Available online on 21.02.2019 at http://jddtonline.info

Journal of Drug Delivery and Therapeutics

Open Access to Pharmaceutical and Medical Research



Open Access Research Article

# Novel Algorithm of CPU-GPU hybrid system for health care data classification

**Safdar Tanweer, Naseem Rao***

Assistant Professors, CSE Department, Hamdard University, Delhi, India

## ABSTRACT

Due to advancements in portable health monitoring technology, such systems have become more and more economical & efficient. This in turn has resulted in a huge amount of data being generated every moment by millions of users of such portable devices. Such voluminous data may include audio, video, and image, and text representing blood pressure, temperature, vocal activity, ECG, sugar level etc. In the Proposed algorithm, first step is assignment, where clusters are assigned to a patient data and the second step is update, which takes the mean of the coordinates of all the data in its cluster. Medical practitioners and service providers can use such data to discover various patterns and useful insights. Such insights can be very useful on understanding various trends during epidemics, such as Malaria, Dengue, Chikungunya and other such outbreaks. A faster and economical way to get such insights is of paramount importance.

**Keywords:** health monitoring; GPU; ECG; epidemics; data mining

**Article Info:** Received 11 Jan 2019; Review Completed 19 Feb 2019; Accepted 19 Feb 2019; Available online 21 Feb 2019

**Cite this article as:**

Tanweer S, Rao N, Novel Algorithm of CPU-GPU hybrid system for health care data classification, Journal of Drug Delivery and Therapeutics. 2019; 9(1-s):355-357 **DOI: http://dx.doi.org/10.22270/jddt.v9i1-s.2445**

***Address for Correspondence:**

Naseem Rao, Assistant Professors, CSE Deptt., Hamdard University, Delhi, India

## INTRODUCTION

Data for health monitoring are received from various healthcare sources such as wearable sensors [1]. Such sensors include sensors for monitoring patient's health parameters, medical tests, and treatment. Normally such data is generated several megabytes per second. Further because there are so many parameters, it is difficult to quickly analyze the data in order to make important decision regarding patient health. Furthermore because of advancements in wearable sensor technology, such systems have become more and more economical & efficient, which has further increased amount of data generated every moment [2-3]. Such voluminous data may include audio, video, and image, and text representing blood pressure, temperature, vocal activity, ECG, sugar level, Oximetry etc. To make sense out of this data, medical practitioners and service providers can apply various data mining algorithm, to discover various patterns and useful insights. Such insights can be very useful on understanding various trends during epidemics, such as Malaria, Dengue, Chikungunya and other such outbreaks [4-10].

## MASSIVELY PARALLEL ARCHITECTURE OF GPU

NVIDIA Graphics Processing Units (GPUs) and other modern GPUs has huge computational power due to a major portion of silicon devoted for computational units called Cores. A core mainly consists of ALU (Arithmetic Logical Units) for performing arithmetic computations and logical operations and other Special Function Units such as for Tanh, Sine etc. These hug number of ALUs and Special Function Units implemented in hardware working in parallel computes much faster as compared to a traditional CPU architecture [5].

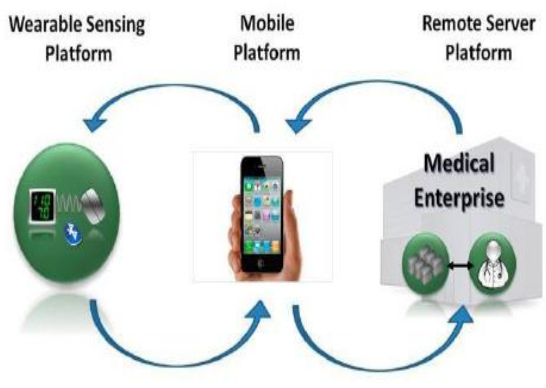


Figure 1: Generic architecture of mhealth monitoring system

Fig shows the layout of the silicon on the chip in case of CPU and GPU. It can be seen that most of the silicon area in GPU is devoted for ALU, shown as small green boxes, while in CPU most of this area is devoted for control and cache. We can see that a GPU provides an economical supercomputing platform for masses! Though there are certain problems such as cache coherence problem. But various solutions have been proposed to solve it [5].

Obviously our algorithm should have parallelism to exploit such computational power. This can be understood from the Algorithm-Hardware Matching perspective as shown in Fig 2.

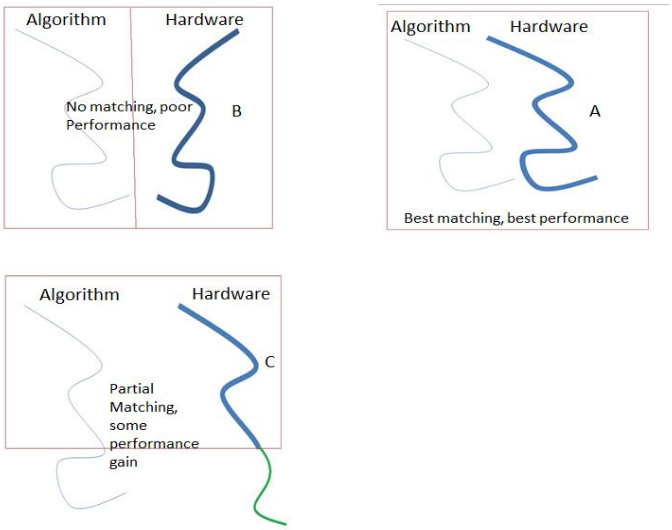


Figure 2: Algorithm-Hardware Matching

It can be seen that in 2(B) and 2(C), the algorithm's pattern and hardware layout is not matching perfectly; therefore we cannot expect much performance gain. But in 2(A) both are matching 100%. This situation is the most favorable from performance point of view. The situation in 2(B) is worst because algorithm and hardware are orthogonal to each other. In simple words we can say, for example, that in the algorithms, when an addition is about to be done, the adder unit is not available in hardware; instead a floating point unit is free. But when in algorithm a floating point calculation is required, the floating point unit is not available (or occupied), instead the adder is available. This is an orthogonal situation we mentioned earlier. The Algorithm presented below discusses more on such issue and data flow.

## ALGORITHM

Our algorithm runs on a CPU-GPU hybrid system. In such a hybrid system information flows from CPU to GPU and then back to GPU in a systematic way. This is shown in Fig 3 below:

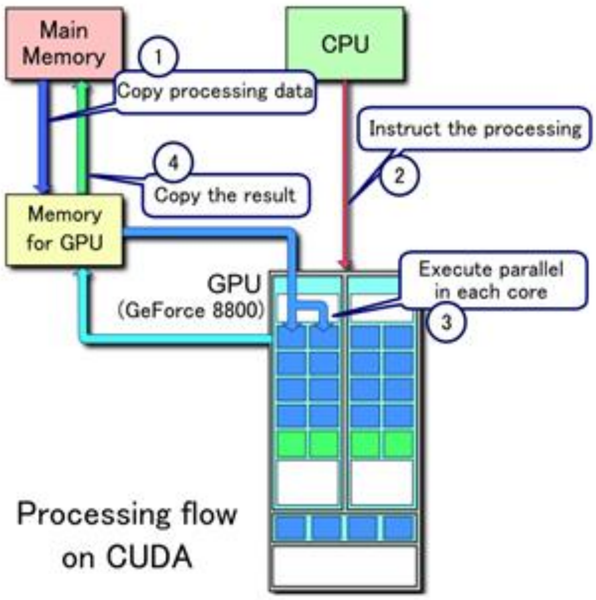


Figure 3: Information flow in CPU-GPU hybrid system

We took the latest work in this area as discussed in [7] as a base. The authors have tried to improve the k-means algorithm from only algorithmic point of view. K-means can provide patterns which can be very useful for making good decisions [8]. While in this work algorithmic as well as architectural point of view is explored. That is we have to make the algorithm suitable for the underlying hardware. The importance of matching algorithm and hardware has already been explained in above section.

The K-means algorithm is iterative [7-8] and unsupervised. It gradually converges to a final value. First we supply some initial guesses for the health care data classification. Then the algorithm will continue iterating until the healthcare data mapped to clusters do not change the clusters to which they belong. Basically the algorithm consists of two steps:

1 The first step is assignment, where clusters are are assigned to a patient data.

2 The second step is update, which takes the mean of the coordinates of all the data in its cluster.

**The following pseudocode gives a general idea of how the algorithm works:**

```
while no healthcare data (points) have changed assigned health group(dataset) :
// Assignment
for each point (or healthcare data point ) in dataset (health group):
point.cluster = nearest cluster
// Update healthcare data group to which this data site belong
for each cluster in k:
for each point in dataset:
if point.cluster == cluster:
cluster.position += point
cluster.position /= count of points in cluster
```

We identified that the distance calculations are preferable to GPU architecture. That is an embarrassingly parallel algorithm. There is inherent parallelism in the algorithm that can be exploited for this purpose.

## CONCLUSION

In this work we demonstrated how to use economical platforms having a GPU can help various medical practitioners, decision makers etc, to understand and discover useful trend from a healthcare data. In the Proposed algorithm, first step is assignment, where clusters are assigned to a patient data and the second step is update, which takes the mean of the coordinates of all the data in its cluster. Further explored more properties of GPU to come up with still better results, which can help make further better decisions, especially in case of various epidemics, including Dengue and Chikungunya etc.

## REFERENCES

[1] Koh H. C., Tan G.Data mining applications in healthcareJournal of Healthcare Information Management, 2005.

[2] Reyes-Ortiz J.-L., Oneto L, Sama A, Parra X, Anguita D, “Transition-aware human activity recognition using smartphones,” Neurocomputing, vol. 171, pp. 754–767, 2016.

[3] Mahdavinejad MS, Rezvan M, Barekatain M, Adibi P, Barnaghi P, Sheth AP, “Machine learning for Internet of Things data analysis: A survey,” Digit. Commun. Netw. 2017.

[4] Patel S, Park H, Bonato P, Chan L, Rodgers M, A review of wearable sensors and systems with application in rehabilitation, Journal of NeuroEngineering and Rehabilitation, 2012.

[5] Zheng J, Zhang Z, Wu T, Zhang G, Emerging Wearable Medical Devices towards Personalized Healthcare , Conference: Proceedings of the 8th International Conference on Body Area Networks, September 2013.

[6] Tanweer S, Mobin A, Alam A, “Environmental Noise Classification using LDA, QDA and ANN Methods”, Indian Journal of Science and Technology, 2016; 9(33) , September. ISSN (Print) : 0974-6846 ISSN (Online) : 0974-5645

[7] Singh I, Shriraman A, WL Fung W, O'Connor M, Aamodt TM, Cache coherence for GPU architectures, IEEE 19th International Symposium on High Performance Computer Architecture (HPCA2013), 2013.

[8] Haraty RA, Dimishkieh M, Masud M, An Enhanced k-Means Clustering Algorithm for Pattern Discovery in Healthcare Data, International Journal of Distributed Sensor Networks vol. 11,June 2015.

[9] Analysis of K-Means and K-Medoids Algorithm For Big Data, Procedia Computer Science, Volume 78, 2016, Pages 507-512 in 1st International Conference on Information Security & Privacy, 2015.

[10] Bottou L., Tesauro BY, Touretzky D. Convergence properties of the k-means algorithms Advances in Neural Information Processing Systems New York, NY, USAMIT,1995.