules is based in the assumption that the presence or absence of each event is unbiased from the presence or absence of different activities. The algorithm consists of two stages: education and classification. The schooling level takes as enter a set of categorized information, where in every point is represented as a tuple <xi.xn, activity >, and calculates the opportunity of every characteristic xi given the interest class j. Later, the m thing is calculated dividing the conditional possibility Pi j(xi/class j) over the conditional possibility of Pik(xi/class k) ∀ the lessons inside the set Cs. similarly, the class step, given a tuple <xi.xn> calculates the score that's the multiplication of the mi ∀ the attributes xi. If the score happens to be more than percentage P (class j/class k) then the hobby are classed as class k else class j. on this algorithm the education step is a good candidate to be computed in clusters; consequently, Map Reduce is appropriate for implementing such step.

1.1.2 Kernel Naïve Bayes Classifier

The Kernel Naive Bayes is same as Naïve Bayes classifier in this a kernel is a weighting function used in non-parametric estimation strategies. Kernels are used in kernel density estimation to estimate random variables' density features, or in kernel regression to estimate the conditional expectation of a random variable. Kernel density estimators belong to a class of estimators referred to as non-parametric density estimators. In contrast to parametric estimators in which the estimator has a hard and fast useful shape (structure) and the parameters of this feature are the most effective statistics we want to store, Non-parametric estimators have no constant shape and rely on all of the records points to reach an estimate. The advantage of the Naive Bayes classifier is that it is requires a small quantity of training statistics to estimate the means and variances of the variables important for class. Because independent variables are assumed, most effective the variances of the variables for each label need to be decided and not the complete covariance matrix. In evaluation to the Naive Bayes operator, the Naive Bayes (Kernel) operator can be carried out on numerical attributes.

1.1.3 K-nearest neighbor's algorithm

K-nearest neighbor's algorithm (k-NN) is a non-parametric method used for classification and regression [1].In both cases, the input consists of the k closest training. The K-Nearest-Neighbors classifies items within the foundation of the distances among unclassified items and labeled objects that exist in the same space. This approach considers a schooling set of labeled factors unfold in a space. Later while an object is to be categorized the approach calculates the gap among the unclassified object and the alternative items within the space, clustering the nearest objects to the unlabeled item collectively. Once the method determines the nearest objects to the unclassified one it takes the okay-closest points and votes with a purpose to label the unclassified item.

4. DESIGN AND IMPLEMENTATION

The design proposes the utilization of parallel computing the use of Map Reduce at the cloud for schooling and spotting human activities primarily based on classifiers that may without problems scale in overall performance and accuracy. The sensor facts is extracted from the cellular, offloaded to the cloud and processed using distinctive class algorithms, Naive Byes Classifier and okay-Nearest-neighbors, the main code is shown in figure 2.

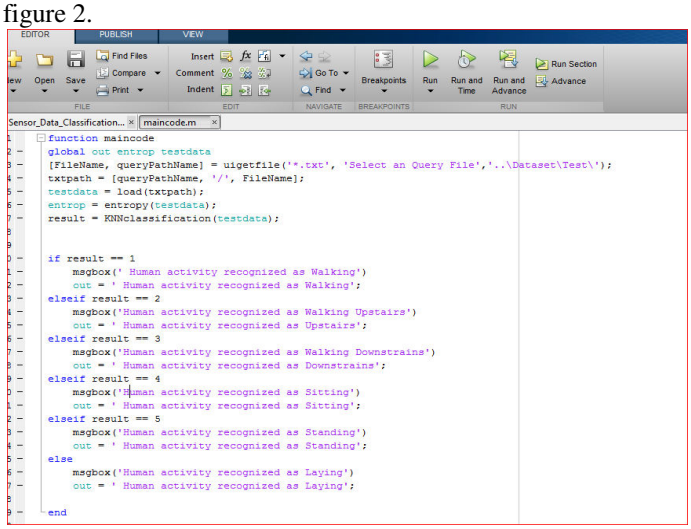

Figure 2. Main the code for recognition of human activity using Matlab

The sensor data is offloaded to the cloud and data is sensed. When the classifier is simulated then it will pop out a message showing data sensing is done, the initial stage of the process which is shown in fig 3.The simulation toolbox have been prepared to classify the sensed data using all the classifiers, KNN, NB & KNB and finally each classifier will show the accuracy and histogram of the popped up result which is shown in figure 4. Than data is classified using KNN by the data set and data is also classified by the Naïve Bayes and gives the accuracy.

Volume: 10 Issue: 03 Pages: 3893-3897 (2018) ISSN: 0975-0290

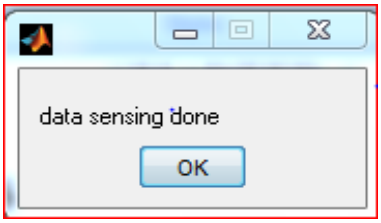

Figure 3. Data sensing.

5. RESULTS AND DISCUSSIONS

Various nodes have been deployed to monitor the human activity, like temperature sensor, pressure sensor etc. These sensors will send the information to the sink node after every one minute of time. Later the data have been sent to cloud server. The data have been classified using the classification algorithms, namely KNN algorithm, Naïve bayes and Kernel naïve bayes algorithms.

The data been classified based on the human regular activity like, sitting, walking, walking upstairs, lying and standing as shown in fig 5, 6, 7, 8, 9 and 10. The histogram of the activity is recognized from the dataset and result has been taken out.

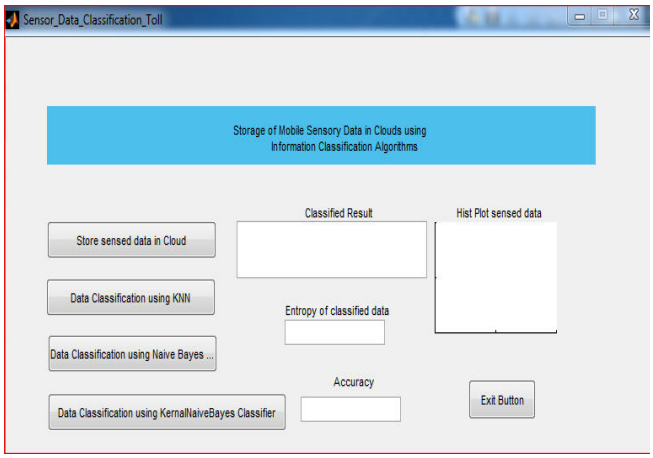

Figure4. Classified Result block

Waveforms of the human activity

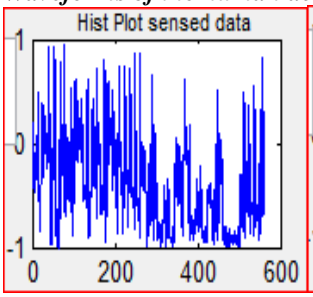

Figure 5.Walking

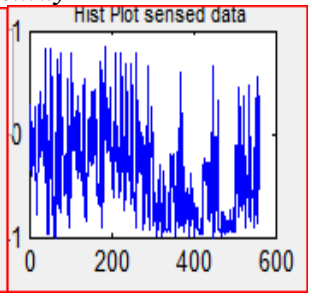

Figure 6.Walking upstairs

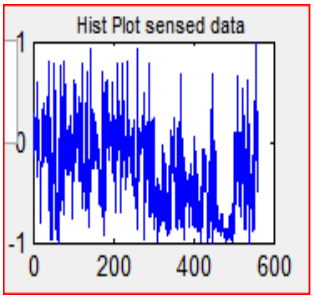

Figure 7.Walking downstairs

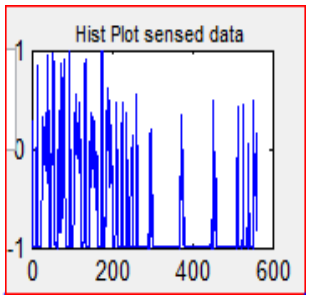

Figure 8.Sitting

Here we are comparing the classifiers between them which one is good; we are using three classifiers KNN, Naive Bayes and kernel naive bayes. Between them we are checking the accuracy and delay time. When we give the inputs in section wise we get the output of the classifiers with accuracy and delay time in seconds.

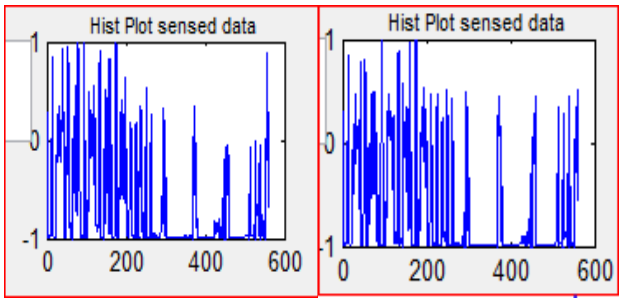

Figure 9. StandingFigure10. Lying

By comparing the output of classifiers with their accuracy we finally conclude that the kernel naive bayes is much better than K-NN by observing the above comparison table 1.

The graph of iterations versus accuracy is as shown in below figure 11. Each classifier is operated with 1000 data and each time the accuracy is calculated by applying all three classifiers using MATLAB. The results shows that, the *Kernel* Naïve bayes showing better accuracy in all the iterations but delay is typically more than KNN which is as shown below.

TABLE 1. COMPARISON TABLE

Classifier	K-NN		Naïve Bays		Kernel NB	
No. of I/P	Accuracy	Delay in sec	Accuracy	delay in sec	Accuracy	Delay in sec
500	94.2	1.8	86	1.4	92	96
500	92.2	1.77	88	1.36	96	92
1000	91.8	1.06	90	1.38	94	94
1000	90.7	1.38	92.7	1.69	98.6	176.3
1000	89.28	1.22	95.2	1.5821	97.9	171.9
1000	91.87	1.05	95.2	1.606	97.9	172.07
1000	89.2	1.85	88.1	1.584	92	182.92
1000	93.28	1.6	94.2	1.63	96.2	182.6
1000	94.2	1.67	82.7	1.6	91.8	182.09
1000	93.7	1.81	89.43	1.8	94.1	214

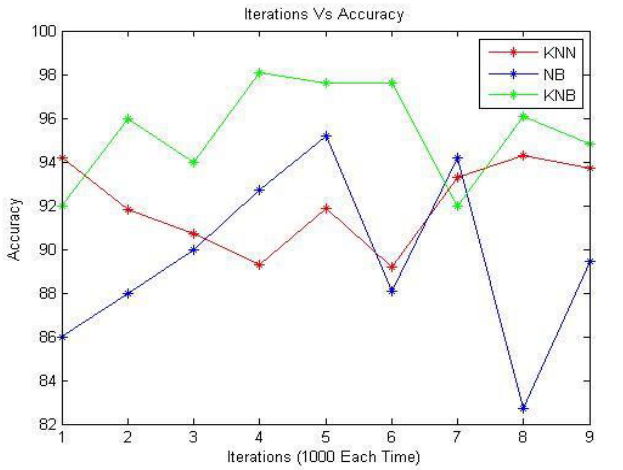

Figure 11. Iterations Vs Accuracy

Volume: 10 Issue: 03 Pages: 3893-3897 (2018) ISSN: 0975-0290

6. CONCLUSION

The design describes the utilization of parallel computing, use of Map Reduce on the cloud for education and recognition of human activities from accelerometer sensor statistics extracted from mobiles, primarily based on classifiers which can without difficulty scale in performance and accuracy. The design considers 3 distinctive nature algorithms, Kernel Naïve Bayes, Naive Bayes Classifier and K-NN. The Map Reduce based algorithms are mentioned in element. The overall performance evaluation of the algorithms suggests that the popularity of activities is feasible with sizable overall performance latencies. The category also blessings from the inherent functions of the cloud like elasticity and horizontal scalability. The final accuracy of the KNN using different application as walking, sitting, standing, etc is around 92.02% and the final accuracy of kernel naive bayes using different application is around 95.8% and the accuracy of naïve bayes classier is around 90.15%. But kernel naïve bayes consumes some time with respect to KNN, with accuracy wise the kernel naïve bayes is good with respect to KNN. So it can be conclude that the kernel naive bays classifier is the good one for sensing human activity using the cloud.

REFERENCES

[1] B. Stenbock Andersen, C. Mangeot, 3-axial accelerometer, *US Patent Ap*p. 20,100/043,552 (Mar. 26 2008).

[2] L. Bao, S. Intille, Activity recognition from user-annotated acceleration data, *Pervasive Computing* (2004) pp.1–17.

[3] J. Dean, S. Ghemawat, Mapreduce: Simplified data processing on large clusters, *Communications of the ACM* 51(1) (2008) pp.107–113.

[4] H. Flores, S. Srirama, C. Paniagua, A generic middleware framework for handling process intensive hybrid cloud services from mobiles, *Proceedings of the 9th International Conference on Advances in Mobile Computing and Multimedia*, ACM, (2011), pp. 87–94.

[5] N. Krishnan, S. Panchanathan, Analysis of low resolution accelerometer data for continuous human activity recognition, *Acoustics, Speech and Signal Processing, ICASSP-2008. IEEE International Conference on, IEEE*, (2008), pp.3337–3340.

[6] N. Ravi, N. Dandekar, P. Mysore, M. Littman, Activity recognition from accelerometer data, *Proceedings of the National Conference on Artificial Intelligence*, Vol. 20, Menlo Park, CA; Cambridge, MA; London; AAAI Press; MIT Press; (1999), pp.1541-1545.

[7] Diego Antolín, Nicolás Medrano, Belén Calvo and Francisco Pérez, A Wearable Wireless Sensor Network for Indoor Smart Environment Monitoring in Safety Applications, Sensors (Basel). (2017) Feb; 17(2), pp.1-14.

[8] Blesson Varghese, Rajkumar Buyya, Next generation cloud computing: New trends and research directions, *Future Generation Computer Systems*, 79, 2018, pp.849–861.

[9] Er. Prachi Jain, Er. Alisha, A brief review on Hadoop architecture and its issues, *International Journal of Engineering Research and General Science*, 5(2), (2017), pp.211-217.

[10] Shuo Xu, Bayesian Naïve Bayes classifiers to text classification, *Journal of Information Science*, 44(1) (2016), pp.48-59.

[11] P. Langley, Stephanie Sage, Induction of selective bayesian classifiers., *Tech. rep., DTIC Document* (1994), pp.399-406.

[12] I. Rish, An empirical study of the naive bayes classifier, *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*, Vol. 3, No. 6. pp.56-58.