Identification of Abdominal Aorta Aneurysm Using Ant Colony Optimization Algorithm

A. Dinesh Kumar[1], R. Nidhya[2], V. Hanah Ayisha[3], Vigneshwar Manokar[4], and Nandhini Vigneshwar Manokar[5]

[1,2] Department of Computer Science and Engineering,
Dr. N.G.P. Institute of Technology, Coimbatore, Tamil Nadu, India
{dineshngpit,nidhuraji88}@gmail.com
[3] Department of Computer Science and Engineering,
MEA College of Engineering, Kerala, India
[4] Department of Computer Science and Engineering,
Faculty of Engineering, Karpagam University, Coimbatore, Tamil Nadu, India
vignesh.manohar@gmail.com
[5] Applied Research – Modeling and Simulation,
Technologies Division, Bharath Corporate, India

Abstract. Abdominal aortic aneurysm (AAA) is a localized dilatation of the abdominal aorta. It occurs when there is a increase in the normal diameter of the blood vessels by more than 50 percent. Approximately 90 percent of abdominal aortic aneurysms occur infrarenally, but they can also occur pararenally or suprarenally. This is because of some catastrophic outcome. Due to this, the blood flow is exaggerated so the blood hemodynamic interaction forces are affected. Therefore this will tends to wall rupture. To identify the AAA, it is important to identify the blood flow interaction and the wall shear stress. The blood and wall interaction is the wall shear stress. Computational fluid dynamics (CFD) is used to get the results for the mechanical conditions within the blood vessels with and without Aneurysms. CFD contains vast computations with Navier Stroke Equations so this will be very time consuming. So to make these CFD computations very efficient, Data mining algorithms are to be used. And also DM algorithms will be a best method to predict the shear stress at the AAA. This will estimate the wall shear stress. There is in need of thousands of CFD runs in a single computer for creating machine learning data so grid computing can be used.
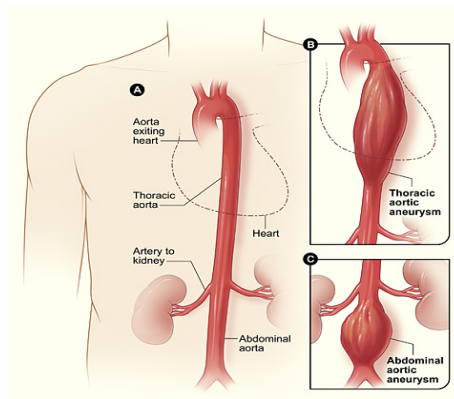
Keywords: Computational fluid dynamics (CFD), data mining (DM), grid computing, hemodynamic parameters, predictive modeling, Ant Colony Optimization (ACO) algorithm.

1 Introduction

Cardiovascular disease is a class of diseases that involve the heart or blood vessels. Cardiovascular disease refers to any disease that affects the cardiovascular system. Cardiovascular diseases remain the biggest cause of deaths worldwide. Age is

an important risk factor in developing cardiovascular diseases. It is estimated that 87 percent of people who die of coronary heart disease are 60 and older. Among these diseases the stenosis process is the most dangerous one. This is a very tedious process which will lead to stroke and aneurysm and now days this is commonly present in patients. In this paper, we are giving importance to Abdominal Aorta Aneurysm (AAA) development that may cause rupture and fatal outcome. An aneurysm is an abnormal widening or ballooning of a portion of an artery due to weakness in the wall of the blood vessel. It is a localized, blood-filled balloon-like bulge in the wall of a blood vessel. Aneurysms can commonly occur in arteries at the base of the brain and an aortic aneurysm occurs in the main artery carrying blood from the left ventricle of the heart. When the size of an aneurysm increases, there is a significant risk of rupture, resulting in severe hemorrhage, other complications or death. Aneurysms can be hereditary or caused by disease, both of which cause the wall of the blood vessel to weaken. Aneurysms may be classified by type, location, and the affected vessel. Other factors may also influence the pathology and diagnosis of aneurysms. A consensus definition of an aneurysm was established in 1991 by the Society of Vascular Surgery and the International Society for Cardiovascular Surgery as a permanent localized dilatation of an artery having at least 50% increase in diameter compared with the expected normal diameter of the artery, or of the diameter of the segment proximal to the dilatation.

To predict the AAA many patient specific studies have been done, from this it's concluded that maximum stress within the vessel wall was more appropriate criterion than maximum diameter and also the shear stress on aneurysm wall is found. Shear stress is a frictional force produced by blood flow, affects biology and structure of the wall. There are many methods to predict this.

Fig. 1. Detailed View of Abdominal Aorta Aneurysm

The following paragraph describes about the methods along with its disadvantages. First, the Shear stress can be obtained by the Computational Fluid dynamics (CFD). This approach has been employed to study wall-shear stress distribution on idealized models of blood vessels. CFD uses numerical methods and algorithms to solve and

analyze problems that involve fluid flows. CFD involves the Navier Stroke Equations and this CFD involves many computations. So this method is very time consuming. Since there are uncertainties in the patients it is difficult to characterize geometric variability using a small number of recorded parameters[2]. Second, Statistical assessment to identify the relationship between flow patterns and geometric attributes. The main concept is to construct the probabilistic models for the input parameters uncertainties that give a reliable output of interest very quickly, without classical CFD calculations[3-5]. An example of this idea is reported by Kolachalama who used Bayesian–Gaussian process emulator to generate a relationship between geometric parameters and maximal wall shear stress (MWSS), and to identify geometries having maximum and minimum of the wall shear stress (WSS)[7]. Third, Statistical Analysis as the Monte Carlo simulation technique. In this method computer runs for the generated random input values and the resultant data is post processed to estimate the output statistics. CPU requirement for a large sample size, this approach becomes computationally prohibitive, particularly when high-fidelity models are used[6].

Fourth, it uses data Mining (DM) model that can be used to avoid maximum time taken by the CFD simulations. The DM model uses data mining algorithms where the input is the aneurysm shape, determined by several independent geometric parameters, while the output parameters are the MWSS over the aneurysm for peak systolic flow, and MWSS over full heart cycle. Also, the shear stress averaged over the aneurysm region and over the heart cycle can show which aneurysms are likely to rupture so that the risk can be properly quantified. The aim of this paper is to introduce Data Mining Algorithm for detecting AAA for multiple CFD runs. An Ant Colony algorithm is used for the training and testing of results.

2   Materials and Methods

The basic task in our approach was first to feed the DM system with necessary data acquired from the CFD simulations and then to perform machine learning. We ran a number of various aneurysm models in order to "teach" the DM system for producing the most accurate predictions. As in any CFD modeling task, there are a number of issues to be considered: geometry, boundary conditions, initial conditions, mesh generation, etc. We are using the CFD finite element (FE) models, which differ only in geometry and rely on random values of ten parameters.

Only the AAA have to be considered among various kinds of artery aneurysms. The shape of the AAA is defined by two splines with half-circle extrusions between them, as shown in Fig. 2 We divide the geometric parameters into three groups.

1) *aneurysm length, A, B, C*—variable parameters;
2) *upleft, up, upright, downleft, down, downright*— quantifying the curvature in the Bezier description fashion, variable parameters;
3) *length, aorta diameter*—parameters considered are constant, and these are taken from the literature for typical AAA, thus this will give ten independent parameters having a ranges as shown in Table 1.
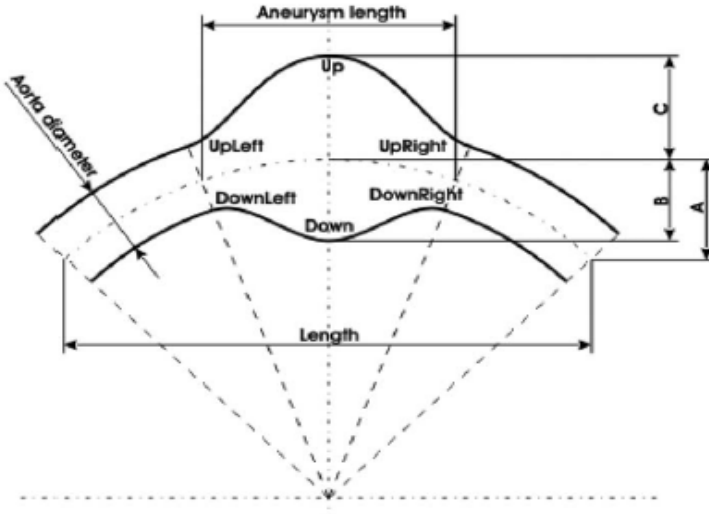
Fig. 2. Parametric model of the aneurysm geometry[1]

Table 1. Ranges of AAA Geometrical Parameters[1]

Parameter	Range
Aneurysm Length (cm)	6 to 15
A (cm)	0 to 3
B (cm)	1.4 to 3
C (cm)	1.4 to 3
Curvature parameters (UpLeft, ...)	0.15 to 0.4

2.1   CFD Model of an Aneurysm

The next step in the prebuilding models for the CFD analysis. This includes the specification of boundary conditions. The identical boundary conditions are prescribed for all combinations of the variable input parameter values. We used an equivalent length at the aneurysm outlet to model the resistance to the blood flow. Other relevant quantities with fixed values are blood density ρ = 1.05 g/cm3, kinematic Viscosity v = 0.035 cm2 /s, length = 24 cm, and aorta diameter D = 2 cm. CFD module generates finite element models for each CFD run. All data are then transferred to Storage Element (SE) from User Interface (UI). Then, each CFD instance reaches destination worker node anywhere in the grid, determined by the Workload Management System (WMS). This procedure significantly saves time, enabling a large number of runs to be executed concurrently (Taken from Reference). The fundamental equations for flow of a viscous incompressible fluid (such as blood) are the Navier–Stokes equations[8].

$$\left(\frac{1}{\Delta t}M_v + {}^{t+\Delta t}K_{vv}^{(i-1)} + {}^{t+\Delta t}K_{\mu v}^{(i-1)} + {}^{t+\Delta t}K'_{vv}^{(i-1)} + {}^{t+\Delta t}J_{vv}^{(i-1)} + K_{\lambda v}\right)\Delta v^{(i)}$$
$$= {}^{t+\Delta t}F'^{(i-1)}_v$$

These Navier–Stokes equations are used to compute the CFD. For the 3-D model of blood flow, eight-node finite element is used with linear interpolation of velocities while the pressure, taken to be constant over the element, is eliminated by a penalty parameter[9]. The incremental iterative form of the FE equilibrium equations. Where the matrices and vectors are defined in a standard FE manner. As an illustration of the CFD solution, the velocity field (left panel) and pressure distribution (right panel) for the peak systole $t/T=0.16$ of an AAA with $D/d=2/75$ (D is aneurism diameter), $d = 12.7$ mm. After initializing input parameters and creating appropriate FE mesh, a specialized CFD FE analysis module performs hemodynamic simulation. Among these quantities, the DM system only considers the wall shear stress values. A total number of models (and therefore FE meshes) constructed using random parameter values and automatic mesh generator was 6000.
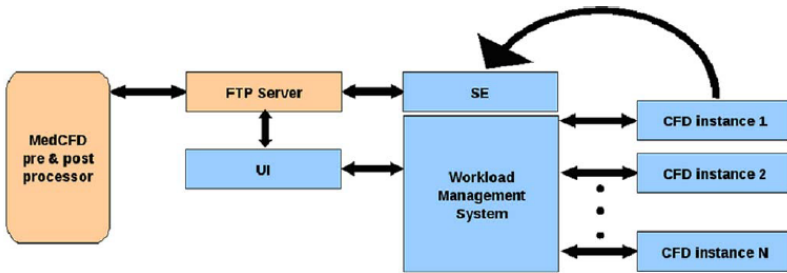

Fig. 3. Grid computing scheme[1]

Since a huge number of CFD finite element analyses had to be executed (for 6000 different geometries based on ten variable parameters) it will take more time (takes even days). So Grid Computing is used to reduce the computation time. Since each CFD run takes around 20 min on a typical personal computer, simple computational time estimation gives around 80 days to run on a single CPU. Equivalent run on a grid platform took 5 h only, while the infrastructure utilization peak during that run was around 600 CPUs at a time.

2.2   DM Approach

After obtaining data from a number of FE runs machine learning process should be done. For this purpose Ant Colony Optimization (ACO) Algorithm is used. The ACO algorithm looks for the minimum of the error function in weight space using the method of gradient descent. The combination of weights which minimizes the error function is considered to be a solution of the learning problem. Since this method requires computation of the gradient of the error function at each iteration step, we must guarantee the continuity and differentiability of the error function.

Training set = all training cases;
         WHILE (No. of cases in the Training set > max_uncovered_cases)
         i=0;REPEAT i=i+1;
Anti incrementally constructs a classification rule;
Prune the just constructed rule;
Update the pheromone of the trail followed by Anti;
UNTIL (i ≥ No_of_Ants) or (Anti constructed the same rule as the previous
No_Rules_Converg-1 Ants)
         Select the best rule among all constructed rules;
         Remove the cases correctly covered by the selected rule from the
         training set;
END WHILE

Ant Colony Optimization (ACO) Algorithm.

We have used the regression model for finding τ systolic (maximum shear stress for the systolic phase) and τ max (maximum shear stress value for the entire cycle domain). Fig. 4 shows these steps the first step is to create additional transformations to the dependent variables, in our case τ systolic and τ max. Sometimes, it is easier to create a regression if the dependent variable Y is first transformed as $\log(Y)$ or $1/Y$. Therefore, in addition to building models for Y, the Bus automatically builds models for $\log(Y)$ and $1/Y$. The second step is to sort the dependent variable in the ascending order and take every tenth out for testing. The third step is to perform variable selection. As a result, it produces up to five different subsets of features, including a solution where all variables are selected. The fourth step is to build regression models
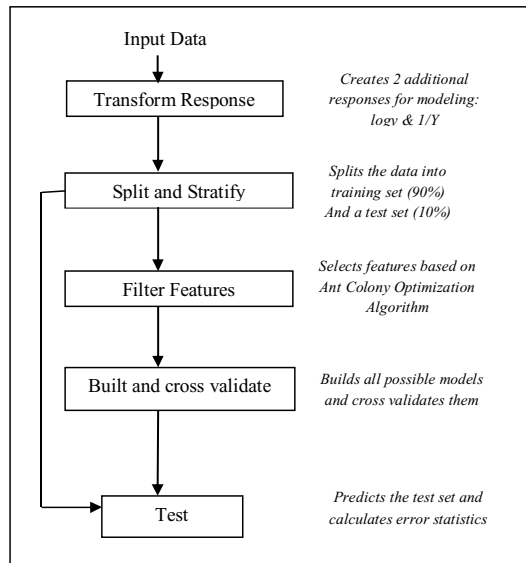
Fig. 4. Overview of the main stages in building and testing

and to cross validate them. All regression models are at the end cross-validated (tenfold) and the following statistics is calculated.

1) Tenfold cross-validated RMSE—root mean square error after tenfold cross validation.
2) Tenfold cross-validated $Q2$—coefficient of determination after tenfold cross validation.
3) Tenfold cross-validated RelSE—relative squared error after tenfold cross validation.
4) Size of the training dataset (Ntr)

The last step is to apply all regression models to the test set and following statistics are calculated.

1)   RMSE—root means square error.

$$RMSE_{train} = \sqrt{\frac{1}{N_{train}} \sum (y_i - y'_i)^2} \qquad RMSE_{test} = \sqrt{\frac{1}{N_{test}} \sum (y_i - y'_i)^2}$$

2.   R2—coefficient of determination.

$$R^2{}_{test} = 1 - \frac{\Sigma(y_i - y'_i)^2}{\Sigma(y_i - y'_{test})^2}$$

y_i ith observed value in the training dataset;
y'_i ith predicted value in the training dataset;
y'_{test} *mean of the observed* values in the testing dataset;
N_{train} number of compounds in the training dataset;
N_{test} number of compounds in the training dataset
From these values we have to produce a performance metrics in which the error rate should be a minimum. RMSE can have an error of 0.001 and R^2 as 0.9.

3   Results and Discussion

From 85 models for property τ systolic and the best model was found for τ −1 systolic. It is a feed forward neural network model that uses all ten input variables. From 82 models for property τmax and the best model was found for τ−1max. It is a feed forward neural network model that uses all ten input variables. The models show a very low error statistics on the test and training sets. In order to better quantify the DM method, we also analyzed an AAA model based on the patient-specific clinical data. The best model among 85 models produced by ACO algorithm (feed forward neural network that uses all ten input variables) has been chosen in order to obtain the best approximation for τ systolic and τ max. As expected, the most realistic model from the first approach consumes the most computing time. The values of τmax and τ systolic were taken as reference values for the parametric and DM models. It can be observed that all relative errors are below 10%, with significant computing time reduction in favor of the DM model.
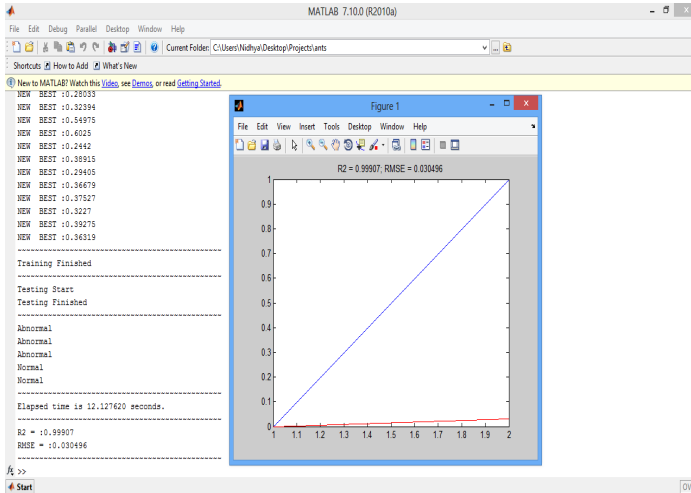
Fig. 5. Experimented Stimulated Output in MATLAB

4   Conclusion

This paper provides a better performance for detecting the aneurysm. In normal methods, the computation takes more than 80 days to run in a single computer and by introducing Grid Computing it has reduced to 5 hours. The CFD calculations were performed using the FE method and the shear stress was evaluated. From these results, it can be stated that a new direction for reducing computational time for patient-specific AAA modeling is suggested. Also, this approach of coupling the computer modeling and DM method (Ant Colony Optimization Algorithm) can further facilitate the development of predictive diagnostic system for clinical practice. For future enhancement efficient data mining algorithm can be used to provide improved accuracy and performance.

References

1. Filipovic, N., Ivanovic, M., Krstajic, D., Kojic, M.: Hemodynamic Flow Modeling Through an Abdominal Aorta Aneurysm Using Data Mining Tools, vol. 15(2) (March 2011)
2. Dobrin, P.B.: Distribution of lamellar deformation implications for properties of the arterial media. Hypertension 33, 806–810 (1999)
3. Taylor, T., Yamaguchi, T.: Three-dimensional simulation of blood flow in an abdominal aortic aneurysm—Steady and unsteady flow cases. ASME J. Biomech. Eng. 116, 89–97 (1994)
4. Finol, E., Amon, C.: Secondary flow and wall shear stress in three dimensional steady flow AAA hemodynamics. In: 2001 Adv. Bioeng., IMECE2001/BED-23013, ASME BED, vol. 51 (2001)

5. Kumar, R., Yamaguchi, T., Liu, H., Himeno, R.: Numerical simulation of 3D unsteady flow dynamics in a blood vessel with Multiple aneurysms. In: 2001 Adv. Bioeng., ASME BED, vol. 50, pp. 475–476 (2001)
6. Rubinstein, R.Y.: Simulation and the Monte Carlo Method. Wiley, Hoboken (1981)
7. Kolachalama, V.B., Bressloff, N.W., Nair, P.B.: Mining data from hemodynamic simulations via Bayesian emulation. BioMed. Eng. Online 6, 47 (2007) 1475-925X-6-47.J
8. Kojić, M., Filipović, N., Stojanović, B., Kojić, N.: Computer Modeling in Bioengineering. Wiley, Hoboken (2008)
9. Filipovic, N., Mijailovic, S., Tsuda, A., Kojic, M.: An implicit algorithm within the arbitrary lagrangian-eulerian formulation for solving incompressible fluid flow with large boundary motions. Comput. Methods Appl. Mech. Eng. 195, 6347–6361 (2006)