International Journal of Research
Available at https://edupediapublications.org/journals
e-ISSN: 2348-6848
p-ISSN: 2348-795X
Volume 04 Issue-17
December 2017

Improving Computational Efficiency in Personalized Healthcare Data by Implementing Sparse Matrices

Macha Deepa & K. Vivek

[1]M. Tech Student, Department of CSE, Vignan's Lara Institute of Technology & Science, Village Vadlamudi, Mandal Chebrolu, District Guntur, Andhra Pradesh, India.

[2]Assistant Professor, Department of CSE, Vignan's Lara Institute of Technology & Science, Village Vadlamudi, Mandal Chebrolu, District Guntur, Andhra Pradesh, India.

ABSTRACT— *A collection of large and complex data sets which are difficult to process using common database management tools or traditional data processing applications. Big data is not just about size and also finds insights from complex, noisy, heterogeneous, longitudinal, and voluminous data. In this paper we are describing various problems that we believe need to be tackled in order to have an effective integration of big data analytics and Virtual Physiological Human (VPH) modeling in healthcare. And also in this paper we are extending our solution by implementing sparse matrix to enhance computational efficiency of big data analysis.*

Keywords: Big Data, Big data Analysis, Virtual Physiological Human Modeling

1. INTRODUCTION

Big data is a relative term describing a situation where the volume, velocity and variety of data exceed an organization's storage or compute capacity for accurate and timely decision making. Some of this data is held in transactional data stores – the byproduct of fast-growing online activity. Machine-to-machine interactions, such as metering, call detail records, environmental sensing and RFID systems, generate their own tidal waves of data. All these forms of data are expanding, and that is coupled with fast-growing streams of unstructured and semi structured data from social media. However, big data is defined less by volume – which is a constantly moving target – than by its ever-increasing variety, velocity, variability and complexity.

Today the healthcare industry is undergoing one of the most important and challenging transitions to date, the move from paper to electronic healthcare records. While the healthcare industry has generally been an incrementally advancing field, this change has the potential to be revolutionarily. Using the data collected from these electronic records exciting tools such as disease recommendation systems have been created to deliver personalized models of an individual's health profile. However despite their early success, tools such as these will soon encounter a significant problem. The amount of healthcare encounter data collected is increasing drastically, and the computational time for these applications will soon reach a point at which these systems can no longer function in a practical timeframe for clinical use.

Big Data as it pertains to health care has emerged at the center of the revolution in personalized medicine. Simply put, the proliferation of data offers great possibilities for more precise diagnosis, as researchers are able to drill

down to see what's happening and create more targeted therapies, specifically at the molecular and tissue levels.

The second half of the twentieth century can be regarded as the era of reductionism in biology, when, driven by revolutions in molecular biology and genomics, the dominant paradigm saw biology as the sum of its parts: the genes and proteins that make up each organism. Most biologists, however, now recognize that an integrative approach – an understanding of how these parts work together in a complex entity – is as important as a reductionist one. The relatively new discipline of systems biology combines this philosophy with an integration of experimental biology with computational biology. One of the most ambitious goals of systems biology is that of modeling the entire human physiology. Human body can be broken down into a series of interlocking organs and systems, with one of the most tractable to model being the heart. Simple mathematical models of cardiac ion channel action have developed over half a century into complex simulations that are being used successfully in drug discovery. Integrating models together, however, requires close international collaboration.

2. RELATED WORK

Eugene V. McCloskey, Helena Johansson, Anders OdenJohn and A. Kanis introduced FRAX algorithm. FRAX is a computer-based algorithm that provides models for the assessment of fracture probability in men and women. The approach uses easily obtained clinical risk factors to estimate 10-year fracture probability, with or without femoral neck Bone Mineral Density (BMD), to enhance fracture risk prediction. It has been constructed using primary data from population-based cohorts around the world. The gradients of fracture risk have been validated in independent cohorts with a similar geographic distribution. The FRAX tool should not be considered as a gold standard, but rather as a platform technology on which to build as new validated risk indicators become available. Notwithstanding, the present models provide an aid to enhance patient assessment by the integration of clinical risk factors alone and/or in combination with BMD. This article describes the steps undertaken in the development of FRAX.

Vinicius C. Rispoli, Jon F. Nielsen , Krishna S. Nayak and Joao L. A. Carvalho proposed a framework for obtaining flow field estimates that are influenced by both PC-MRI measurements and a fluid physics model. The results showed that the proposed technique provides better agreement with the PC-MRI measurements than pure CFD simulations, and has reduced computation time (faster convergence). MRI-guided CFD can be used to correct the MRI-measured flow field, forcing it to satisfy the fluid mechanics equations. It can also be used as a means of reducing noise in the PC-MRI measurements, and has potential as a method for reducing scan time. The proposed framework offers a general approach to in vivo blood flow assessment that is complementary to improvements in PC-MRI acquisition and reconstruction techniques, and can be applied to the study and diagnosis of a broad range of cardiovascular flow mapping applications.

Shi et al. studied that a simplified method algorithm was employed to calculate FFRCTA; we observed good correlation and an acceptable mean difference between FFR_{CTA} and invasive FFR, as well as a better diagnostic performance of FFR_{CTA} in diagnosing ischaemia-causing stenosis in the clinic. By implementing this new boundary condition, the simplified FFR_{CTA} calculated with pulsatile

flow has the potential to be an alternate and accurate diagnostic parameter in the assessment of the haemo dynamic characteristics for coronary stenosis.

3. VIRTUAL PHYSIOLOGICAL HUMAN

In recent years major advances have been made in understanding certain aspects of the human body, as a complex system. Computer 'models' exist which can reproduce the complexities of a beating heart, for example, and furthermore tell you what will happen to that heart if it is exposed to a certain drug or kind of trauma. Models of other parts of the body describing kidney function, joint and limb mechanics, and blood flow (amongst others) all exist. The challenge faced by researchers is how to connect these models to create a more effective and realistic simulation or the human body, both in health and disease. It is the goal of the VPH initiative to create the research environment in which progress in the investigation of the human body as a single complex system may be achieved. VPH models of the human body are both descriptive and predictive. They are formed by large collections of anatomical, physiological, and pathological data stored in digital format, by predictive simulations developed from these collections, and by services aimed to support the researchers in the creation and maintenance of these models. They also include services aimed to empower clinical, industrial and societal users in the use of VPH-related information. As such coordination of the VPH Initiative is required on many levels: from how the models work and connect, to how data is stored and accessed, to who can use and develop these models, and how they will be used to improve healthcare and drug discovery.

4. IMPLEMENTATION OF SPARSE MATRIX FOR BIG DATA ANALYSIS

Analysis of big data on large-scale distributed systems often necessitates efficient parallel graph algorithms that are used to explore the relationships between individual components. Graph algorithms use the basic adjacency list representation for graphs, which can also be viewed as a sparse matrix. This correspondence between representation of graphs and sparse matrices makes it possible to express many important graph algorithms in terms of basic sparse matrix operations, where the literature for optimization is more mature.

In the case of a sparse matrix, substantial memory requirement reductions can be realized by storing only the non-zero entries. Depending on the number and distribution of the non-zero entries, different data structures can be used and yield huge savings in memory when compared to the basic approach. The trade-off is that accessing the individual elements becomes more complex and additional structures are needed to be able to recover the original matrix unambiguously. Sparse matrices also have significant advantages in terms of computational efficiency. Unlike operations with full matrices, operations with sparse matrices do not perform unnecessary low-level arithmetic, such as zero-adds. The resulting efficiencies can lead to dramatic improvements in execution time for programs working with large amounts of sparse data.

5. BIG DATA FOR COMPUTATIONAL BIOMEDICINE

The majority of big data applications deal with data that do not refer to an individual person. This does not exclude the possibility that their aggregated information content might not be socially sensitive, but very rarely is it

possible to reconnect such content to the identity of an individual. In the cases where sensitive data are involved, it is usually possible to collect and analyze the data at a single location; so this becomes a problem of computer security; within the secure box, the treatment of the data is identical to that of non sensitive data. Healthcare poses some peculiar problems in this area. First, all medical data are highly sensitive, and in many developed countries are considered legally owned by the patient, and the healthcare provider is required to respect patient confidentiality.

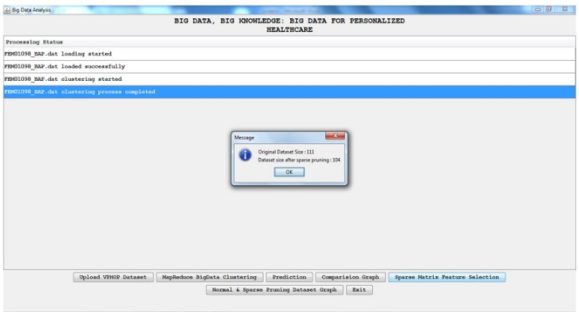
Genomics and postgenomics technologies produce very large amounts of raw data about the complex biochemical processes that regulate each living organism; nowadays a single deep sequencing dataset can exceed 1 TB. More recently, we have started to see the generation of "deep phenotyping" data, where biochemical, imaging, and sensing technologies are used to quantify complex phenotypical traits and link them to the genetic information.

Enormous progress has been made in the integration of image processing and engineering analysis, with many applications in healthcare across the spectrum from orthopaedics to cardiovascular systems and often multiscale models of disease processes, including cancer, are included in these analyses. Very efficient methods, and associated workflows, have been developed that support the generation of patient-specific anatomical models based on exquisite three and four-dimensional medical images. The major challenge now is to use these models to predict acute and longer-term physiological and biological changes that will occur under the progression of disease and under candidate interventions, whether pharmacological or surgical. There is a wealth of data in the clinical record that could support this, but its transformation into relevant information is enormously complex.

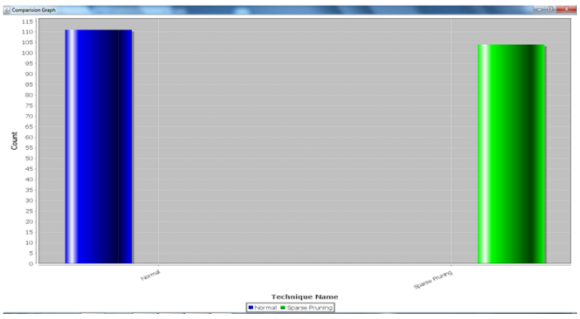
6. EXPERIMENTAL RESULTS

In this experiment, we used VHOP dataset to parse dataset. Here, first we need to upload the dataset and we can cluster the dataset by using MapReduce big data clustering. The status of the processing we can view on the processing status screen. After completion of the clustering, we can perform the prediction on another uploaded file as well we can see the comparison chart for classification results.

Next we can run the Sparse Matrix Feature selection, then it will return the before sparse dataset size as well after sparse dataset.



We can see the comparison between the normal pruning of the dataset and sparse pruning of the dataset.



7. CONCLUSION

In this paper, we studied about the big data analytics that is although sometimes overhyped, big data technologies do have great potential in the domain of computational biomedicine, but their development should take place in combination with other modeling strategies, and not in competition. This will minimize the risk of research investments, and will ensure a constant improvement of in silico medicine, favoring its clinical adoption. However, we need to improve the computational efficiency of the big data analysis in personalized health data. Hence we implemented sparse matrix in this paper to reduce the computational complexity of the big data in healthcare systems.

REFERENCES

[1] D. Laney, "3D data management: Controlling data volume, velocity, and variety," META Group, 2001.

[2] O. Terzo, P. Ruiu, E. Bucci, and F. Xhafa, "Data as a service (DaaS) for sharing and processing of large data collections in the cloud," in Proc. Int. Conf. Complex Intell. Softw. Intensive Syst., 2013, pp. 475–480.

[3] B. Wixom, T. Ariyachandra, D. Douglas, M. Goul, B. Gupta, L. Iyer, U. Kulkarni, J. G. Mooney, G. Phillips-Wren, and O. Turetken, "The current state of business intelligence in academia: The arrival of big data," Commun. Assoc. Inform. Syst., vol. 34, no. 1, p. 1, 2014.

[4] J. Ginsberg, M. H. Mohebbi, R. S. Patel, L. Brammer, M. S. Smolinski, and L. Brilliant, "Detecting influenza epidemics using search engine query data," Nature, vol. 457, no. 7232, pp. 1012–1014, 2009.

[5] J. Manyika, M. Chui, B. Brown, J. Bughin, R. Dobbs, C. Roxburgh, and A. H. Byers, "Big data: The next frontier for innovation, competition, and productivity," Mckinsey Global Inst., 2011.

[6] J. W. Fenner, B. Brook, G. Clapworthy, P. V. Coveney, V. Feipel, H. Gregersen, D. R. Hose, P. Kohl, P. Lawford, K. M. McCormack, D. Pinney, S. R. Thomas, S. Van Sint Jan, S. Waters, and M. Viceconti, "The EuroPhysiome, STEP and a roadmap for the virtual physiological human," Philos. Trans. A Math. Phys. Eng. Sci., vol. 366, no. 1878, pp. 2979–99, Sep. 2008.

[7] F. C. Horn, B. A. Tahir, N. J. Stewart, G. J. Collier, G. Norquay, G. Leung, R. H. Ireland, J. Parra-Robles, H. Marshall, and J. M. Wild, "Lung ventilation volumetry with same-breath acquisition of hyperpolarized gas and proton MRI," NMR Biomed., vol. 27, pp. 1461–1467, Sep. 2014.

[8] 10] A. Brandts, J. J. Westenberg, M. J. Versluis, L. J. Kroft, N. B. Smith, A. G. Webb, and A. de Roos, "Quantitative assessment of left ventricular function in humans at 7 T," Magn. Reson. Med., vol. 64, no. 5, pp. 1471– 1477, Nov. 2010.

[9] E. Schileo, E. Dall'ara, F. Taddei, A. Malandrino, T. Schotkamp, M. Baleani, and M. Viceconti, "An accurate estimation of bone density improves the accuracy of subject-specific finite element models," J. Biomech., vol. 41, no. 11, pp. 2483–91, Aug. 2008.

[10] P. Lamata, S. Niederer, D. Barber, D. Norsletten, J. Lee, R. Hose, and N. Smith, "Personalization of cubic Hermite meshes for efficient biomechanical simulations," Med. Image. Comput. Comput. Assist. Interv., vol. 13, no. Pt 2, pp. 380–387, 2010