ULAB JOURNAL OF SCIENCE AND ENGINEERING

Vol. 7, November 2016

(Contents Continued from Front Cover)

Quality of ICT Facilities at the Tertiary Level Education in Bangladesh: Public versus Private University 41
H M Jahirul Haque and Faisal Mohammad Shahriar

A Note for Contributors 48

Copyright Form 49

ULAB
UNIVERSITY OF LIBERAL ARTS BANGLADESH

House 56, Rd 4/A @ Satmasjid Road Dhanmondi, Dhaka-1209, Bangladesh
Phone: 966-1255, 966-1301, 0171-309-1936. Web: www.ulab.edu.bd

ULAB JOURNAL OF SCIENCE AND ENGINEERING

A RESEARCH PUBLICATION OF ULAB

November 2016 Vol. 7 ISSN 2079-4398 (Print)
ISSN 2414-102X (Online)

CONTENTS

Editorial 1

Numerical Study on the Stability of Finite Difference Schemes for Solving Advection Diffusion Equation 2
T M A K Azad and L S Andallah

Environmental Problems of Bangladesh: A Review 11
Nishita Ivy

On the Independence of Sample Mean and Variance 17
Anwar H. Joarder and M. Hafidz Omar

The appropriate grid size of N layer loosely coupled hexagonal model for equilibria of Boltzmann equation 22
Wahida Zaman Loskor

Optimizing SQL Performance in a Parallel Processing DBMS Architecture 29
Nayem Rahman and Leonard Sutton

(Contents Continued on Back Cover)

School of Science & Engineering
UNIVERSITY OF LIBERAL ARTS BANGLADESH

G P

Editorial
Vol. 7, 2016

WE are pleased to present the seventh volume of the ULAB's Journal of Science and Engineering (JSE). It's been another successful year for us. Since 2010, ULAB's JSE has maintained its position as the most prestigious national publication in the field of science and engineering. We feel proud that the journal has crossed the national boundary and becomes international, as now we are getting papers from abroad.

In the first volume, we found many ordinary papers were being submitted, but the quality of submitted paper is gradually improving in the subsequent volumes. This year on the basis of peer-review we have accepted 6 papers.

The first paper of this volume presents a numerical study with two explicit finite difference schemes for solving the advection diffusion equation along with stability analysis. In addition, it did a simulation to show accuracy of the schemes.

The second paper focuses on the deterioration of environment of Bangladesh to provide an overall scenario of major environmental problems for improving its environmental quality.

The third paper shows two new proofs of the independence of sample mean and variance for two independently, identically and normally distributed random variables with a view to improve the confidence of students and instructors about the fundamental theorems of statistics.

The fourth paper presents a simulation for the equilibrium analysis of Boltzmann equation through a N-layer loosely coupled hexagonal model.

The fifth paper discusses the techniques of optimizing SQL performance in a parallel processing DBMS architecture with a view to empower SQL developers and business intelligence community to write efficient queries for maintaing a stable data warehousing environment.

Finally, the sixth paper focuses on the students' perception about the quality of the ICT facilities available in public and private universities of Bangladesh for a comparative study through utilizing scientific methodology.

According to the publication policy all papers submitted to this journal have been subject to a rigorous peer-review. We continuously strive to publish original research that contains elements with technical novelty in a timely manner. The journal's focus is on traditional areas of both theoretical and practical applications of physics, mathematics, statistics, environmental science, electronics, computer science, information and communication engineering. In addition, we shall gladly accept submissions on emerging technologies and other emerging areas related to the above fields.

You are most welcome to read this issue of the ULAB Journal of Science and Engineering. In order to continue publishing a high-quality journal, JSE's editorial board seeks excellent contributions containing original research or reviews. Our editorial board works tirelessly to provide contributors with a prompt and thorough review process.

We would like to extend our heartfelt thanks to every author, reviewer and reader for your support and dedication to JSE. We strongly believe that together, we shall elevate the journal to even higher levels of quality, impact, and reputation.

Mohammad Shorif Uddin
Editor-in-Chief

Rezaul Karim Mazumder
Editor

Sazzad Hossain
Editor

Anwar H. Joarder
Editor

Numerical Study on the Stability of Finite Difference Schemes for Solving Advection Diffusion Equation

[1]T M A K Azad and [2]L S Andallah
[1]Department of Computer Science & Engineering,
University of Liberal Arts Bangladesh, Dhaka-1209,
Email: abul.kalam@ulab.edu.bd
[2]Department of Mathematics,
Jahangirnagar University, Savar, Dhaka 1342, Bangladesh

Abstract

This paper presents a numerical study with two explicit finite difference schemes FTBSCS (forward time backward space and centered space) and FTCS (forward time and centered space) for solving the advection diffusion equation (ADE). Stability conditions for the schemes are studied, and numerical experiments are performed by applying the stability criterions obtained in this study. Error comparisons with analytical solutions of ADE are presented graphically to show the accuracy of the schemes.

Keywords—Advection Diffusion Equation, Explicit Finite Difference Schemes, Stability condition.

© University of Liberal Arts Bangladesh
All rights reserved.

Manuscript received on 10 July 2016 and accepted for publication on 30 August 2016.

1 Introduction

The most general statement of conservation of contaminant mass in a control volume subject to advective and diffusive flux across its boundaries is $\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} = D\frac{\partial^2 C}{\partial x^2}$, where c(x,t)=solute concentration, D=solute diffusivity, and u= mean flow velocity assumed to be constant. The equation assumes incompressible ambient fluid, and adopts Fick's law of simple proportionality between diffusive contaminant and the concentration gradient. It is a parabolic type partial differential equation, and is derived on the principle of conservation of mass using Fick's law. This equation reflects physical phenomena where in the diffusion process particles are moving with certain velocity from higher concentration to lower concentration. This process is described by the right hand term of the Advection diffusion equation. Second and right hand terms represent the concentration of the contaminant particles as respect to the change in distance and the acceleration in velocity gained over distance, respectively. Since stability results for many common schemes for approximating, the wave equation and the heat equation are well known, an often used practical strategy is to take the more restrictive of the two stability constraints for the wave and heat equations as the stability condition for the advection diffusion equation. The definition of stability that we employ here is a generalization of the classical Neumann stability conditions and is designed to guarantee that the computed solution inherits one important property of the analytical solution. In this paper, we are interested in the numerical study on the stability criteria of approximation schemes for solving this equation.

Many researchers have already been worked on it. Derivation of ADE, analytical solution, and numerical simulation of ADE are studied in the literatures [1], [2], and papers [10-12], [17]. Atul Kumar, Dilip Kumar Jaiswaland Naveen Kumar [5] presented an analytical solution of one dimensional advection diffusion equation with variable coefficients in a finite domain by using Lap

lace transformation technique. In that process new independent space and time variables have been introduced. Ogata and Banks [6] obtained analytical solution of the one dimensional ADE by reducing the original ADE into a diffusion equation by applying moving coordinates. F.B. Agusto and O. M. Bamingbola [7] obtained a Numerical Treatment of the Mathematical Models for Water Pollution. Young-San Park, Jong-Jin Baik [8] presented an analytical solution of the advection diffusion equation for a ground level finite area source. Al-Niami and Ruston, 1977 [9] obtained analytical solution of the one dimensional ADE by reducing the original ADE into a diffusion equation by introducing another dependent variable. L.F. Leon, P.M. Austria [13] presented Stability Criterion for Explicit Scheme on the solution of Advection-Diffusion Equation. T. F. Chan [14] presented Stability analysis of finite difference schemes for the advection diffusion equation. Alain Rigal [15] obtained Stability analysis of finite difference schemes for the Navier-Stokes equations. K.W. Morton [16] obtained Stability and convergence in fluid flow problems.

With the above discussion in view, in the present paper, Numerical Study on the Stability of Finite Difference Schemes with FTBSCS and FTCS techniques for solving the advection diffusion equation is presented. Numerical experiments are performed to verify the stability criterions obtained in this study. The schemes are compared with an analytical solution of ADE graphically to show accuracy of the solutions.

2 Numerical Schemes for Governing equation

We consider the one–dimensional ADE as an initial and boundary value problem $\frac{\partial c}{\partial t}+u\frac{\partial c}{\partial x}=D\frac{\partial^2 c}{\partial x^2}$ with initial condition $c(t_0,x)=c_0(x)$; $a\le x\le b$ and boundary conditions $c(t,a)=c_a(x)$; $t_0\le t\le T$ $c(t,b)=c_b(x)$; $t_0\le t\le T$. Finite difference techniques for solving the one dimensional advection diffusion equation can be considered according to the number of spatial grid points involved, the number of time levels used, whether they are explicit or implicit nature.

2.1 Explicit finite difference scheme for ADE

For the numerical solution of the one –dimensional linear advection- diffusion equation we consider the IBVP $\frac{\partial C}{\partial t}+u\frac{\partial C}{\partial x}=D\frac{\partial^2 C}{\partial x^2}$…….. (1) With initial condition $c(x,0)=f(x)$ $0\le x<l$ and Neumann boundary conditions $\frac{\partial c}{\partial x}(x=0,t)=0$; $0<t\le T$ and $\frac{\partial c}{\partial x}(x=l,t)=0$; $0<t\le T$. In order to develop the scheme, we discretize the x-t plane by choosing a mesh width h≡Δx space size and a time step size k≡Δt. The finite difference methods, we will develop, produce approximations $c_i^n\in R^n$ to the solution $c(x_i,t_n)$ in the discrete points by x_i= ih, i= 0, 1, 2, 3, … … … and t_n= nk, n= 0, 1, 2, 3, … … …Let the solution $c(x_i,t_n)$ be denoted by c_i^n and its approximate value by c_i^n.

2.2 Explicit Centered Difference Scheme by FTBSCS Techniques

By Explicit forward time difference formula $\frac{\partial c}{\partial t}=\frac{c_i^{n+1}-c_i^n}{\Delta t}$ … … … (2), next use the backward space difference formula $\frac{\partial c}{\partial x}=\frac{c_i^n-c_{i-1}^n}{\Delta x}$ … … … (3) and centered space difference formula $\frac{\partial^2 c}{\partial x^2}=\frac{c_{i+1}^n-2c_i^n+c_{i-1}^n}{\Delta x^2}$ … … … (4). Substituting equations (2- 4) into equation (1) and rearrange according the time level, lead to $\frac{c_i^{n+1}-c_i^n}{\Delta t}+u\frac{c_i^n-c_{i-1}^n}{\Delta x}=D\frac{c_{i+1}^n-2c_i^n+c_{i-1}^n}{\Delta x^2}$

Which leads to $c_i^{n+1}=c_i^n-\frac{u\Delta t}{\Delta x}(c_i^n-c_{i-1}^n)+\frac{D\Delta t}{\Delta x^2}(c_{i+1}^n-2c_i^n+c_{i-1}^n)$

$$c_i^{n+1}=\left(\frac{u\Delta t}{\Delta x}+\frac{D\Delta t}{\Delta x^2}\right)c_{i-1}^n+\left(1-\frac{u\Delta t}{\Delta x}-2\frac{D\Delta t}{\Delta x^2}\right)c_i^n+\frac{D\Delta t}{\Delta x^2}c_{i+1}^n$$

Implies to $c_i^{n+1}=(\gamma+\lambda)c_{i-1}^n+(1-\gamma-2\lambda)c_i^n+\lambda c_{i+1}^n$ … … … (5) in which, $\gamma=\frac{u\Delta t}{\Delta x}$, $\lambda=\frac{D\Delta t}{\Delta x^2}$

2.3 Explicit Centered Difference Scheme by FTCS techniques

By Explicit forward time difference formula $\frac{\partial C}{\partial t} = \frac{c_i^{n+1} - c_i^n}{\Delta t}$ … … … (6), next use the centered space difference formula $\frac{\partial C}{\partial x} = \frac{c_{i+1}^n - c_{i-1}^n}{2\Delta x}$ … … … (7) and $\frac{\partial^2 C}{\partial x^2} = \frac{c_{i+1}^n - 2c_i^n + c_{i-1}^n}{\Delta x^2}$ … … … (8)

Substituting equations (6 - 8) into equation (1) and rearrange according the time level, lead to

$$\frac{c_i^{n+1} - c_i^n}{\Delta t} + u\frac{c_{i+1}^n - c_{i-1}^n}{2\Delta x} = D\frac{c_{i+1}^n - 2c_i^n + c_{i-1}^n}{\Delta x^2}$$

Which leads to $$c_i^{n+1} = c_i^n - \frac{u\Delta t}{2\Delta x}(c_{i+1}^n - c_{i-1}^n) + \frac{D\Delta t}{\Delta x^2}(c_{i+1}^n - 2c_i^n + c_{i-1}^n)$$

$$c_i^{n+1} = \left(\frac{u\Delta t}{2\Delta x} + \frac{D\Delta t}{\Delta x^2}\right)c_{i-1}^n + \left(1 - 2\frac{D\Delta t}{\Delta x^2}\right)c_i^n + \left(\frac{D\Delta t}{\Delta x^2} - \frac{u\Delta t}{2\Delta x}\right)c_{i+1}^n$$

Implies to

$c_i^{n+1} = (\gamma/2 + \lambda)c_{i-1}^n + (1 - 2\lambda)c_i^n + \left(-\frac{\gamma}{2} + \lambda\right)c_{i+1}^n$ … (9), in which, $\gamma = \frac{u\Delta t}{\Delta x}$, $\lambda = \frac{D\Delta t}{\Delta x^2}$

Now, we can write the general form of an explicit 1st order scheme as $c_i^{n+1} = L_0 c_{i-1}^n + L_1 c_i^n + L_2 c_{i+1}^n$

where, the values of the coefficients are given at below:

Coefficient of explicit central difference scheme			
Scheme	L_0	L_1	L_2
FTBSCS	$\gamma + \lambda$	$1 - \gamma - 2\lambda$	λ
FTCS	$\frac{\gamma}{2} + \lambda$	$1 - 2\lambda$	$\lambda - \frac{\gamma}{2}$

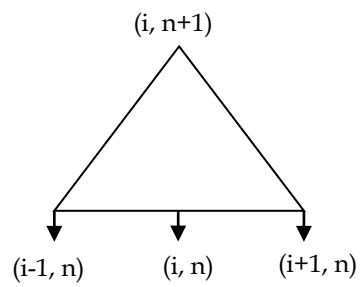


Figure 2.3: The stencil for explicit method for the ADE.

So, from above equations, we observe that knowing the values at time n we can obtain the corresponding ones at time n+1 using this recurrence relation c_0^n and c_i^n must be replaced by the boundary conditions, in this example they are both 0 and 1.

3. Stability Analysis

After surveying the relevant literature on the subject, we discover that no practical stability criterion exists for (5) and (9). We developed simultaneous stability conditions for both the schemes and maintaining the criteria we verify the results of the schemes by setting an example.

3.1 Stability conditions for the scheme with FTBSCS techniques are given by

$$0 \le \frac{D\Delta t}{\Delta x^2} \le 1 \text{ and } -\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 1 - 2\frac{D\Delta t}{\Delta x^2}.$$

In which $\gamma = \frac{u\Delta t}{\Delta x}$, $\lambda = \frac{D\Delta t}{\Delta x^2}$

3.2 Stability conditions for the scheme with FTCS techniques are given by

$$0 \le \frac{D\Delta t}{\Delta x^2} \le \frac{1}{2} \text{ and } -2\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 2\left(1 - \frac{D\Delta t}{\Delta x^2}\right),$$

In which, $\gamma = \frac{u\Delta t}{\Delta x}$, $\lambda = \frac{D\Delta t}{\Delta x^2}$

4. Numerical Simulation and results discussions

Various finite difference equations were used to represent the parabolic model equation (1). It is extremely important to experiment with the application of these numerical techniques. It is hoped that by writing computer codes and analyzing the results, additional insights into the solution procedures are gained. Therefore, this section proposes an example and presents solutions by the described schemes.

4.1 Problem description: Estimation of pollutant in a river of length l = 6 meter at all time t = 1minute to t = 6 minutes with fluid velocity, u = 0.01 m/s = 36 m/h and diffusion coefficient, D = 0.01m2/s=36 m2/h. The advection diffusion equation for this problem is $\frac{\partial C}{\partial t} + u\frac{\partial C}{\partial x} = D\frac{\partial^2 C}{\partial x^2}$. Various values of spatial nodes size and time steps are to be used to investigate the numerical schemes and the effect of steps on stability and accuracy. An attempt is made to solve the stated problem subject to the imposed initial and Neumann boundary conditions by the following: The FTBSCS and FTCS schemes with

$\Delta x = 0.05$, $nx = 120$ $\Delta t = 0.07$, $nt = 3600$

$\Delta x = 0.15$, $nx = 40$ $\Delta t = 0.1$, $nt = 3600$

$\Delta x = 0.05$, $nx = 120$ $\Delta t = 0.122$, $nt = 2948$

Solutions

Case I. When the step sizes are $\Delta x = 0.05$, $\Delta t = 0.07$.

In this case, both the schemes are to be used as stated previously: The stability requirements of the FTBSCS scheme are $0 \le \frac{D\Delta t}{\Delta x^2} \le 1$ and $-\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 1 - 2\frac{D\Delta t}{\Delta x^2}$ (the terms $\frac{u\Delta t}{\Delta x} = \gamma$ and $\frac{D\Delta t}{\Delta x^2} = \lambda$ are known as the advection number and diffusion number respectively).

For this particular application, $\lambda = \frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.07}{(0.05)^2} = 0.28$

$$\gamma = \frac{u\Delta t}{\Delta x} = \frac{0.01 \times 0.07}{0.05} = 0.014$$

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.07}{(0.05)^2} = 0.28 \le 1 \text{ and } -\frac{0.01\times 0.07}{(0.05)^2} \le \frac{0.01\times 0.07}{0.05} \le 1 - 2\times\frac{0.01\times 0.07}{(0.05)^2}$$

or,

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.07}{(0.05)^2} = 0.28 \le 1 \text{ and } -0.28 \le 0.014 \le 0.44.$$

And the stability requirements of the FTCS scheme are $0 \le \frac{D\Delta t}{\Delta x^2} \le \frac{1}{2}$ *and* $-2\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 2\left(1 - \frac{D\Delta t}{\Delta x^2}\right)$.

For this particular application, $\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.07}{(0.05)^2} = 0.28 \le \frac{1}{2}$ and $-2\times\frac{0.01\times 0.07}{(0.05)^2} \le \frac{0.01\times 0.07}{0.05} \le 2\left(1 - \frac{0.01\times 0.07}{(0.05)^2}\right)$

or,

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.07}{(0.05)^2} = 0.28 \le \frac{1}{2} \text{ and } -0.56 \le 0.014 \le 1.44.$$

Therefore, the stability conditions for both the schemes are satisfied, and a stable solution is expected. The velocity profiles are to be obtained up to t = 4 minutes are shown in Figure 4.1.

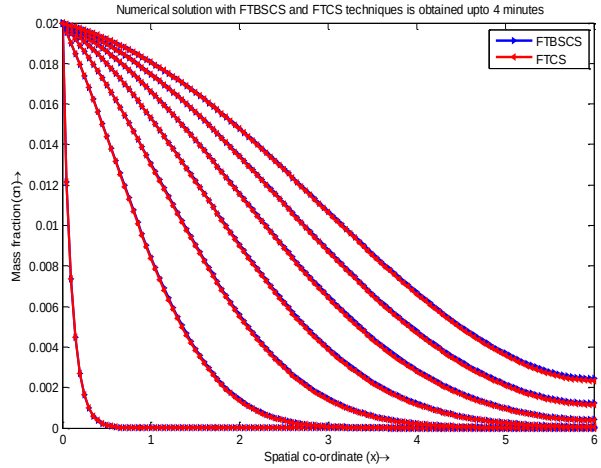


Figure 4.1: Velocity profiles with the schemes, Δx=0.05, Δt=0.07

Case II. When the step sizes are increased to Δx = 0.15, Δt = 0.1, the stability requirements of the FTBSCS scheme are $0 \le \frac{D\Delta t}{\Delta x^2} \le 1$ and $-\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 1 - 2\frac{D\Delta t}{\Delta x^2}$. For this particular application, $\lambda = \frac{D\Delta t}{\Delta x^2} = \frac{0.01\times0.1}{(0.15)^2} = 0.044$ and $\gamma = \frac{u\Delta t}{\Delta x} = \frac{0.01\times0.1}{0.15} = 0.007$

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times0.1}{(0.15)^2} = 0.044 \le 1 \text{ and } -\frac{0.01\times0.1}{(0.15)^2} \le \frac{0.01\times0.1}{0.15} \le 1 - 2\times\frac{0.01\times0.1}{(0.15)^2}$$

or

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times0.1}{(0.15)^2} = 0.044 \le 1 \text{ and } -0.044 \le 0.007 \le 0.912$$

And the stability requirements of the FTCS scheme are $0 \le \frac{D\Delta t}{\Delta x^2} \le \frac{1}{2}$ and $-2\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 2\left(1 - \frac{D\Delta t}{\Delta x^2}\right)$.

For this particular application,

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times0.1}{(0.15)^2} = 0.044 \le \frac{1}{2} \text{ and } -2\times\frac{0.01\times0.1}{(0.15)^2} \le \frac{0.01\times0.1}{0.15} \le 2\left(1 - \frac{0.01\times0.1}{(0.15)^2}\right)$$

or

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times0.1}{(0.15)^2} = 0.044 \le \frac{1}{2} \text{ and } -0.088 \le 0.007 \le 1.912$$

Therefore, the stability condition is satisfied, and a stable solution is expected. The velocity profiles are to be obtained up to t = 6 minutes are shown in Figure 4.2.

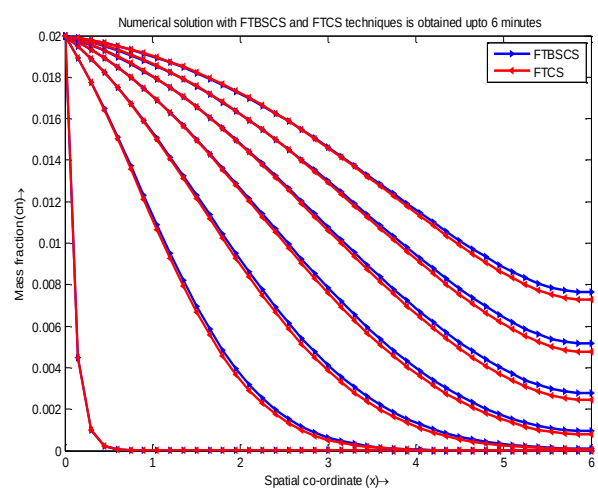


Figure 4.2: Velocity profiles with the schemes, Δx=0.15, Δt=0.1

Case III. When the step sizes are increased to $\Delta x = 0.05$, $\Delta t = 0.122$, which is only a fraction of an increase over preceding cases.

In this case, the stability requirement of the FTBSCS scheme are $0 \le \frac{D\Delta t}{\Delta x^2} \le 1$ and $-\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 1 - 2\frac{D\Delta t}{\Delta x^2}$.

For this particular application, $\lambda = \frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.122}{(0.05)^2} = 0.488$ and $\gamma = \frac{u\Delta t}{\Delta x} = \frac{0.01\times 0.122}{0.05} = 0.0244$

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.122}{(0.05)^2} = 0.488 \le 1 \text{ and } -\frac{0.01\times 0.122}{(0.05)^2} \le \frac{0.01\times 0.122}{0.05} \le 1 - 2\times\frac{0.01\times 0.122}{(0.05)^2}$$

or

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.122}{(0.05)^2} = 0.488 \le 1 \text{ and } -0.488 \le 0.0244 \le 0.024, \text{ which exceeds the stability requirement.}$$

And the stability requirements of the FTCS scheme are $0 \le \frac{D\Delta t}{\Delta x^2} \le \frac{1}{2}$ and $-2\frac{D\Delta t}{\Delta x^2} \le \frac{u\Delta t}{\Delta x} \le 2\left(1 - \frac{D\Delta t}{\Delta x^2}\right)$.

For this particular application,

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.122}{(0.05)^2} = 0.488 \le \frac{1}{2} \text{ and } -2\times\frac{0.01\times 0.122}{(0.05)^2} \le \frac{0.01\times 0.122}{0.05} \le 2\left(1 - \frac{0.01\times 0.122}{(0.05)^2}\right)$$

or

$$\frac{D\Delta t}{\Delta x^2} = \frac{0.01\times 0.122}{(0.05)^2} = 0.488 \le \frac{1}{2} \text{ and } -0.976 \le 0.0244 \le 1.024.$$

Therefore, at this stage one of the stability conditions for FTBSCS is not satisfied, and an unstable solution is appeared. With the step sizes indicated, an unstable solution is developed. The velocity profiles are to be obtained at t = 6 minutes are shown in Figure 4.3.

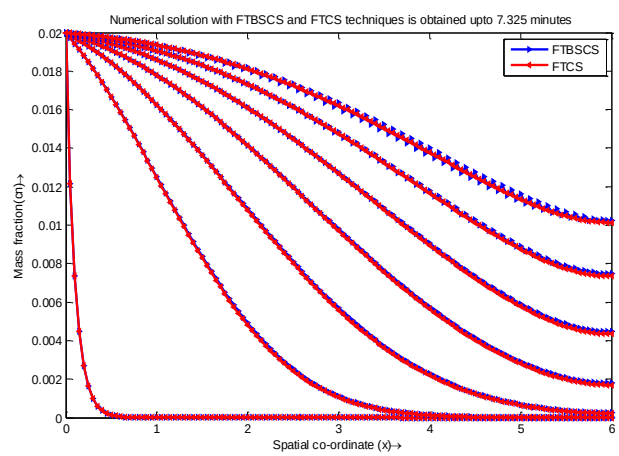


Figure 4.3: Velocity profiles with the schemes, Δx=0.05, Δt = 0.122

5. Analysis

In the preceding section, two explicit finite difference schemes are applied to the advection diffusion equation and the solutions are presented. The effect of the stability imposed by the diffusion number on the FTBSCS and FTCS explicit schemes are clearly indicated. Therefore, for these schemes the selection of step sizes is limited due to the stability requirement. However, the accuracy requirement limits the use of large time steps, since an increase in time steps will increase the truncation errors introduced in the approximation process of the PDE.

For the simple problem under consideration, an analytical solution may be obtained. The analytical solution of ADE with the imposed initial and boundary conditions is as follows-

5.1 The Advection-diffusion equation as an IBVP

The one-dimensional advection-diffusion equation [1] is given as $\frac{\partial c}{\partial t} + u\frac{\partial c}{\partial x} = D\frac{\partial^2 c}{\partial x^2} \ldots\ldots\ldots (10)$ where, c represents the solute con-

centration [ML-3] at x along longitudinal direction at time t, D is the solute dispersion, if it is independent of position and time, is called dispersion coefficient [L2T-1], t=time[T]; x= distance[L] and, u is the mean flow velocity [LT-1] assumed to be constant.

Appended with initial condition $c(x,0) = f(x) \qquad 0 \le x < l \ldots\ldots\ldots (11)$

and boundary conditions $c(x=0,t) = g_0(x) \qquad 0 < t \le T \ldots\ldots\ldots (12)$

$c(x=l,t) = g_1(x) \qquad 0 < t \le T \ldots\ldots\ldots (13)$

the ADE formulates an initial boundary value problem (IBVP).

5.2 Analytic solution

By coordinate transformation, the exact solution [1] of the advection-diffusion equation in unbounded is given by

$$c(x,t) = \frac{M}{A\sqrt{4\pi D\mathrm{t}}} exp\left(-\frac{(\mathrm{x} - (\mathrm{x}_0 + \mathrm{ut}))^2}{4D\mathrm{t}}\right) \ldots\ldots\ldots (14)$$

5.3 Error Estimation and Convergence

An error term is in L1-norm as $err = \frac{\|c_e - c_n\|_1}{\|c_n\|} \ldots\ldots\ldots (15)$ where, ce is the exact solution, and cn is the numerical solution computed by the finite difference schemes for time $t \in [0,6]$.The following figure 5.1 shows the convergence of relative error by the scheme FTBSCS.

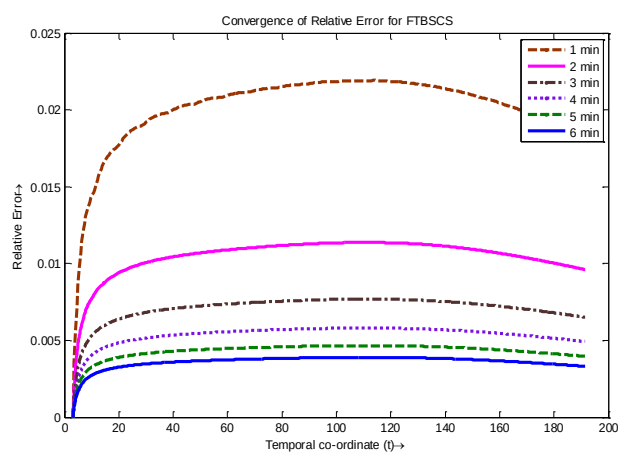


Figure 5.1 Rate of Numerical feature of Convergence

The following figure 5.2 shows the convergence of relative error by the scheme FTCS.

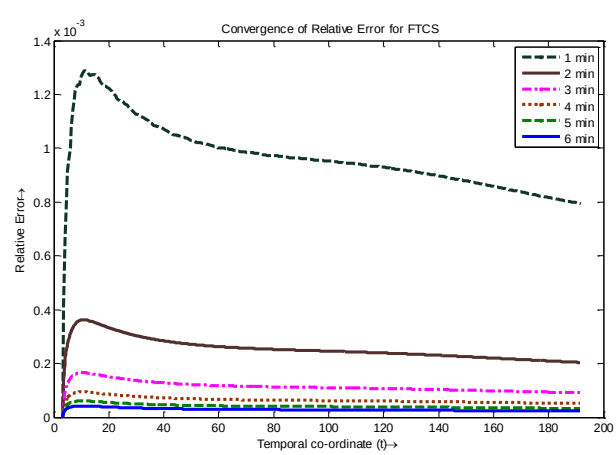


Figure 5.2 Rate of Numerical feature of Convergence

The following figure 5.3 shows the comparison of relative errors for the both schemes.

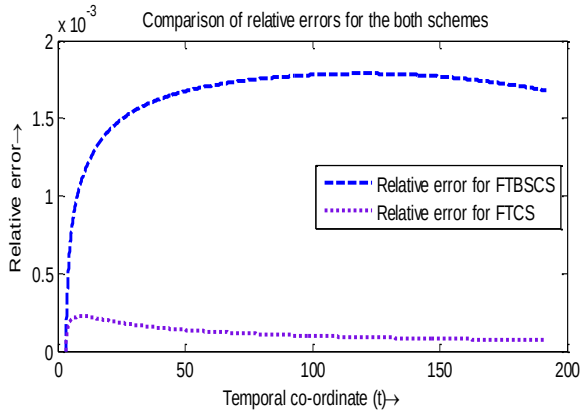


Figure 5.3 Comparison of relative errors for the both schemes

6. Conclusion

We have presented stability analysis, analytical solutions and numerical solutions for 1D advection diffusion equation with an initial condition and Neumann boundary conditions. Numerical experiment is presented graphically. The analytical result is used for code validation and for error comparison of both schemes. In addition, it is used to study the effect of step size on the accuracy of solutions. The results shown in Figure 5.1 – 5.3 are the error terms as defined above at time level [1, 6]. Two points to emphasize with regard to Figure 5.1-5.3 are: (1) for this application, the FTCS scheme has minimum error in comparison with FTBSCS scheme, and the amount of error is decreased for the both schemes as the solution is marched in time. This error reduction is due to a decrease in the influence of the initial data.

REFERENCES

[1] Scott A. Socolofsky & Gerhard H. Jirka, "Environmental Fluid Mechanics", part 1, 2nd Edition, 2002.

[2] Yijun Liu, "Introduction to the Finite Element Method", lecture Notes; University of Cincinnati, U.S.A.

[3] Anderson, Mary P., and William W. Woessner. "Applied Groundwater Modeling, Simulation of Flow and Advective Transport". Academic Press Inc., San Diego, 1992. Pages 6-68.

[4] Bender, Edward A. "An Introduction to Mathematical Modeling". John Wiley & Sons Inc., New York, 1978. Pages 1-140.

[5] Atul Kumar, Dilip Kumar Jaiswal and Naveen Kumar,"Analytical solution of one dimensional Advection diffusion equation with variable coefficients in a finite domain", J.Earth Syst. Sci. 118, No.5, pp. 539-549, October 2009.

[6] Ogata A, Banks R.B.,"A Solution of the differential equation of longitudinal dispersion in porous media", US Geological Survey, Paper 1961; 411-A; 1961.

[7] F.B. Agusto and O. M. Bamingbola, "Numerical Treatment of the Mathematical Models for Water Pollution", Research Journal of Applied Sciences 2(5): 548-556, 2007.

[8] Young-San Park, Jong-Jin Baik,"Analytical solution of the advection diffusion equation for a ground level finite area source", Atmospheric Environment 42, 9063-9069, 2008.

[9] Al-Niami A N S and Ruston K R 1977, "Analysis of flow against dispersion in porous media", J.Hydrol. 3387-97.

[10] Aral M M and Liao B 1996, "Analytical solutions for two-dimensional transport equation with time-dependent dispersion coefficients", J. Hydrol. Engg. 1(1) 20-32.

[11] Nicholas J.Higham, "Accuracy and stability of Numerical Algorithms", Society of Industrial and Applied Mathematics, Philadelphia, 1996.

[12] M. Thongmoon and R.Mckibbin, "A comparison of some numerical methods for the advection-diffusion equation", Int. Math. Sci. 2006, Vol. 10, pp. 49-52.

[13] L.F. Leon, P.M. Austria, "Stability Criterion for Explicit Scheme on the solution of Advection-Diffusion Equation", Maxican Institute of Water Technology.

[14] T. F. Chan, "Stability analysis of finite difference schemes for the advection diffusion equation", SIAM J. Numer. Anal., 21:272-284, 1984.

[15] Alain Rigal, "Stability analysis of finite difference schemes for the Navier-Stokes equations", Internat. J. Numer. Meth. engng., 14 (1979), pp. 617-628.

[16] K.W. Morton. "Stability and convergence in fluid flow problems", Proc, Roy. Soc. London A, 323: 237-253, 1971.

[17] Changjun Zhu and Shuwen Li, "A numerical Simulation of River Water Pollution using Grey Differential Model". Journal of Computer, No.9, September 2010.

T. M. Abul Kalam Azad was born on February 15, 1966 in Sirajgonj, Bangladesh. He received his MPhil degree on Ring Theory, MSc (Pure Mathematics) and BSc(Hons) degrees from the Department of Mathematics, Rajshahi University. He has been working in the Department of Computer Science and Engineering, University of Liberal Arts Bangladesh (ULAB), Dhaka from 2010 as a full time faculty. Previously, he worked at home, 'Rajshahi University of Engineering & Technology (RUET)' and abroad. Now, he is doing his PhD degree courses on Computational fluid dynamics in the Department of Mathematics, Jahangirnagar University. He has published a number of papers in peer-reviewed international and national journals. He is a life member of Bangladesh Mathematical Society.

Laek Sazzad Andallah was born on July 12, 1967 in Nilphamari, Bangladesh. He received his PhD degree on Applied Mathematics from TU Ilmenau Institute, Germany and MSc degree on industrial Mathematics from University of Kaiserslautern, Germany. He received his MSc, BSc(hons) degrees from the Department of Mathematics, Jahangirnagar University. He joined at the Department of Mathematics, Jahangirnagar University on 1992 as teaching associate. Now, he is working in the same department as a professor. He has published a number of papers in peer-reviewed international and national journals and presented original research articles in numerous international conferences. He received various scholarships, research and travel grants as recognition of his research works. He is a member of Bangladesh Mathematical Society.

Environmental Problems of Bangladesh: A Review

Nishita Ivy
Institute of Forestry and Environmental Sciences
University of Chittagong, Chittagong, Bangladesh
E-mail: nishitaivy@gmail.com

Abstract

Bangladesh is one of the developing countries of Asia which has been facing serious environmental deterioration due to disposal of solid and liquid wastes here and there in the nature without any treatment, gaseous emissions to the atmosphere and exposure to noise. Air, water and noise pollution and disposal of solid wastes are the major environmental problems of Bangladesh. Environment of Bangladesh is mainly polluted due to deficiency of standard waste management system and land use, rapid urbanization, over population, exploitation of natural resources, industrialization and capitalization. It is needed to take necessary preventive measures and actions to polluters to improve the environmental quality for ensuring healthy living of citizens of the country.

Keywords—Environmental Problems and Pollution, Air Pollution, Water Pollution, Disposal of Solid Waste, Noise Pollution.

© University of Liberal Arts Bangladesh
All rights reserved.

Manuscript received on 28 July 2016 and accepted for publication on 11 September 2016.

1 Introduction

IN developing countries of Asia, rapid growth of economy, population, urbanization and dependency on vehicles are causing increased risk on environment [1]. Deterioration of environment has most hostile effects on Asian, African and Latin American countries due to poverty than the other countries of the world [2]. As an overpopulated developing country of Asia, the population of Bangladesh will be about 17 crore by the year of 2020 [3]. Environmental condition of the country is deteriorated due to different sources of environmental pollution. Rapid urbanization and industrialization lead the country for major environmental problems. Air, water and soil pollution, disposal of solid waste and exposure to noise are the key problems of environment in Bangladesh. Moreover, Bangladesh is on high risk of climate change and global warming.

The current environmental situation of the country could not be balanced thoroughly because acute environmental pollution are affecting public health, ecosystem and growth of economy [4]. Environmental problems occur in all major cities of the country [5]. Environment is being deteriorated and natural resources are being depleted in this country because of rapid growth of population, poverty, and absence of proper knowledge on environmental conservation [6]. Presence of proper effluent treatment plants has not been found in most of the industrial establishments of Bangladesh so that disposal of industrial effluent pollute the soil and water and cause emissions to the air [7]. Moreover, emissions from motor vehicles, aero planes, trains, industrial establishments, power generation set up, brick fields, uncovered burning, incinerating, solid waste dumping area and particulate matters are causing to atmospheric pollution [8]. Year to year, this country has been enduring deterioration process of environment that is a serious matter of thinking [2]. So environmental problems due to different sources of pollution should be controlled to reduce the effects of environmental pollution and to ensure the healthy environmental condition of the country. Avoidance of the certain limit of deterioration of environment in developmental process is also a subject of argument [9].

The natural environment of Bangladesh has been under continuous pressure due to unplanned urbanization and industrialization [10]. Although environmental rules and regulations are formulated and enforcing in some extent, it is needed to take more necessary steps for improving environmental quality and to prevent deterioration of environment in Bangladesh. People of this country are not also aware that they are polluting the environment because of lack of knowledge and awareness. The effort of this article was to focus on deterioration of environment and to recognize the key concerns of environment of Bangladesh to provide an overall scenario of major environmental problems of the country for improving its environmental situation. However, more emphasis should also be given on environmental improvement for continuing to achieve millennium development goals as well as achieving sustainable development goals. For that reason, effective environmental management system and policy framing can only be the way of preventing and mitigating deterioration of environmental quality.

2 Materials and Methods

This review study has been carried out thoroughly based on secondary information. A systematic review of available research studies was carried out for identifying major environmental problems and for giving an outline to address these problems. In Bangladesh, many methodical qualitative and quantitative research works have been carrying out on different perspectives of environmental problems. Many research works were conducted on water quality, air pollution, solid waste management, noise pollution and industrial pollution in the country. In this article, mainly available qualitative and quantitative research works and articles were collected from Internet which are related to air, water, soil and noise pollution of Bangladesh. Collected documents were analyzed to retrieve information and to give an overview of major problems in case of environmental degradation in Bangladesh.

3 Results and Discussions

As demonstrated in this document, the numbering for sections upper case Arabic numerals, then upper case Arabic numerals, separated by periods. Initial paragraphs after the section title are not indented. Only the initial, introductory paragraph has a drop cap.

3.1 Air Pollution

Air pollution is an issue of significant danger in numerous developing countries which is causing annual loss of life of about 20, 00,000 people globally [11]. Polluted air causes respiratory problems, bronchitis, headaches and dizziness, nasal congestion, and renal damage and a lot of emissions also contribute to the greenhouse effect and thus inducing global warming and sea level rise [10]. Air pollution is a great environmental concern for Bangladesh particularly in big metropolitan cities like Dhaka and Chittagong [12]. Ambient levels of Particulate Matter, SO_2 and Pb far exceed the levels of Bangladesh air quality standards and WHO guidelines [8]. Recorded data on air pollution in Dhaka city confirms that the annual mean of PM10 and PM2.5 is highly significant [13]. According to Mahmood [11], the volume of lead concentration in atmosphere of Dhaka city is 463 nanograms per cubic meter which is ten times higher than the permissible limit. Burning of fossil fuels, emissions from brick kiln, industrial emissions and emissions from vehicles are the main causes of atmospheric pollution in Bangladesh [3], [14]. According to Dewan *et al.* [15], "Leather, food, pulp and paper, textile industries around the Dhaka city largely contribute to SO_2, NO_2, CO, PM and Volatile Organic Chemicals (VOC) to the air, for example, the estimated emission of SO_2 by the brick kilns and manufacturing industries is 28.8% and 10%, respectively."

As discussed by DoE [12], there are some remarkable projects for air quality management. These are:

i. Air Quality Management Project (AQMP) implemented by the DoE with support from the World Bank during 2000-2007.
ii. Clean Air and Sustainable Environment (CASE) project supported by the World Bank.
iii. Bangladesh Air Pollution Management (BAPMAN) project.
iv. A project for implementation of Male Declaration.

In Dhaka city, air quality improvement was found in some extent because of CNG introduction and banning 2-stokes 3 wheelers [14]. Although some initiatives have been taken to prevent air pollution, such initiatives are not sufficient for improving the air quality and more initiatives are needed for controlling and monitoring air emissions from vehicles, industries, brick kilns and other possible sources of air pollution.

3.2 Water Pollution

Bangladesh, located mostly in the flood plains of the Ganges, the Brahmaputra rivers and the Meghna rivers, is one of the largest deltas in the world and crossed by 405 rivers draining an area of 1,750,000 Km^2 [16]. Water has significant influence in the economy of this country [10]. The reality of water resources in Bangladesh is however a reason for grave concern [17].

Water is crucial for maintaining the quality of life and inevitable source of sustaining existence [2], [18]. But, disposal of urban and medical wastages and hazardous releases of pollutants from industrial establishments degrade the quality of both surface and ground water [4]. In Bangladesh, most of the industries release hazardous wastewater to the surrounding water courses apart from sense of environment [19]. In most of the cases, wastewater are discharged from point sources without any treatment. Moreover, the residential and commercial establishments near the river banks directly discharge wastewater into the rivers or their discharges subsequently find way into the rivers [20]. The main water pollutants in Bangladesh are a) liquid organic and inorganic wastes, b) nutrient substances, c) synthetic compounds, d) inorganic chemicals, e) silt and sediment, f) hot water and i) industrial, municipal and urban wastes [21].

According to the Institute of Environment and Development Studies [22], "about 900 polluting industries in Bangladesh dispose their untreated industrial wastes directly into the rivers, although the effluent contain 10 to 100 times higher than the allowable permissible limits." It was found that the rivers around the Dhaka Metropolitan City were highly polluted in the months from January to May [16]. Seven thousands industries situated near the rivers discharge daily 15,00,000 m^3 of wastewater into the Buriganga, the Shitalakhya, the Balu and the Turag rivers of Dhaka city and other 5,00,000 m^3 of wastewater come into these rivers from different non-industrial origins [10]. Halder and Islam [23] found, "the maximum concentration of turbidity, BOD, hardness,

TDS and COD of Turag river of Dhaka is much higher than the acceptable limits." Study revealed that the concentration of DO and BOD has exceeded beyond the standard limits in cases of the Buriganga and the Shitalakhya rivers during 2000 to 2010 [10]. The Karnaphuli is the main river of commercial capital Chittagong district of Bangladesh where the river is polluted in several ways particularly through industrial and sewerage disposals and municipal wastes from different drainage systems without any treatment [24] [25]. Quality of riverine water is polluting day by day [17]. From 1990 and onwards, the Government of Bangladesh has been pushing the industrial proponents towards pollution abatement measures and installations of Effluent Treatment Plants (ETPs) at their premises [10]. Although industries are installing ETPs, structure of ETPs should be improved more for tertiary treatment along with primary and secondary treatments. Moreover, it is needed to take necessary steps to treat municipal sewerage and domestic wastewater before final disposal to the water bodies.

Arsenic is another problem regarding ground water pollution in Bangladesh. This utmost serious issue is caused by the deterioration of environment [4]. Numerous tubewells has found polluted with arsenic in ninety's decade [6]. According to World Bank [10], "about 35 million of Bangladesh population are affected by the arsenic contamination." According to UNICEF [26], "nationwide, approximately 20 per cent of shallow tubewells are contaminated and in more than 8,000 villages where 80 per cent of all tube wells are contaminated; about 20 million people in Bangladesh are using tube wells with more than 50ppb of Arsenic." Hence, it is needed to focus more to mitigate water pollution from different sources of pollutants along with Arsenic contamination of ground water.

3.3 Disposal of Solid Waste

Unsanitary situation of solid waste is terrible in Bangladesh and this situation will be worst with rapid growth of population in municipal area [27]. Consumption as well as production patterns are changing rapidly causing an increasing quantity of wastes [10]. About 4000 to 4500 tons of solid waste are generated daily and this waste generation is accomplished with dispersing on roadways, spillage, obstructing water courses, undiscriminating disposal in empty areas and create degradation of environment [4]. Per capitia waste generation rate is 0.41 kg/day in the municipal area of Bangladesh although in total waste collection situation is not very satisfactory [3]. In Dhaka city, each day over 3000 tons of household waste is being produced, but Dhaka City Corporation collects less than half of it [28]. Enayetullah *et al.* [29] mentioned, "existing infrastructure for waste management shows that waste collection efficiency in different urban areas varies from 37% to 77% with an average of 55%."
Waste management is a serious environmental concern for Bangladesh, although waste could be transformed in resources [10]. From decomposition of solid waste, methane emission occurs in the atmosphere which is one of the main gases of greenhouse effect. Bad odor and smell also cause serious threat to the urban environment due to mismanagement of solid waste. Spread of communicable diseases is one of the major effect of mismanagement of municipal waste and poorly managed wastes cause air-, water- and land-pollution, damage of ecosystems, decrease of soil fertility and loss of aesthetic beauty [10]. It is extremely needed to take more initiatives like 3Rs (Reduce, Reuse and Recycle) and 4Rs (Recover, Reduce, Reuse and Recycle) for management and policy formulation of solid wastes, which ultimately improve the environment.

3.4 Noise Pollution

Noise is undoubtedly a source of atmospheric pollution [30]. Excessive noise pollution has become one of the major concerns of urban life in Bangladesh [31]. Due to industrial and urban growth, noise pollution is increasing in municipal areas of the country [32]. It was found from study that some selected residential areas of Dhaka have maximal sound pollution than the standard sound level prescribed by the Department of Environment (DoE) of Bangladesh [33]. As mentioned by Ahmed amd Rahman [30], "noise level of busy streets in Dhaka city has been estimated 60 to 80 dB, with the sound of vehicles being 95 dB, loud speakers 90 to 100 dB, mills and factories 80 to 90 dB , restaurants and cinema halls 75 to 90 dB, scooter or motorbike 87 to 92 dB, trucks and buses 92 to 94 dB. However, the desired sound level is 25 dB in the bedroom, 40 dB in the dining or drawing room, 35-40 dB in the office, 30-40 dB in the classroom, 35-40 dB in the library, 20-35 dB in hospital, 40-60 dB in a restaurant and 45 dB in the city at night."

Horns from vehicles, aeroplanes, railway stocks, vans, public transport facilities and other means of transports all producing excessive noise and because of exposure to excessive noise, numerous health problems are happened [34]. Because of noise pollution, impairment of hearing ability and malfunctioning of other body systems is more frequent [35]. Khan [30] mentioned, "exposure to noise pollution on a regular basis at a high decibel level causes deafness, memory loss, nausea, peptic ulcer, blood pressure, heart ailments and mental and physical disorders of children."

3.5 International Emphasis on Environment

First step for pollution prevention was taken by the Water Boundary Treaty in 1909 which was an intermediation between the United States and Canada. In 1972, first international conference focusing on environment named by the 'United Nations Conference on Human Environment' initiated steps for preventing human impact on environment [37]. After this conference, some international pacts were signed: (i) Geneva Convention on Transboundary Air Pollution in 1979, (ii) Vienna Convention for the Protection of the Ozone Layer in 1985, (iii) Montreal Protocol on Substances that Deplete the Ozone Layer in 1987, (iv) Basel Convention on the Control of Transboundary Movements of Hazardous Wastes and their Disposal in 1989, (v) UN Conference on Environment and Development (UNCED) in Rio de Janerio, Brazil in 1992 which is also known as 'Earth Summit', (vi) Agenda 21 - a

platform for sustainable development and (vii) United Nations Framework Convention on Climate Change (UNFCCC) in 1995 [37], [38], [39].

In 2001, Stockholm Convention on Persistent Organic Pollutants (POPs) banned specific chemicals and intercept the usage of other chemicals and also prescribed disposal and management of chemicals in environmental friendly procedures [39]. The United Nations Climate Change Conferences (UNFCCC) has been held annually from 1995 that lead to establish 'Kyoto Protocol' in 1997 for protecting the world from global warming through cut down of emissions of greenhouse gases and the last effort of UFCCC was 'Paris Agreement' in 2015 for bounding earth temperature beneath 2 or 1.5 degree Celsius [40].

3.6 Steps for Protecting Envionment and Mitigating Environmental Problem in Bangladesh

In Bangladesh, policies, rules and laws concerning environmental problems have been formulated. Some important steps were taken to protect environment like adoption of National Environment Policy in 1992, initiation of National Environment Management Action Plan (NEMAP), enactment of the Environmental Conservation Act 1995 and obtaining environmental clearance for operation of each and every industrial units and projects [10]. Major environmental policies, rules and laws of Bangladesh are given below:

a. The Motor Vehicles Rules, 1940 (Extracts)
b. The Building Construction Act, 1952
c. Water Pollution Control Ordinance, 1970 [Repealed by Ord. XIII of 1977]
d. Environmental Pollution Control Ordinance, 1977 [Repealed by Act I of 1995]
e. Motor Vehicles Ordinance, 1983 (Extracts)
f. Industrial Policy, 1991
g. Bangladesh National Environmental Policy, 1992
h. National Conservation Strategy, 1992
i. National Environmental Management Action Plan (NEMAP), 1995
j. Bangladesh Environmental Conservation Act, 1995 [Act I of 1995]
k. Bangladesh Environmental Conservation Rules, 1997
l. National Policy for Safe Water Supply and Sanitation, 1998
m. The National Water Policy, 1999
n. Bangladesh Environmental Court Act, 2000
o. Bangladesh Environmental Conservation (Amendment) Act, 2000
p. Bangladesh Environmental Conservation (Amendment) Act, 2002
q. Bangladesh Environmental Court (Amendment) Act, 2002
r. Ozone Depleting Substance (Control) Rules, 2004
s. National Energy Policy, 2004
t. Sound Pollution Rules, 2006
u. Building Construction Rules, 2006
v. Medical Waste (Management and Handling) Rules, 2008
w. Mobile Court Act, 2009
x. Bangladesh Environmental Conservation (Amendment) Act, 2010
y. Bangladesh Environmental Conservation (Amendment) Rules, 2010
z. Bangladesh Environmental Court Act, 2010
aa. National Urban Sector Policy, 2011
bb. Hazardous Waste and Ship Breaking Waste Management Rules, 2011
cc. Bangladesh Water Act, 2013 [41]

It is urgent need to enforce and implement effectively the above mentioned rules, laws and policy for protecting deterioration of environmental condition. The Ministry of Environment and Forest (MoEF) is in charge of devising appropriate policies, plans and programs and to coordinate the activities for protection and improvement of the environment and the Department of Environment (DoE) under the MoEF is the regulatory body and technical wing of the government responsible for enforcing environmental laws to ensure environmental conservation and sustainable development [10]. Moreover, strong monitoring is also necessary to observe violation of legislation by different industries, brick kilns, establishments and vehicles.

4 Conclusion

Sound environmental condition plays an important role for fulfilling basic requirements of living of people [2]. Bangladesh is the 8th biggest populous and 6th most densely populated country of the world [10]. Air, water and soil are the major abiotic parts of environment which are needed to consider for any development of the country. But by ignoring the environmental components, Bangladesh is facing serious environmental degradation because of pollution problems with rapid growth of population and industrialization. Deterioration of environment is in dreadful situation over the current periods of time [2]. Environment pollution has adverse effect on healthy living, prosperity and survival of people [9].

Although initiatives were taken by the government as well as by the policy making and enforcing Department of Environment

(DoE), but more initiatives should be taken to abate environmental pollution. Sector wise management system and implication of legislation to reduce air, water and soil pollution and to ensure proper management of solid waste shall be effective in such case. Policy makers are also needed to be more active and aware about environmental problems for effective decision making. We should adhere the message of the United Nations Research Institute for Social Development [9]: "Environmental problems must be understood as part of the larger social framework, as an integral part of social integration, and must be addressed from this perspective." Without proper implementation of legislation and enforcement of law, it is impossible to reduce environmental problems. As the people of this country are not fully aware about the importance of conserving environment, so environmental awareness programs are needed for the people. Engagement of civil society in creating environmental awareness and for taking part in environmental development is also essential in this respect. Moreover, as a low lying country, Bangladesh is one of the vulnerable countries for global warming, climate change and sea level rise and also natural and man-made disasters. Environmental deterioration has also been leading the possibility of more disastrous situations for the country. So it is the demand of time to take necessary steps (preventive, precautionary and corrective measures) to improve the environmental quality of Bangladesh.

REFERENCES

[1] HEI Scientific Oversight Committee, "Outdoor air pollution and health in developing countries of Asia: a comprehensive review," Special Report 18, Health Effects Institute, Boston, MA, Nov. 2010.

[2] M. Jahan, "The impact of environmental degradation on women in Bangladesh: an overview," Asian Affairs, vol.30, no. 2, pp.5-15, 2008.

[3] K.M. Bahauddin, M.H. Uddin, "Prospect of solid waste situation and an approach of Environmental Management Measures (EMM) model for sustainable solid waste management: case study of Dhaka city," Journal of Environmental Science and Natural Resources, vol. 5, no. 1, pp. 99-111, 2012.

[4] G.M.J. Alam,"Environmental pollution of Bangladesh – it's effect and control," Proc. of the International Conference on Mechanical Engineering 2009 (ICME2009), Dhaka Bangladesh, 26-28 Dec. 2009.

[5] M.M. Karim, "Status of air quality and State-of-art control measures in Dhaka, Bangladesh," 94th Annual Conference and Exhibition of Air and Waste Management Association, 24-28 June 2001.

[6] Discovery Bangladesh, "Bangladesh: environmental issues," http://www.discoverybangladesh.com/meetbangladesh/env_issue.html. 2016.

[7] A.F.M.A. Satter (Ed.), "Environmental pollution and health hazards: present and future perspective of Bangladesh," Journal of Current and Advance Medical Research, vol. 1, no. 1, pp. 1-2, 2014.

[8] M.H. Rahman and A. Al-Muyeed, "Urban air pollution: a Bangladesh perspective," WIT Transactions on Ecology and the Environment, vol. 82, pp. 605-614, 2005.

[9] United Nations Research Institute for Social Development, "Environmental degradation and social integration," UNRISD Briefing Paper No. 3, World Summit for Social Development, Nov. 1994.

[10] DoE, "Bangladesh environment and climate change outlook," Department of Environment, Dhaka, Bangladesh, June 2013.

[11] S.A.I. Mahmood, "Air pollution kills at least 15,000 Bangladeshis each year: the role of public administration and government integrity," Journal of Public Administration and Policy Research. vol. 3, no. 4, pp. 129-140, 2011.

[12] Department of Environment (DoE), "Air pollution reduction strategy of Bangladesh," Final Report, Department of Environment, Dhaka, Bangladesh, Dec. 2012.

[13] M.A. Motalib, and R.D. Lasco, "Assessing air quality in Dhaka city," International Journal of Science and Research, vol. 4, no. 12, pp. 1908-1912, 2015.

[14] K.M.T. Ahmed, and D.A. Begum, "Air pollution aspects of Dhaka city," Proc. of International Conference on Environmental Aspects of Bangladesh, Japan, Sep. 2010.

[15] A.M. Dewan, M.H. Kabir, K. Nahar and M.Z. Rahman, "Urbanization and degradation in Dhaka Metropolitan Area of Bangladesh," Int. J. Environment and Sustainable Development, vol. 1, no. 2, pp.118-147, 2012.

[16] Anon, "Report on the national stakeholder consultation on water supporting the post-2015 development agenda," The Post 2015 Water Thematic Consultation, Dhaka, 30 Mar. 2013.

[17] M.I. Sarwar, A.K. Majumder and M.N. Islam, "Water quality parameters: a case study of Karnaphuli river Chittagong, Bangladesh," Bangladesh Journal of Scientific and Industrial Research, vol. 45, no. 2, pp.177-181, 2010.

[18] M.N. Uddin, M.S. Alam, M.A. Mobin, and M.A. Miah, "An assessment of the river water quality parameters: a case study of Jamuna river" Journal of Environmental Science and Natural Resources, vol. 7, no. 1, pp. 249-256, 2014.

[19] N. Ivy, M.K. Hossain and M.L. Hossain, "Effects of industrial effluents on germination and early growth of selected agricultural crops," Journal of Agronomy, vol. 14, no. 1, pp. 43-48, 2015.

[20] M.A. Rahman, and D.A. Bakri, "A study of selected water quality parameters along the river Buriganga, Bangladesh," Iranica Journal of Energy and Environment, vol. 1, no. 2, pp. 89-92, 2010.

[21] C. Chakraborty, M.M. Huq, S. Ahmed, T. Tabassum and M.R. Miah, "Analysis of the causes and impact of water pollution of Buriganga river: a critical study," International Journal of Scientific and Technology Research, vol. 2, no. 9, pp. 245-252, 2013.

[22] Institute of Environment and Development Studies, "Aquatic ecology and dangerous substances: Bangladesh perspectives," Diffuse Pollution Conference, Dublin, 2003.

[23] J. N. Halder and M.N. Islam, "Water pollution and its impact on the human health," Journal of Environment and Human, vol. 2, no. 1, pp. 36-46, 2015.

[24] S. Dey, J. Das, and M.A. Manchur, "Studies on heavy metal pollution of Karnaphuli river, Chittagong, Bangladesh," IOSR Journal of Environmental Science, Toxicology and Food Technology, vol. 9, no. 8, pp.79-83, 2015.

[25] B. Das, Y.S.A. Khan, and M.A.K. Sarkar, "Trace metal concentration in water of the Karnaphuli river estuary of the Bay of Bengal" Pakistan Journal of Biological Sciences, vol. 5, pp. 5, pp. 607-608, 2002.

[26] UNICEF, "Arsenic Mitigation in Bangladesh" Available online, http://www.unicef.org/bangladesh/Arsenic.pdf, Accessed on 3 January 2016.

[27] N. Ivy, M.M. Uddin and M.K. Hossain, "People's perception on using waste bins in reduce, reuse and recycle (3Rs) process for solid waste management (SWM) in Chittagong, Bangladesh," International Journal of Applied Science, Technology and Engineering Research, vol. 2, no. 3, pp. 30-40, 2013.

[28] T.A. Chowdhury, and S.R. Afza, "Waste management in Dhaka city-a theoretical marketing model" BRAC University Journal, vol. 3, no. 2, pp. 101-111, 2006.

[29] I. Enayetullah, A.H.M.M. Sinha, and S.S.A. Khan, "Urban solid waste management scenario of Bangladesh: problems and prospects," Waste Concern Technical Documentation, Waste Concern, 6 June, 2005.

[30] T. Ahmed, and T. Rahman, "Non-Auditory health hazard vulnerability to noise pollution: assessing public awareness gap," American Journal of Engineering Research, vol. 4, no. 4, pp. 143-147, 2015.

[31] M. Sultan, "Noise pollution: a major concern of urban life," The Financial Express, Available online, http://print.thefinancialexpressbd.com/old/more.php?news_id=124351&date=2012-03-23, Accessed on 23 March 2012.

[32] M.A. Haq, M.M. Islam, M.S. Ali, M.F. Haque and M.M.R. Akhand, "Status of noise pollution in mixed areas of Dhaka city: a GIS approach," Journal of Environmental Science and Natural Resources, vol. 5, no. 1, pp. 9-17, 2012.

[33] A.M. Husain, S. Yusuf, T.H. Rini and M. Hasan, "Noise pollution in major places in Dhaka and proposing a device to keep a data log," Journal of Modern Science and Technology, vol. 3, no. 1, pp. 20-30, 2015.

[34] S. Q. Chowdhury, "Noise pollution," Available online, http://en.banglapedia.org/index.php?title=Noise_Pollution, Accessed on 12 March 2015.

[35] S.C. Chowdhury, M.M. Razzaque, and M.M. Helali, "Assessment of noise pollution in Dhaka city," 17th International Congress on Sound and Vibration (ICSV17), Cairo, Egypt, 18-22 Jul. 2010.

[36] M.A. Khan, "Noise pollution: the silent killer," Available online, http://www.reportsbd.com/noise-pollution-the-silent-killer/, Accessed on 4 January 2016.

[37] Anon, "Handbook on international environmental agreements: an Indian perspective," Center for Environmental Law, WWF – India, 172 – B, Lodi Estate, New Delhi – 110003, pp. 2-4, 2006.

[38] J. G. Speth, "Environmental pollution: a long term perspective," World Resource Institute, 1709 New York Avenue, N.W., Washington, D.C. 20006 USA, pp. 281, 1988.

[39] D. B. Magraw and J. Gorin, "Laws and Regulations: International," Available online, http://www.pollutionissues.com/Ho-Li/Laws-and-Regulations-International.html, Accessed on 5 January 2016.

[40] "United Nations Framework Convention on Climate Change," Available online, https://en.wikipedia.org/wiki/United_Nations_Framework_Convention_on_Climate_Change, Accessed on 5 January 2016.

[41] Department of Environment (DoE), Available online: http://www.doe.gov.bd/, Accessed on 5 January 2016.

Nishita Ivy was born in 1990 in Chittagong, Bangladesh. She received B.Sc. (Honors) and M.S. degrees in Environmental Science from the Institute of Forestry and Environmental Sciences, University of Chittagong in 2013 and 2014, respectively. She received University Grants Commission (UGC) Government Merit Scholarship and also got nomination for the Prime Minister Gold Medal for her excellent academic result in B.Sc. (Honors) Examination. At present, she is working as an Environmental Management Trainee Officer in a reputed organization. In addition to this paper she also published two other scientific papers. Her research is focused on environmental management, industrial effluents and solid waste management.

On the Independence of Sample Mean and Variance

[1]Anwar H. Joarder and [2]M. Hafidz Omar
[1]School of Business, University of Liberal Arts Bangladesh, Dhanmondi, Dhaka 1209, Bangladesh,
Email: anwar.joarder@ulab.edu.bd
[2]Department of Mathematics and Statistics, King Fahd University of Petroleum and Minerals Dhahran 1261,
Saudi Arabia, Email: omarmh@kfupm.edu.sa

Abstract

We present two new proofs of the independence of sample mean and variance for two independently, identically and normally distributed random variables.

Key Words: Sample mean, sample variance, independence, moment generating function

© University of Liberal Arts Bangladesh
All rights reserved.

Manuscript received on 31 July 2016 and accepted for publication on 10 September 2016.

1 Introduction

The independence of sample mean and variance of independently, identically and normally distributed variables is essential in the basic definition of Student t-statistic [1] and also in the development of many statistical methods. For a sample of size n, it is usually proved using the independence of $\bar{X}$ and $(X_1 - \bar{X},\ X_2 - \bar{X}, \ldots, X_n - \bar{X})$, (see e.g. Theorem 1, p.340, [2]), but this requires background on independence of functions of random variables (Theorem 2, p.121, [2]) that may not be easily accessible to beginning undergraduate students.

There are several proofs of the independence of sample mean and variance, the popular of which seems to be the one due to[3]. It uses moment generating function of chi-square variable conditional on sample mean and as such not straightforward. Therefore, proofs accessible to undergraduate students have been an issue of discussion. See for example, [4] and also American Statistician, 1992, Volume 46, No. 1, pp. 72-75.

In this note, we present two new proofs. The first one is a direct proof of independence of sample mean and variance but for a sample of size 2 with a view to shedding light on the topic and to inspire students and instructors. Though we use moment generating function but unlike Shuster's proof we avoid the use of conditional distributions. The second one seems to be simplest of all other proofs though it requires the evaluation of a double integral and a sense of origin invariance of sample variance.

2 The Importance of the Result

The derivation of *t*-distribution by [1] required the independence of sample mean and variance. Though it was not clearly mentioned in the paper, [5] figured out that the independence of sample mean and variance was implicitly assumed. The independence of mean and variance was proved by [3] using a mathematical tool provided by Fisher. This happened one year before Student passed away. [7] points that Helmert proved that sample mean and variance are independent. Because of this historical fact, [8] recommended calling the joint distribution of the two random variables "Helmert's Distribution". Lukacs (1942) presented an easier proof of the result:

If the variance (or second moment) of a population distribution exists, then a necessary and sufficient condition for the normality of the population distribution is that sample mean and variance are mutually independent. Since then many authors including [8], [9], [10], [11], [3], [4] and [12] came up with different proofs and characterizations.

The independence of mean and variance extends to a broader class of distributions when the iid restriction of the sample is relaxed. For example, it is well known that if $(X_1, X_2, ..., X_n)$ has an *n*-variate normal distribution with an exchangeable covariance pattern, that is, $V(X_j) = \sigma^2$,

$$Cov(X_j, X_k) = \rho\sigma^2, \ -\frac{1}{n-1} < \rho < 1, \ < 1, \text{ for } j = 1, 2, ..., n \text{ and } k = 1, 2, ..., n$$

(c.f. [10], pp. 196–197). [13], in an expository article, also demonstrated independence of mean and variance based on non-iid samples from populations with specific mixture structures. It has been proved, see for example, [14] that *t*-statistic based on elliptically symmetric distributions has Student *t*-distribution. The *t*-statistic based on many skew normal distributions has also Student *t*-distribution. See, for example, [15].

Also *t*-statistics based on a joint distribution proposed by [16] follow Student *t*-distribution. [17] came up with vertical density representation that includes a brod class of distributions that guarantees that *t*-statisic has a Student *t*-distribution.

The above proves that normailty is not a necessity as has been thought of over the decades.
However, even with all these developments, we feel there is no direct and simple proof of the independence of mean and variance.

3 Main Results

Let X_1 and X_2 have an arbitrary 2-dimensional joint distribution. We define the sample mean $\bar{x}$ and variance s^2 by $n\bar{x} = x_1 + x_2$ and $2s_x^2 = (x_1 - x_2)^2$ respectively. Since $\bar{X} \sim N(\mu, \sigma^2/n)$, the moment generating function $M_{\bar{X}}(t_1) = E[\exp(t_1\bar{X})]$ of $\bar{X}$ is known to be

$M_{\bar{X}}(t_1) = \exp\left(t\mu + \frac{t_1^2}{2} \times \frac{\sigma^2}{n}\right)$, which simplifies to $M_{\bar{X}}(t_1) = \exp\left(t_1\mu + \frac{t_1^2\sigma^2}{4}\right)$ for a sample of size 2.

The moment generating function $M_{S^2}(t_2) = E[\exp(t_2 S^2)]$ of S^2 with degrees of freedom $m = 1$ is given by

$$M_{S^2}(t_2) = E\left[\exp\left(\frac{mS^2}{\sigma^2}\frac{t_2\sigma^2}{m}\right)\right] = M_U\left(\frac{t_2\sigma^2}{m}\right),$$

where, U has a χ^2 with $m = 1$ degrees of freedom. Then for $n = 2$, we have,

$$M_{S^2}(t_2) = \frac{1}{\sqrt{1 - 2t_2\sigma^2}}, \text{ where } -\infty < t_2 < \frac{m}{2\sigma^2}.$$

For the history of the distribution of sample variance or of chi-square based on normal distribution, see [6].

Theorem 3.1 Let the random variables X_1 and X_2 be independently, identically and normally distributed with $E(X_1) = \mu$ and $Var(X_1) = \sigma^2$. Then the joint moment generating function of the sample mean $\bar{X}$ and variance S_X^2 satisfies $M_{\bar{X}, S_X^2}(t_1, t_2) = M_{\bar{X}}(t_1) M_{S_X^2}(t_2)$ and hence $\bar{X}$ and S_X^2 are independent.

First Proof. For two observations $2s_x^2 = (x_1 - x_2)^2$. The joint moment generation function of $\bar{X} = (X_1 + X_2)/2$ and s_x^2 is given by $M_{\bar{X}, s_X^2}(t_1, t_2) = E(e^{t_1\bar{X} + t_2 s_X^2})$. Letting $X = \mu + \sigma Z$, or, $X_j = \mu + \sigma Z_j$, $j = 1, 2$, we have $M_{\bar{X}, s_X^2}(t_1, t_2) = E\exp[t_1(\mu + \sigma\bar{Z}) + t_2(\sigma^2 s_Z^2]$, where, $X_j = \mu + \sigma Z_j$, $j = 1, 2$. Obviously $s_Z^2 = \frac{1}{2}[(\mu + \sigma Z_1) - (\mu + \sigma Z_2)]^2 = \frac{1}{2}(Z_1 - Z_2)^2$, and $f_{Z_j}(z_j)$ is the density function of standard normal variable Z_j, $j = 1, 2$. Then we have,

$$M_{\bar{X}, S^2}(t_1, t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \exp\left[t_1(\mu + \sigma\bar{z})] + t_2\sigma^2 s_Z^2\right] f_{Z_1}(z_1) f_{Z_2}(z_2) dz_1 dz_2,$$

which can be written as,

$$M_{\bar{X},S^2}(t_1,t_2) = e^{t_1\mu}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp\left[\frac{1}{2}t_1\sigma(z_1+z_2)+\frac{1}{2}t_2\sigma^2(z_1-z_2)^2\right]f_{Z_1}(z_1)f_{Z_2}(z_2)dz_1dz_2. \tag{3.1}$$

Since the exponent of the exponential function in the above integrand is

$$q = \frac{1}{2}t_1\sigma(z_1+z_2)+\frac{1}{2}t_2\sigma^2(z_1-z_2)^2-\frac{1}{2}(z_1^2+z_2^2),$$

(3.1) can be written as

$$M_{\bar{X},S^2}(t_1,t_2) = e^{t_1\mu}\exp\left(\frac{t_1^2\sigma^2}{4(1-t_2\sigma^2)}\right)\int_{-\infty}^{\infty} f_W(w)M_Y(-t_2\sigma^2 w)dw, \tag{3.2}$$

where, $W \sim N\left(\frac{t_1\sigma}{2(1-t_2\sigma^2)},\ \frac{1}{1-t_2\sigma^2}\right)$ and $Y \sim N\left(\frac{t_1\sigma}{2(1-t_2\sigma^2)},\ \frac{1}{1-t_2\sigma^2}\right)$.

Having written out $M_Y(-t_2\sigma^2 w)$ and used in (3.2), we have

$$M_{\bar{X},S^2}(t_1,t_2) = \exp\left(t_1\mu+\frac{t_1^2\sigma^2}{4}\right)\frac{1}{\sqrt{1-2t_2\sigma^2}}I(t_2,\sigma),$$

where, $I(t_2,\sigma) = \int_{-\infty}^{\infty}\frac{\sqrt{1-2t_2\sigma^2}}{\sqrt{2\pi(1-t_2\sigma^2)}}\exp\left[-\frac{1-2t_2\sigma^2}{2(1-t_2\sigma^2)}\left(y-\frac{t_1\sigma}{2}\right)^2\right]dy.$

The above integral is 1 as the integrand is the density function of $Y \sim N\left(\frac{t_1\sigma}{2},\ \frac{1-t_2\sigma^2}{1-2t_2\sigma^2}\right)$.

Thus we have,

$$M_{\bar{X},S^2}(t_1,t_2) = \exp\left(t_1\mu+\frac{t_1^2\sigma^2}{4}\right)\frac{1}{\sqrt{1-2t_2\sigma^2}},$$

which is the product of $M_{\bar{X}}(t_1) = \exp\left(t_1\mu+\frac{t_1^2\sigma^2}{4}\right)$ and $M_{S^2}(t_2) = \frac{1}{\sqrt{1-2t_2\sigma^2}}$.

That is, $M_{\bar{X},S^2}(t_1,t_2) = M_{\bar{X}}(t_1)M_{S^2}(t_2)$. By uniqueness property of moment generating function, it implies the independence of sample mean and variance.

The proof for n observations seems to be straightforward but it is quite involved.

Second Proof. Without any loss of generality, we may assume that $\mu = 0$ and $\sigma = 1$. Then the joint density function of the sample is given by $f(x_1,x_2) = \frac{1}{2\pi}\exp\left(-\frac{1}{2}(x_1^2+x_2^2)\right)$, so that the joint moment generating function of $\bar{X}$ and s_X^2 is given by

$$M_{\bar{X},s_X^2}(t_1,t_2) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(t_1\bar{x}+t_2s_x^2)\times\frac{1}{2\pi}\exp\left(-\frac{1}{2}(x_1^2+x_2^2)\right)dx_1dx_2,$$

which can be written as,

$$M_{\bar{X},s_X^2}(t_1,t_2) = \exp\left(\frac{t_1^2}{4}\right)\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\exp(t_2s_x^2)\exp\left[-\frac{1}{2}\left(x_1-\frac{t_1}{2}\right)^2-\frac{1}{2}\left(x_2-\frac{t_1}{2}\right)^2\right]dx_1dx_2.$$

Then, with the transformation $u_i = x_i - t_1/2,\ (i=1,2),$ we have $s_x^2 = s_u^2,\ dx_1dx_2 = du_1du_2$, and hence

$$M_{\bar{X},s_X^2}(t_1,t_2) = \exp(t_1^2/4)I(t_2;u), \tag{3.3}$$

$$\text{where } I(t_2;u) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}\exp\left(t_2 s_u^2\right)\exp\left(-\frac{1}{2}(u_1^2+u_2^2)\right)du_1du_2.$$

The above integral is exactly the same as,

$$I(t_2;x) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{2\pi}\exp\left[\frac{1}{2}t_2(x_1-x_2)^2\right]\exp\left(-\frac{1}{2}(x_1^2+x_2^2)\right)dx_1dx_2 = M_{S_X^2}(t_2).$$

Since $M_{\bar{X}}(t_1) = \exp(t_1^2/4),$ it follows from (3.3), that $M_{\bar{X},s_X^2}(t_1,t_2) = M_{\bar{X}}(t_1)M_{s_X^2}(t_2).$ By uniqueness property of the moment generating function, this proves that the sample mean and variance are independent. The proof for n observations seems to be straightforward.

4. Conclusion

The independence of sample mean and variance has been proved in Section 3 for a sample of size 2. Since the proofs are simple and direct, it will make students and instructors confident about the fundamental theorems of statistics. An open problem is to generalize the proofs of this paper to any sample size. Another challenging problem would be to identify the family of distributions that enjoy the property of independence of sample mean and variance. This will prove the robustness of *t*-test under broader distributional assumptions.

Acknowledgement

The authors gratefully acknowledge the research facilities provided by King Fahd University of Petroleum and Minerals, Saudi Arabia through the Project IN131015.

References

[1] Student. The probable error of a mean. *Biometrika*, 6: 1-25, 1908.
[2] Rohatgi, V.K. and Saleh, A.K.M.E. *Probability and Statistics*. Wiley, 2001.
[3] Shuster, J. A simple method of teaching the independence of $\bar{X}$ and S^2. *The American Statistician*, 27(1), 29-30, 1973.
[4] Zehna, P.W. On Proving that $\bar{X}$ and S^2 are independent. *The American Statistician*, 45(2), 121-122, 1991
[5] Fisher, R.A. Student. *Annals of Eugenics*, 9, 1-9, 19390
[6] Geary, R.C. The distribution of Student's ratio for non-normal samples. *Supplement to the Royal Statistical Society*, 3, 178-184, 1936.
[7] Johnson, N.L.; Kotz, S. and Balakrishnan, N. Continuous Multivariate Distributions, v-2. New York: John Wiley and Sons, 1994.
[8] Kruskal, W.H. Helmert's distribution. *American Mathematical Monthly*, 53, 435-438, 1946.
[9]Zinger, A.A. On independent samples from a normal population. Uspekhi Matem. Nauk, 6:5(45), 172-175, 1951.
[10] Kagan, A.M.; Linnik, Yu. V. and Rao, C.R. (1973). *Characterization Problems in Mathematical Statistics*. John Wiley and Sons: New York, 1973.
[11] Rao, C. R. *Linear Statistical Inference and Its Applications* (2nd ed.). New York: Wiley, 1973.
[12]Stigler, Stephen M. Kruskal's proof of the joint distribution of $\bar{X}$ and S^2. *The American Statistician*, 46, 72-73, 1992.
[13] Mukhopadhyay, N., "Dependence or Independence of the Sample Mean and Variance in Non-iid or Nonnormal Cases and the Role of Some Tests of Independence," in *Recent Advances in Applied Probability, IWAP 2002 116 Teacher's Corner* Downloaded by [32.209.230.42] at 18:03 09 September 2014 *Caracas Proceedings*, eds. R. Baeza-yates, J. Glaz, J. Husler, and J. L. Palacios, New York: Springer-Verlag, pp. 397–426, 2005.
[14] Fang, K.T; Kotz, S. and Ng, K.W. *Symmetric Multivariate and Related Distributions.* Chapman and Hall, 1990.
[15] Gupta, A.K.; Omar, M. Hafidz and Joarder, A.H. On a Generalized Mixture of Standard Normal and Skew Normal Probability Distributions. *Pan American Mathematics Journal,* 23(3), 1-14, 2013.
[16] Gupta, Rameshwar D. and Richards, Donald St. P. Multivraite Liouville Distributions –V. 377-396. Advances in the Theory and Practice of Statistics: A Volume in Honor of Samuel Kotz. Eds: Norman L. Johnson and N. Balakrishnan (1997). John Wiley and Sons, New York, 1997.
[17] Yang, Zhenhai; Fang, Kai -Tai and Kotz, Samuel. On the Student's *t*-distribution and the *t*-statistic. *Journal of Multivariate Analysis*, 98, 1293-1304, 2007.

Anwar H. Joarder obtained PhD and MS degrees in Statistics from the University of Western Ontario in 1992 and 1988, respectively. Currently he is a professor at University of Liberal Arts Bangladesh (ULAB). Before joining ULAB, he worked at many institutions including the University of Western Ontario, Monash University and University of Sydney. He serves on the editorial board of about 10 international research journals. He has authored or co-authored about 80 methodical papers published in internationally reputed research journals. His research area includes robustness of statistical methods, correlated statistical analysis, multivariate parametric estimation, survey sampling and pedagogical statistics. He has been a lifelong Fellow of Royal Statistical Society, England, since 2005.

M. Hafidz Omar obtained PhD degree in Educational Measurement and Statistics in 1995 from the University of Iowa, USA. He obtained MS degree in Statistics in 1992 and BS in Actuarial Science in 1990 from the same university. He worked in companies and taught statistics at several institutions in the USA including the University of Iowa, the University of Kansas, and Riverside Publishing Company. Currently he is an Associate Professor of Statistics and Actuarial Science at King Fahd University of Petroleum and Minerals. He has refereed many research papers from reputable international journals. He has authored or co-authored more than 50 papers published in internationally reputed research journals. His research area includes applied multivariate statistics, correlated statistical analysis, probability modeling, mathematical statistics, educational Statistics, actuarial statistics, and statistics education.

The Appropriate Grid Size of *N* Layer Loosely Coupled Hexagonal Model for Equilibria of Boltzmann Equation

Wahida Zaman Loskor
Faculty of Science & Humanities
Bangladesh Army International
University of Science & Technology (BAIUST)
Comilla Cantonment, Comilla-3501, Bangladesh.
Email: loskor.baiust@gmail.com, loskor_gb@yahoo.com

Abstract

The discrete equilibrium solutions (equilibria) f of the Boltzmann equation can be expressed in terms of four parameters characterizing mass, (x, y)-momenta and kinetic energy. We present an error estimation by comparing the discrete equilibria with the corresponding Maxwellian which leads to determine the appropriate grid size N of the N-layer loosely coupled hexagonal grid for given temperature and bulk-velocity. We calculate numerical results showing how temperature depends on the parameter μ, characterizing kinetic energy and bulk-velocity depends on the parameter κ's, characterizing (x, y) momenta and μ as well.

Keywords— The Boltzmann equation, loosely coupled hexagonal grid, discrete equilibrium solution.

© University of Liberal Arts Bangladesh
All rights reserved.

Manuscript received on 20 July 2016 and accepted for publication on 30 August 2016.

1 Introduction

To derive an efficient scheme for the numerical simulation of the Boltzmann collision operator, one immediately faces two major difficulties: (i) The computational cost for the numerical simulation of the Boltzmann equation is very high. This is due to the complexity of the five dimensional integral Boltzmann collision operator, which has to be numerically evaluate at each point (x, y) in the (discretize) six-dimensional space and (ii) The discretization of the Boltzmann collision operator has to be modeled carefully so that it satisfies the features of kinetic theory like conservation laws, H-theorem, properties of linearized collisions operator etc. Several numerical techniques have been proposed in recent years to deal with the complexity of the Boltzmann collision operator. Simplified Collision model have been introduced in rectangular grid in [1][2][3], but there appears many artificial invariants which have to be eliminated by further techniques.

Therefore, in [4] has developed a kinetic theory for the discrete Boltzmann equation based on hexagonal grid in $\mathbf{R}^2$. The Boltzmann collision sphere can be much more suited in the hexagonal grid than rectangular grid model. The system of binary collisions contains artificial invariants and to avoid the artificial invariants a ternary interaction law is introduced. It is shown that the conservation laws, the H-theorem, the correct number of invariants and the properties of linearized operator are satisfied for the discrete Boltzmann equation in the hexagonal grid. In [5], has developed an automatic generation of the Boltzmann collision operator based on a hexagonal grid and made some numerical simulations based on the grid in. In [6], introduced a discrete model Boltzmann equation based on a loosely coupled hexagonal discretization of $\mathbf{R}^2$. The model satisfies the basic features of kinetic theory like conservation laws, H theorem, correct dimension of the null-space of the linearized collision operator etc. In [7] showed that the model Boltzmann equation, based on only binary collision law, discretized on the loosely coupled hexagonal grid in $\mathbf{R}^2$ provides two artificial invariants. In [8], developed a generalized layer-wise construction of a loosely coupled N-layer hexagonal mesh for a discrete model Boltzmann equation. This work also described some properties of the N-layer loosely coupled hexagonal grid and identify the regular hexagons belonging to the mash in

order to generate collision model for the Boltzmann equation.

The discrete equilibrium solutions (equilibria) of the Boltzmann equation based on a generalize N-layer loosely coupled hexagonal grid is determine in [9][10]. The equilibria f of the discrete Boltzmann equation can be expressed in terms of four parameters characterizing mass, (x, y)-momenta and kinetic energy. A necessary algorithm for the computation of the equlibria is also constructed by them.

In this article, several numerical simulations for a 6-layer loosely coupled hexagonal grid in $\mathbf{R}^2$ are performed. Also we present an error estimation by comparing the discrete equilibria with the corresponding Maxwellian which leads to determine the appropriate grid size N of the N-layer loosely coupled hexagonal grid for given temperature and bulk-velocity. We also observe temperature depends on the kinetic energy and bulk-velocity depends on the (x, y) momenta and kinetic energy.

2 Boltazmann Equation

The Boltzmann equation is a prominent representative of kinetic equations, describes the evolution of rarefied gases. With the conservation of momentum and energy, the dynamics of the Boltzmann equation is given by a free flow step and a particle interactions step. The free flow step is modeled by the Liouville-equation and the particle interaction step is modeled by the Boltzmann collision operator. As a simple mathematical consequence of the Boltzmann equation is $\left(\partial_t + v.\nabla_x\right) f\left(t, x, v\right) = \mathrm{J}[\mathrm{f}, \mathrm{f}]$, where, $\mathrm{J}[\mathrm{f}, \mathrm{f}] = \int_{R^d} \int_{S^{d-1}} k(\mathrm{v}-\mathrm{w}, \eta)[\mathrm{f}(\mathrm{v}')\,\mathrm{f}(\mathrm{w}') - \mathrm{f}(\mathrm{v})\,\mathrm{f}(\mathrm{w})] d^{d-1}\eta\, d^d \mathrm{w}$ is the Boltzmann collision operator and $f = f(t,x,v)$, a density function which depends on time, space and velocity. Here $k(.,.)$ is the collision kernel in the operator satisfying some symmetry properties, the post collision velocities $(\mathrm{v}', \mathrm{w}')$ result from the pre-collision velocities (v, w) satisfying the collision relations, conservation of momentum, $\mathrm{v} + \mathrm{w} = \mathrm{v}' + \mathrm{w}'$ and conservation of kinetic energy, $(|\mathrm{v}|^2 + |\mathrm{w}|^2) = (|\mathrm{v}'|^2 + |\mathrm{w}'|^2)$.

3 N-Layer Loosely Coupled Hexagonal Mesh

Any two neighboring hexagons have only one common vertex on a loosely coupled hexagonal discretization in $\mathbf{R}^2$. A generalized layer-wise construction of a loosely coupled N-layer hexagonal mesh for a discrete velocity model Boltzmann equation is presented in this section. A 72-velocity model which is constructed by adding two-layer of hexagons centering to a center one and called two layers loosely coupled hexagonal mesh is shown in Fig.3.1. Similarly, by adding one more layer of regular basic hexagons, one can obtain a 3-layer hexagonal mesh and so on. In general, we may call this a N-layer loosely coupled hexagonal mesh and the collision model based on the mesh can be called a N-layer hexagonal model which is a regular collision model so that it satisfies the basic kinetic features and can be divided into six symmetric partitions (Fig.3.1).

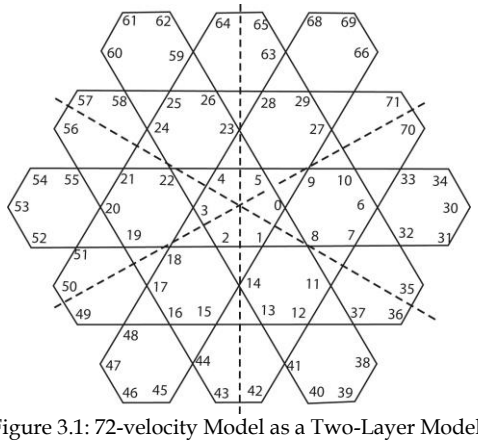


Figure 3.1: 72-velocity Model as a Two-Layer Model.

4 Equilibria For a N-Layer Model

Strictly positive density vectors $\mathrm{f} = \left(f_i\right)_{i=0}^{(3N+2)^2 + (3N+1)}$ is said to be the equilibrium solutions (equilibria) if $\mathrm{J}[\mathrm{f}, \mathrm{f}] \equiv 0$ for a N-layer hexagonal model. The ith equilibrium of the nth layer is $f(n,i) = z\mu^{m(n,i)}\, \kappa_{0+}^{\bar{\kappa}_0(n,i)}\, \kappa_{1+}^{\bar{\kappa}_1(n,i)}\, \kappa_{2+}^{\bar{\kappa}_2(n,i)}$; where $z, \kappa_{0+}, \kappa_{1+}, \kappa_{2+} > 0$ are satisfying arbitrary quantities $\kappa_{0+}\kappa_{1-}\kappa_{2+} = 1$. The equilibria at the six nodes of 0-st layer

(i.e. at the nodes of the central basic hexagon) are givenby $\left(f_0, f_1, f_2, f_3, f_4, f_5\right) = z.\left(\kappa_{0+}, \kappa_{1+}, \kappa_{2+}, \kappa_{0-}, \kappa_{1-}, \kappa_{2-}\right)^{\mathrm{T}}$.

For a 3-layer model, presents the equilibria for the nodes of the partition as

$z(\mu\kappa_{0+}\kappa_{1+},\ \mu^3\kappa_{0+}\kappa_{1+}^2,\ \mu^4\kappa_{1+}^3,\ \mu^3\kappa_{1+}^2\kappa_{2+}) \in$ 1st layer

$z(\mu^9\kappa_{0+}^2\kappa_{1+}^3,\ \mu^6\kappa_{0+}\kappa_{1+}^3,\ \mu^{10}\kappa_{0+}\kappa_{1+}^4,\ \mu^{12}\kappa_{1+}^5,\ \mu^{10}\kappa_{1+}^4\kappa_{2+},\ \mu^6\kappa_{1+}^3\kappa_{2+},\ \mu^9\kappa_{1+}^3\kappa_{2+}^2) \in$ 2nd layer

$z(\mu^{13}\kappa_{0+}^3\kappa_{1+}^3,\ \mu^{18}\kappa_{0+}^3\kappa_{1+}^4,\ \mu^{19}\kappa_{0+}^2\kappa_{1+}^5,\ \mu^{15}\kappa_{0+}\kappa_{1+}^5,\ \mu^{21}\kappa_{0+}\kappa_{1+}^6,\ \mu^{24}\kappa_{1+}^7,\ \mu^{21}\kappa_{1+}^6\kappa_{2+},\ \mu^{15}\kappa_{1+}^5\kappa_{2+},\ \mu^{19}\kappa_{1+}^5\kappa_{2+}^2,\ \mu^{18}\kappa_{1+}^4\kappa_{2+}^3) \in$ 3rd layer

Where z parameterizes mass, $(\kappa_{0+}, \kappa_{2+})$ characterize non-vanishing bulk-velocity, μ responsible kinetic energy.

Fig. 4.1 shows each nth layer of a partition has $(3n+1)$ nodes and the node numbering is from the top to bottom of at each layer.

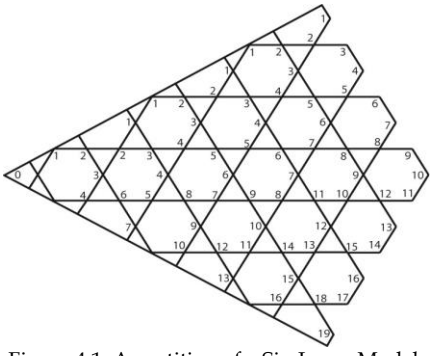


Figure 4.1: A partition of a Six-Layer Model.

We generalize nodes values of a partition for equilibra of *N*-layer model as a proposition in [9], [10]. Also establish a theorem in [9][10] for equilibria of *N*-layer loosely coupled hexagonal model and prove the theorem in [9].

5 Computation of Equilibrium

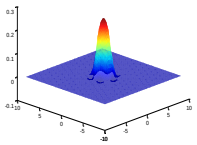
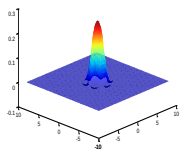
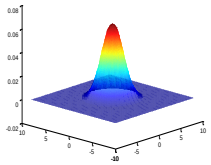
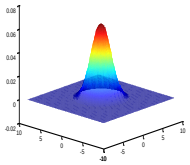
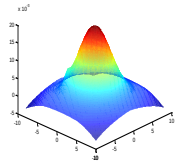
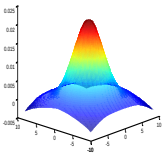
Here the discrete equilibria are computed, which are described by the parameters $z, \mu, \kappa_{0+}, \kappa_{2+}$ characterizing respectively mass, temperature, and bulk-velocity. It is evident that $\bar{v}_x = 0,\ <0,\ >0$ according as $\kappa_{0+}\kappa_{2+} = \kappa_{1+} = 1,\ <1,\ >1$; $\bar{v}_y = 0,\ <0,\ >0$ according as $\kappa_{0+} - \kappa_{2+} = 0,\ <0,\ >0$ and $\kappa_{0+} - \kappa_{2+} = \kappa_{0+} - \dfrac{\kappa_{1+}}{\kappa_{0+}}$.

Now the discrete equlibria $\tilde{\mathrm{f}}^h \in \varepsilon$ given by the theorem in [9], [10] for the case of zero bulk-velocity on a 4-layer grid (of 210 grid points) with discretization parameter $h = 1$ for three different values of $\mu = 0.25,\ 0.55,\ 0.95$ is computed and the corresponding Maxwellian given by $\tilde{f} = \dfrac{\rho}{(2\pi T)^{d/2}} \exp\left(\dfrac{-(\mathrm{v} - \bar{\mathrm{v}})^2}{2T}\right),\ d = 2$, where $\rho = \sum_i \tilde{f}_i^h$ is the mass density, $\bar{\mathrm{v}} = \left(1/\rho\right)\sum_i \mathrm{v}_i \tilde{f}_i^h$ is defined as the bulk velocity, $T = \left(1/2\rho\right)\sum_i (\mathrm{v}_i - \bar{\mathrm{v}})^2 \tilde{f}_i^h$ is the temperature. The normalized discrete equilibrium state f^h is compared with the normalized Maxwellian f and calculated the error $err = \left\|\mathrm{f} - \mathrm{f}^h\right\|_1$ and the moments, temperature as shown in the table 1.1.

Here three interesting cases of three different temperatures with zero bulk-velocities are presented.

1. In the first case, for $\mu = 0.25$ the calculated temperature is $T = 0.8293$. The main part of the configuration is centered at the origin with a small radius and a small part of the mass occurred on the grid. That is the resolution of the grid is too low to present such low temperature and this causes a noticeable 5% error.
2. This is a good situation because the main part of the mass of the function f lies inside of the domain. In this case it occurs very little error for $\mu = 0.55$ in which the calculated temperature $T = 1.6759$.
3. Here the grid is not large enough to present such high temperature $T = 11.0488$ for a given $\mu = 0.95$ and a significant fraction of the mass of the function f is cut down the boundary which causes 20% error. Thus to avoid this error it requires to further extension of the 4-layer grid model.

Table 1.1
In varying temperatures Discrete equilibria and Maxwellian are on a 4-layer grid

f^h = **Discrete equilibria**	f = **Maxwellian**	**Err =** $\|f - f^h\|_{L_1}$
		$\rho = 1.000$ $\bar{v}_x = 0$ $\bar{v}_y = 0$ $\mu = 0.25$ $T = 0.8293$ $err = 0.0501$
		$\rho = 1.000$ $\bar{v}_x = 0$ $\bar{v}_y = 0$ $\mu = 0.55$ $T = 1.6759$ $err = 0.0013$
		$\rho = 1.000$ $\bar{v}_x = 0$ $\bar{v}_y = 0$ $\mu = 0.95$ $T = 11.0488$ $err = 0.1920$

It is thus seen that a larger model is needed to restrict the error to a reasonable range for high temperature otherwise a noticeable error due to boundary effect.

6 Appropriate Grid Size

In order to determine the appropriate size grid for given values of μ, κ_{0+}, κ_{2+}, we compute the error $err = \|f - f^h\|_1$ for different sizes N of the N-layer model and chose the smallest N as an appropriate size for which the error restricted to a given tolerance. Fig.6.1 shows the error with respect to the size N for zero bulk velocity with four increasing values of temperature for $\mu = 0.3,\ 0.5,\ 0.7,\ 0.75$. In the first case for $\mu = 0.3$, it shows that the error goes below 1.2% for $N = 2$ and thus for given tolerance 0.012. $N = 2$ is the appropriate size for $\mu = 0.3$. Similarly, the rest three cases show that $N = 3, 4, 5$ are the appropriate size for given values of $\mu = 0.5,\ 0.7,\ 0.75$ respectively. For given temperature and bulk-velocity we can determine the appropriate size (az) of the model by the few steps as shown in algorithm 6.1.

Initialize $N = 1$
Calculate $err = \|f - f^h\|_{L1}$
If $err < tol$
$az = N$
BREAK
ELSE
$N = N + 1$
END
CONTINUE

Algorithm 6.1

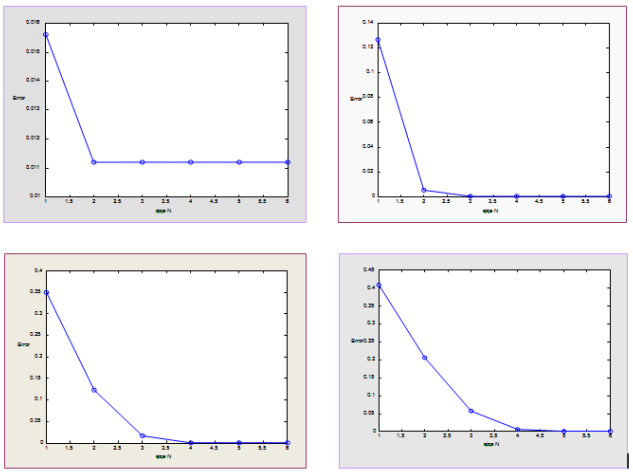


Figure 6.1: Error w.r.to size N for increasing temperature T in zero bulk-velocity v.

Now Figure 6.2, shows temperature depends on the parameter μ, characterizing kinetic energy.

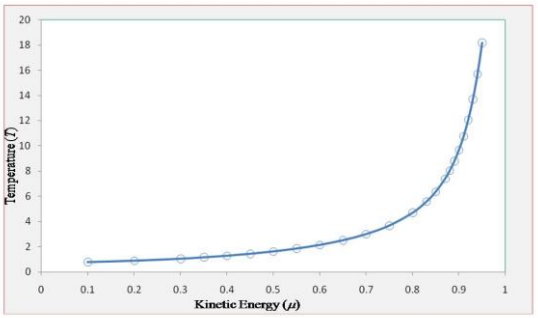


Figure 6.2: temperature depends on the parameter μ

It is clearly seen in the Figure 6.2 that temperature depends upon the values of the kinetic energy μ which is expected.

Figure 6.3, shows the calculated bulk-velocity for given $\mu \in [0.1,\ 0.9]$ at the three different choices of $(\kappa_{0+}, \kappa_{2+})$.

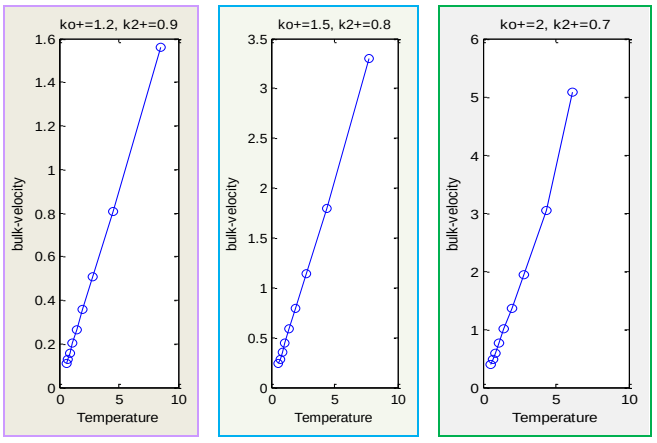


Figure 6.3: Bulk-velocity depends on $T(\mu)$ and κ's.

As expected, it is clearly seen in the Figure 6.3 that the modulus of the bulk-velocity $|\mathrm{v}|$ depends upon the choice of temperature as well as the values of the parameters κ_{0+}, κ_{2+}.

Figure 6.4 shows appropriate size $N(T,|\mathrm{v}|)$ for two different choices of the pairs $(\kappa_{0+},\kappa_{2+})$ and some varying values of μ. For both the cases of $(\kappa_{0+},\kappa_{2+})=(2,1),\ (3,1)$ (in the figure the upper and lower respectively), we observe that the temperature profiles are the same but the velocity profiles are changing as expected.

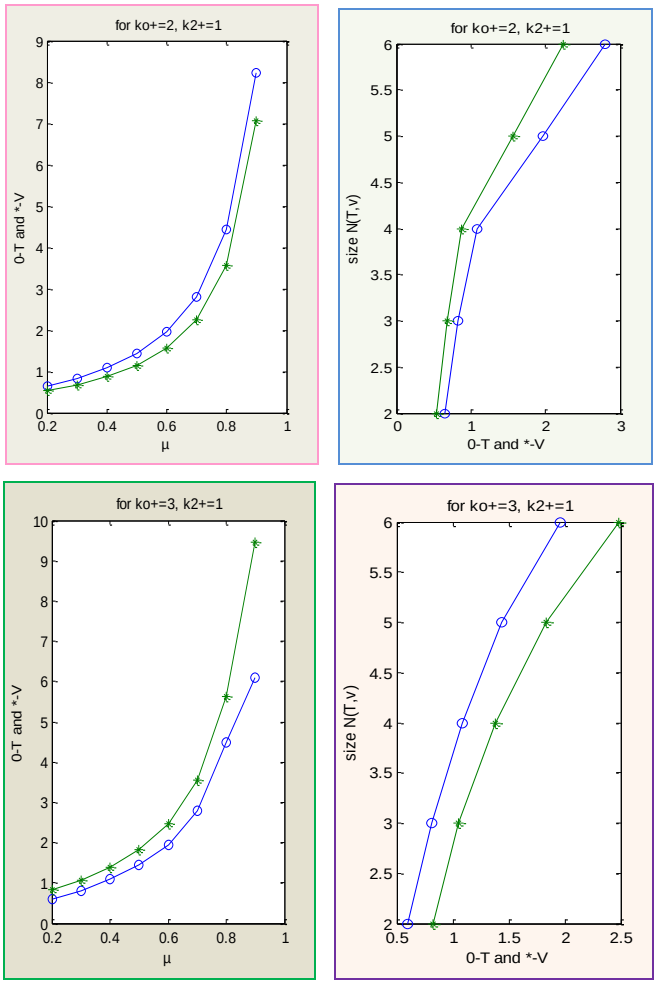


Figure 6.4: Size $N(T,|\mathrm{v}|)$ depends on temperature T and bulk-velocity v.

7 Conclusion

Using four parameters, characterizing mass, (x,y)-momenta and kinetic energy, the equilibria f of the Boltzmann equation can be expressed. The computation of discrete equilibria of the Boltzmann equation is effected by varying temperature and bulk-velocity. The temperature depends on the parameter μ, characterizing kinetic energy. The bulk-velocity depends on the parameters κ's, characterizing (x,y) momenta and μ as well. Error estimation of the discrete equilibria of the discrete model leads to determine the appropriate grid size for a given mass, bulk-velocity and temperature. Restricting the error in a given tolerance one can investigates efficient numerical scheme to solve the space inhomogeneous Boltzmann equation which we may investigate in our future work.

Acknowledgment

I sincerely acknowledge Jahangirnagar University, Bangladesh for providing all sorts of support to carry out this research work.

REFERENCES

[1] H. Babovsky, "Hierarchies of reducible kinetic models, Discrete Modeling and Discrete Algorithms in Continuum Mechanics", *Th. Sonar and I. Thomas (Eds.), Logos Verlag, Berlin, 2001*.

[2] H. Babovsky, "A kinetic multiscale model", *Math. Models Methods Appl. Science*, Vol. 12, pp.309-331, 2002.

[3] L. S. Andallah, "A hexagonal collision model for the numerical solution of the Boltzmann equation", Ph.D. thesis, University Library, TU Ilmenau, Germany, 2005.

[4] L. S. Andallah, and H. Babovsky, "A discrete Boltzmann equation based on hexagons", *Math. Model Methods Appl. Science*, Vol.13, No.11, pp.1-28, 2003.

[5] L. S. Andallah, "On the generation of a hexagonal collision model for the Boltzmann equation", *Comp. Meth. in Appl.Math.*,4(3), 267-285, 2004.

[6] L. S. Andallah, and H. Babovsky, "A discrete Boltzmann equation based on a loosely coupled hexagonal discretization of R^2" *Jahangirnagar University Journal of Science*, Vol.28, 2005.

[7] L. S. Andallah, H. Babovsky and A. Z. Khan "Spurious invarients of the Boltzmann operator discretized on a loosely coupled hexagonal grid" *Jahangirnagar University Journal of Science*, Vol.29, pp.51-62, 2006.

[8] L. S. Andallah "Generation of mesh for the numerical solution of the Boltzmann equation based on a loosely coupled hexagonal discretization", *Jahangirnagar University Journal of Science*, Vol.31, No.1, pp.59-68, 2008.

[9] W. Z. Loskor "Equilibrium Solution and Error Estimation of a Boltzmann Equation Discretized on a Loosely Coupled Hexagonal Grid", Ph.D. thesis, Jahangirnagar Library, Savar, Dhaka, 2012.

[10] W. Z. Loskor and L. S. Andallah "Generation of Equilibria of the Boltzmann Equation Discretized on N-Layer Loosely Coupled Hexagonal Grid" *Jahangirnagar University Journal of Information Technology*, Vol. 3, pp. 9-14, 2014.

Wahida Zaman Loskor received the B. Sc. and M. Sc. degrees in Mathematics from Jahangirnagar University, Savar, Dhaka. She awarded M. Phil. Degree in Mathematics from BUET, Dhaka and Ph. D degree from Jahangirnagar University, Savar, Dhaka. Her Ph.D research is on the Equilibrium Solution and Error Estimation of a Boltzmann Equation Discretized on a Loosely Coupled Hexagonal Grid. She worked as a full time faculty member in the Department of Computer Science & Engineering of Gono Bishwabidyalay, Savar, Dhaka from 2001 to 2015. Now, she has been working in the Faculty of Science & Humanities, Bangladesh Army International University of Science & Technology (BAIUST), Comilla Cantonment, Comilla, Bangladesh from 2015 as a full time faculty. She has published a number of papers in peer reviewed journals.

Optimizing SQL Performance in a Parallel Processing DBMS Architecture

[1]Nayem Rahman and [2]Leonard Sutton
[1]Cross Enterprise Systems IT, Intel Corporation, Hillsboro, OR, USA,
Email: nayem.rahman@intel.com
[2]Independent Teradata Consultant, The Boeing Company, Seattle, WA, USA,
LeonardSuttonLLC@bctonline.com

Abstract

A Database Management System (DBMS) with a parallel processing architecture is different from conventional database systems. Accordingly, writing SQL for a parallel processing DBMS architecture requires special attention to maintain parallel efficiency in DBMS resources usage such as CPU and I/O. In a large data warehouse, a large number of SQL queries are executed by different user groups on a daily basis. Query response time needs to be minimal. Many batch jobs run to refresh data warehouse subject areas several times a day. To allow batch cycles run more frequently and keep the data warehouse environment stable the database system's resource utilization must be optimal. Running efficient queries is critical to keep resource utilization manageable. This article discusses the techniques of SQL writing, tuning, utilization of index, data distribution techniques in a parallel processing DBMS architecture. We hope that these techniques will empower SQL developers and business intelligence community to write efficient queries which will help maintain a stable data warehousing environment.

Keywords—Database Architecture, DBMS, Computing Resources, CPU, I/O, Data Warehouse, Parallel Processing

An earlier version, "SQL Optimization in a Parallel Processing Database System" byRahman, N. appeared in Proceedings of CCECE '13. ©2013 IEEE.

© University of Liberal Arts Bangladesh
All rights reserved.

Manuscript received on 29 August 2016 and accepted for publication on 17 September 2016.

1 Introduction

TODAY'Sbusiness organizations use data warehouse as a central repository of data that come from internal operational sources as well as external sources (includes big data) [15]. As business organizations become global, there is a need to run business operations twenty four-by-seven so business decisions could be made faster. The data warehouse plays a prominent role in providing business intelligence (BI) capabilities [46]. It has proved to be one of the key infrastructures of information technology for an organization to better manage and leverage its information [31, 45]. Data warehouses are used for target marketing, financial reporting, customer services, inventory management, and more. They keep changing the way business is conducted [9]. Research suggests that data warehouses are increasingly being used by medium and large companies as these organizations are realizing its benefits [36].

Due to global nature of business and increased competition the data warehouse users and analytical community want to get near real-time information for strategic and tactical decision making. With the increased capabilities of advanced database technologies and massive parallel processing systems, it is now possible to load, maintain, and access databases of terabyte size [14] in reasonable times. In order to maintain a stable data warehousing environment data warehouse design, SQL writing, and load techniques all need to be efficient [2, 27]. Strategies are needed to save database management system (DBMS) resources during load processes in order to make the DBMS available to analytical tools and query processing while data warehouse refreshes continue at the same time. The data warehouse SQL queries for both load process and reporting need to be efficient. In this article we propose a comprehensive list of SQL query optimization techniques. We argue that data warehouse resource consumption could be made optimal by taking advantage of parallel processing architecture of database system.

This article is organized as follows: Section 2 briefly discusses related work done in this area. Section 3 discusses our proposed techniques of performance optimization. Section 4 discusses SQL parallel efficiency implementation steps and DBMS resource savings. Section 5 provides SQL query performance metrics of use cases. Section 6 summarizes and concludes the article.

2 Literature review

Researches have been conducted in different area of data warehousing. These include design issues [11, 13, 17, and 19], extract-transform-load (ETL) tools [23, 43], temporal data updates [18], data warehouse automation [24], data maintenance [1], implementation issues [21] and implementation effectiveness [36]. In this paper, we attempt to address the question of how to make a data warehousing environment stable and how to keep resource consumption by individual queries optimal by virtue of efficient SQL writing.

In data warehousing and data management systems parallel processing architecture is considered as a key capability [10]. Database system performance and SQL query optimization are important in any database system [39, 47]. In real world, the SQL queries that get executed are often quite complex and for data mining tasks queries are even more complex and resource intensive [38]. Hence, SQL query optimization is very critical for data warehouse stability.

In order maintain a stable data warehouse system in terms of resource utilization researchers and industry technical leaders propose many tools, techniques, algorithms and strategies. Here we take a cursory look at them. Ghazal et al. [20] present an algorithm that dynamically chooses between saving and re-using compiled plans and minimize re-compiling queries. Ganguly et al. [16] show that a cost model can predict response time with features of query execution parallelism. Kashem et al. [25] present a query optimization algorithm in rank aware queries to efficiently answer to the queries with join of N relations. Rahman and Rutz [24] assert it is critical to ensure that processes in the data warehouse are automated and optimized for performance. The authors propose using automation tools in a data warehouse ETL process, SQL block generations for views, stored procedures and macros wherever possible.

Elnaffar et al. [12] state that a DBMS workload could be considered as a determinant of performance tuning techniques. The authors argue that DBMS workloads are different in terms of OLTP and DSS. They propose reconfiguring DBMS resources by automatically identifying the DBMS work load. DSS queries process huge volume of data. Hence, they take more resources than OLTP queries. Dayal et al. [7] and Sharaf and Chrysanthis [41] propose managing database workloads with mixture of OLTP-like queries that run for fraction of a second and on the other hand, business intelligence queries that run for a longer time. The standard benchmark for Decision Support Systems comprises database workload and query performance metrics [42]. Powley et al. [40] present query throttling techniques as method to control workload. Kerkad et al. [26] propose a query beehive algorithm for data warehouse buffer management and query scheduling to improve data warehouse system's performance and scalability. Rahman [9] proposes a balanced scorecard approach for measuring performance of data warehouse operations.

Meng et al. [34] propose logically splitting large queries so each of them deal with small set of data and cause less impact on the overall warehouse environment and thus avoid consuming huge resource by one single large query. Narasayya et al. [35] propose a buffer pool page replacement algorithm that effectively shares buffer pool memory in multi-tenant relational database-as-a-service (DaaS). VanderMeer et al. [44] propose a cost-based database request across a cluster of databases to spread workload and resource usage. Neumann [37] asserts that query optimizer needs to be more efficient to efficiently handle different types of SQL queries. The author argues that query optimizer has larger impact than that of runtime system.

Hill and Ross [22] present a method to make outer joins efficient in order to improve query performance and response time. Rahman [18] discusses performance improvement of load and report queries, and maintenance of views with temporal data. Armstrong [4] proposes reduction of data movement to increase user accessibility, minimize data latency and improve performance of the entire data warehouse. Krompass et al. [28, 29] propose a workload management system for managing the execution of individual queries based on customer service level objectives. Rahman [49] proposes strict governance in data warehouse maintenance and operations to bring discipline and control. This includes defining guidelines for application developers and IT integration engineers to follow. The author presents a set of data warehouse governance best practices with insights from real-world experience and research findings from industry and academic papers.

Allen and Parsons [3] demonstrate that anchoring and adjustment during query reuse by novice query writers can lead to queries that are less accurate than those written from scratch. This suggests that in real-world SQL queries could be written by users and developers of varieties skill-set. A significant number of them could be badly written. Hence, SQL queries need to run through some SQL score-card process [1] to ensure parallelism of query runs. Lee et al. [30] propose a Statistical Process Control (SPC) charts to detect database performance anomalies and identify their root causes. However, performance anomalies could be prevented from happening if each SQL queries could be run

through a SQL performance scorecard process [1].

In this article, we focus on writing efficient SQL that conforms to parallel processing architecture. We address the problem of DBMS resource consumption and stability issue by taking care of SQL efficiency, defining indexes and many other SQL optimization techniques. By taking advantage of database parallelism architecture the problem of SQL query response time could be minimized [8]. This helps in achieving database system resources (CPU and I/O) saving [5].

3 Performance Optimization in a Parallel Processing DBMS

In a parallel processing DBMS architecture a large number of individual Access Module Processors (AMP) are used. We can think of these as "Units of Parallelism". Each "unit" will have dedicated Disk and dedicated CPU. The goal of the Physical Database Design, and the design of the SQL submitted, is to force the processing to be as well distributed across all the AMPs as possible. Because the CPU and other resources are shared with other jobs across the system, the actual impact of any given process is: the highest amount of resources used on any one AMP, times the number of AMPs on that particular system. If the high AMP uses 80 CPU and 2000 I/O, and we have a 100 AMP system, then the real impact of that job is as if it used 8000 CPU and 20,000 I/O, even if the total resources used by all the AMPs appears to be a much smaller amount. When a query executes, each step in the process waits until all the AMPS are finished for a given step, before the next step starts. For this reason, the most efficient processes are the ones which have about the same amount of resources used on each AMP.

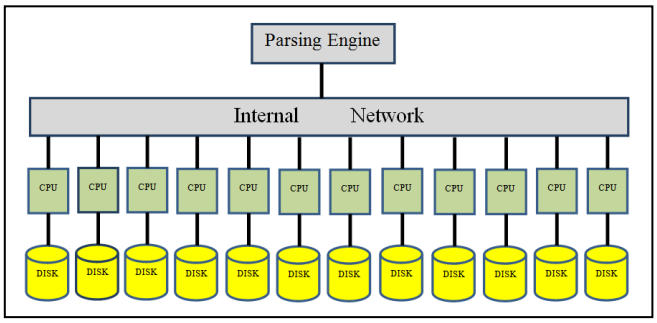


Figure 1: A Parallel Processing DBMS Architecture.

Skewed processing is when there is significant difference between the resources used by the "high" AMP and average of all AMP's. If the Physical Database Design has been verified to be optimizes, then attention can be given to the SQL being submitted. This document deals with different way to optimize the SQL.

3.1 Row Redistribution in a Parallel Processing Architecture

In most large systems, a typical report will need to look at many tables. In a parallel processing database system, care must be taken in choosing Primary and Secondary Indexes, to try to avoid "redistribution" steps in the SQL Parsing steps. The optimizer joins 2 tables at a time and puts the result into a spool file. Then it joins that to another table or spool file, and so on until all the tables are joined. On each of these joins the rows to be joined on each table must reside on the same AMP. If the 2 tables have the same primary index (PI) then all the rows that will join together already reside on the same AMP. If the 2 tables have different PI's the DBMS needs to do one of two things: either duplicate (one of the) table(s) on all AMPs or redistribute one of the tables (using a PI that is the same as the other table) so that the rows being joined now reside on the same AMP. So the reason for redistribution is always that the 2 tables being joined do not have the same PI. Sometimes we cannot do anything about this; it is just the way it works. Other times, we can build a derived table, narrowing the selection of rows to a smaller number, and try to make the optimizer duplicate the table on all AMPS. If it does not disturb other processes; the best way to eliminate redistribution is to build the tables being joined with the same PI (this is not always possible). Depending on the choice of indexes, this join process can have very different paths to get to the desired results.

3.2 Duplicating on all AMPs and Product JOINs

Sometimes, the optimizer sometimes builds a copy of a table on each unit of parallelism to facilitate parallel processing. There are many cases where this proves to be the best path for the optimizer to take. To ensure that this duplicating takes less resource, a derived table can be used in the SQL, creating a reduced set of rows and/or columns for the optimizer to work with.

Sometimes a Product-Joins occur when the optimizer needs to join a large and a small table. To improve performance: narrow down the rows and columns of that small table; if the smaller table contains static data with few records in that case column values could be placed in memory variables (Figure 2). That way, a JOIN with the smaller table could be entirely eliminated.

```
,CASE
        WHEN ANLC.fscl_yr_mo_nbr = :fscl_yr_mo_nbr THEN :last_day_curr_mo
        WHEN ANLC.fscl_yr_mo_nbr = :Mo1 THEN :Dt1
        WHEN ANLC.fscl_yr_mo_nbr = :Mo2 THEN :Dt2
        WHEN ANLC.fscl_yr_mo_nbr = :Mo3 THEN :Dt3
        WHEN ANLC.fscl_yr_mo_nbr = :Mo4 THEN :Dt4
        WHEN ANLC.fscl_yr_mo_nbr = :Mo5 THEN :Dt5
        WHEN ANLC.fscl_yr_mo_nbr = :Mo6 THEN :Dt6
        WHEN ANLC.fscl_yr_mo_nbr = :Mo7 THEN :Dt7
        WHEN ANLC.fscl_yr_mo_nbr = :Mo8 THEN :Dt8
        WHEN ANLC.fscl_yr_mo_nbr = :Mo9 THEN :Dt9
        WHEN ANLC.fscl_yr_mo_nbr = :Mo10 THEN :Dt10
        WHEN ANLC.fscl_yr_mo_nbr = :Mo11 THEN :Dt11
        WHEN ANLC.fscl_yr_mo_nbr = :Mo12 THEN :Dt12
        WHEN ANLC.fscl_yr_mo_nbr = :Mo13 THEN :Dt13
        WHEN ANLC.fscl_yr_mo_nbr = :Mo14 THEN :Dt14
        WHEN ANLC.fscl_yr_mo_nbr = :Mo15 THEN :Dt15
        WHEN ANLC.fscl_yr_mo_nbr = :Mo16 THEN :Dt16
        WHEN ANLC.fscl_yr_mo_nbr = :Mo17 THEN :Dt17
        WHEN ANLC.fscl_yr_mo_nbr = :Mo18 THEN :Dt18
        WHEN ANLC.fscl_yr_mo_nbr = :Mo19 THEN :Dt19
        WHEN ANLC.fscl_yr_mo_nbr = :Mo20 THEN :Dt20
        WHEN ANLC.fscl_yr_mo_nbr = :Mo21 THEN :Dt21
        WHEN ANLC.fscl_yr_mo_nbr = :Mo22 THEN :Dt22
        WHEN ANLC.fscl_yr_mo_nbr = :Mo23 THEN :Dt23
        WHEN ANLC.fscl_yr_mo_nbr = :Mo24 THEN :Dt24
        ELSE DATE '9999-12-31'
END AS depr_hist_clndr_dt
```

Figure2: Eliminate a skewed JOIN and populate column with memory variable values.

3.3 Parallel Efficiency

Skewed data distribution and skewed processing adversely affect parallel efficiency. Poor parallel efficiency occurs when the join field is highly skewed. Rows are redistributed to AMPs based on the join column values; a disproportionate number of rows may end up on one AMP on or a few AMPs operation. Highly non-unique PIs cause uneven row distribution. More than 1000 occurrences of a value in a Non-Unique Primary Index (NUPI) value begin to cause performance degradation problems: Increased I/O's for updates and inserts of over-represented values; Poor CPU parallel efficiency on full table scans and bulk inserts.

```
CREATE TABLE Capital_DRV_MET.fact_purch_doc_line
(
  asof_src_dt DATE NOT NULL,
  asof_src_ts TIMESTAMP(0) NOT NULL,
...
  src_sys_nm CHAR(20) CHARACTER SET LATIN NOT NULL,
  purch_doc_nbr CHAR(10) CHARACTER SET LATIN NOT NULL,
  purch_doc_line_nbr CHAR(5) CHARACTER SET LATIN NOT NULL,
  purch_doc_line_shrt_dsc VARCHAR(40) CHARACTER SET LATIN,
...
)
PRIMARY INDEX xfact_purch_doc_line02 (purch_doc_nbr,purch_doc_line_nbr);
UNIQUE INDEX xfact_purch_doc_line01 (purch_doc_nbr ,purch_doc_line_nbr,src_sys_nm );
```

Figure 3: Primary Index defined with two columns for better row distribution.

Figure 3 shows a Primary Index (PI) with two columns to make sure rows are distributed to all AMPs. Initially we defined index with a single column, that is, with 'purch_doc_nbr' only. But, since there are a large number of the same 'purch_doc_nbr' we redefined PI consisting of two columns. Addition of the second column, 'purch_doc_line_nbr' made data distribution much better. Table load performance has improved significantly. If there is still a need for an index on purch_doc_nbr, we can build a secondary index

3.4 Primary Index Choice Criteria

There are several things to consider when choosing primary and secondary indexes in a parallel processing environment. Because some indexes are chosen based on usage of the data in reporting, there might be some testing needed, later in the development process to arrive at the best possible set of indexes. The primary index of the table does not necessarily need to be a unique index.

Access Demographics: Columns that would appear with a value in a WHERE clause. Choose the column most frequently used for access to maximize the number of direct, single-row access operations. Distribution Demographics: The more unique the index, the better the distribution.

```
PRIMARY INDEX xfact_pr02(pr_nbr)
UNIQUE INDEX xfact_pr01 (src_sys_nm, pr_nbr, pr_line_nbr, acct_asgn_nbr);
```

Figure 4: Primary Index defined on a column most often used as a filter.

Volatility: The data values should not change quite often. Any changes to PI values may result in heavy I/O overhead. Join activity should dictate the PI definition. For large tables, the number of Distinct Primary Index values should be much greater than the number of units of parallelism.

3.5 Synchronizing Source and Target Primary Indexes

Common indexes between source and target tables help bulk inserts. The optimizer performs index-based MERGE JOINs. In a large join operation, a merge join requires less I/O CPU time than nested join. A merge join usually reads each block of the INNER table only once, unless a large number of hash collisions occur. In a real world scenario we noticed that due to missing common primary indexes, the SQL of a stored procedure became 90% skewed. It pulled records from two large tables with several join columns. Run time was 5 hours and 6 minutes to load 9 million rows. After PI synchronization the run-time dropped to 1 minute 11 seconds.

3.6 Deriving Common PI's Between Source Tables

Creating and populating a Global Temporary table helps in avoiding uneven PIs and eliminate LEFT OUTER JOIN in the Final INSERT-SELECT step (second INSERT in Figure 5).

```
INSERT INTO TABLE2.GT_Asset_Cost_Centr
  SELECT AstDrv.Asset_Nbr,AstDrv.Asset_Sub_Nbr
  ,AstDrv.Co_Cd,CC.Prft_Centr_Sap_Nbr
  ,CC.Prft_Centr_Char_Nbr
  FROM Asset.v_asset_DRV  AstDRV
  LEFT OUTER JOIN
  Capital_Analysis.v_dim_cost_centr_curr  CC
  ON AstDRV.cost_centr_cd = CC.cost_centr_sap_nbr;
INSERT INTO
  FROM Asset.v_asset_DRV  AstDRV
  INNER JOIN
  TABLE1.GT_Asset_Cost_Centr CC
  ON AstDrv.Asset_Nbr = CC.Asset_Nbr
AND AstDrv.Asset_Sub_Nbr = CC.Asset_Sub_Nbr
AND AstDrv.Co_Cd = CC.Co_Cd;
```

Figure 5: Deriving a Common PI for Parallel Efficiency.

In a simulation of SQL-run we found that total CPU consumption dropped to 122 second, yielding a 44.68 sec savings. The total I/O operation dropped to 111,192, yielding a 70,632 savings.

3.7 Temporary Tables versus Derived Tables

The solution to some of the resource intensive queries includes conversion of a derived table (DT) to a global temporary table (GTT). This is because the GTT can have statistics collected whereas the DT cannot. The GTT approach makes the optimizer plans more aggressive and rely more heavily on collected statistics as opposed to sampled statistics. As in all of life, there is trade-offs: relying on collected stats would produce better running queries than the random samples. With data skew, the random samples were often wrong and caused wrong choices to be made. We can achieve better performance plans for tables (GTT) with collected statistics. We cannot collect statistics on derived nor volatile tables so these do not perform as well. Statistics collection on join and filter columns improve SQL query performance [6]. Figure 6 shows performance results of an SQL that used derived tables. The result shows that per evaluation criteria the SQL failed in terms of computing resources such as CPU, IO and spool space usage. Their parallel efficiencies are very poor.

SQL Scorecard Results

Project Name: Procure Analysis MSTR
Developer Name: Wally Heaven
System used for Testing: TDCNS1

ETL		CPU Evaluation		I/O Evaluation		Spool Evaluation		EXPLAIN Evaluation		System Rating			
SQL ID	Procedure Name (SQL)	Total CPU	CPU Parallel Efficiency (%)	Total I/O	I/O Parallel Efficiency (%)	Total Peak Spool	Spool Parallel Efficiency (%)	Statistics on all Joins & Filters?	Joins or Filters on Derived Attributes?	CPU : I/O Ratio	Resource Usage Rating	Parallel Efficiency Rating	Overall Score
1	pr_Fpo_Imt_rpt	1,642	71.65	1,231,559	31.5	5,610,799,616	5.71	YES	NO	1.33	FAIL	FAIL	FAIL

Figure 6:Resource Usage with an SQL that uses derived tables.

Figure 7 shows that each SQL passed in terms of performance evaluation criterion. Computing resources consumption such CPU, IO and spool usage is much lower compared to the resources used shown in Figure 6. Each SQL alsoshows that they higher parallel efficiency.

SQL Scorecard Results

Project Name: Procurement Analysis MSTR
Developer Name: Wally Heaven
System used for Testing: TDCNS1

ETL		CPU Evaluation		I/O Evaluation		Spool Evaluation		EXPLAIN Evaluation		System Rating			
SQL ID	Procedure Name (SQL)	Total CPU	CPU Parallel Efficiency (%)	Total I/O	I/O Parallel Efficiency (%)	Total Peak Spool	Spool Parallel Efficiency (%)	Statistics on all Joins & Filters?	Joins or Filters on Derived Attributes?	CPU : I/O Ratio	Resource Usage Rating	Parallel Efficiency Rating	Overall Score
1	gt_po_lmt_rpt - TEMP INSERT	7	69.79	53,001	96.18	40,266,240	93.86	YES	NO	0.14	PASS	PASS	PASS
2	gt_po_lmt_rpt1 - TEMP INSERT	14	57.74	59,305	84.05	57,574,240	96.65	YES	NO	0.23	PASS	PASS	PASS
3	gt_po_lmt_rpt2	164	87.53	183,323	85.59	1,634,344,448	60.94	YES	NO	0.89	PASS	PASS	PASS
4	gt_po_lmt_rpt2 - UPDATE	6	64.82	25,538	80.36	350,208	89.06	YES	NO	0.25	PASS	PASS	PASS
5	po_lmt_rpt - FINAL INSERT	9	62.37	12,428	68.5	71,610,368	72.77	YES	NO	0.75	PASS	PASS	PASS

Figure 7: Resource Usage with SQL's that use GTT.

3.8 Handling NULLs for Better Parallel Efficiency

When performing a LEFT OUTER JOIN operation with a column with so many common values performance of join operation degrades. It is important that NULL or blank values be filtered out in the SQL while doing join operation. An MPP (Massively Parallel Processing) machine can get very slow when there is nothing to "parallel process".

```
FROM Procurement.v_funds_rsrv_line AS LINE
LEFT OUTER JOIN
Material.v_mtrl_CRS AS mtrl
ON line.mtrl_id = mtrl.mtrl_id

Results:
-------------
mtrl_id                    Count(*)
000000000000000001          2
NULL/blank                  463950
```

Figure 8: Avoiding NULLs for Parallel Efficiency.

Figure 8 shows a scenario in which case only 2 rows with useful values. The rest of the rows show NULL/BLANK which severely impact database optimizer to perform an MMP.

3.9 Avoiding Updates between Large Tables

When tables are large SELECT/INSERT performs much better. Update is good when source table has fewer rows. An example provided in Figure 9.

```
UPDATE TrgtTbl
  FROM Table1.v_fact_depr_hist_base  TrgtTbl
       ,Table2.gt_fact_depr_hist_anlp4  SrcTbl
  SET calc_curr_qtr_mo1_ordnry_amt = SrcTbl.depr_to_be_post_ordnry_amt
     ,calc_curr_qtr_mo1_unpln_amt = SrcTbl.depr_to_be_post_unpln_amt
  WHERE TrgtTbl.co_cd  = SrcTbl.co_cd AND TrgtTbl.asset_nbr = SrcTbl.asset_nbr
  AND TrgtTbl.asset_sub_nbr = SrcTbl.asset_sub_nbr
AND TrgtTbl.depr_area_cd = SrcTbl.depr_area_cd
  AND TrgtTbl.fscl_yr_mo_nbr_drv1 = SrcTbl.fscl_yr_qtr_mo1_nbr ;
```

Figure 9: Performance Degrades with large volume of Updates.

In one scenario the UPDATE operation by joining large source table caused CPU consumption of 1,013 seconds. If we need to use a subset of data from large tables using global temporary tables will help in computing resource consumption. In a simulation we noticed that by using global temporary tables for a sub-set of data in source table the UPDATE operation took only 600 CPU seconds.

3.10 Partitioned Primary Index

If they are available, Partitioned Primary Indexes (PPI) can be very productive. A PPI is equivalent to row level partitioning. Queries which specify a restrictive condition on the partitioning column avoid full table scans. Larger tables are good candidates for partitioning. The greatest potential gain derived from partitioning a table is the ability to read

a small subset of the table instead of the entire table.

Current commercial databases have come up with efficient indexes to improve query performance. When a query is run with filters on PPI columns the DBMS will directly pull data based on particular bucket(s) instead of scanning the whole table. Based on a SQL score-card on both PPI and non-PPI tables it was found that the SQL uses only 33% of the resources to pull rows from a PPI table in relation to a non-PPI table. The run time is also less in the same proportion. The potential gain derived from partitioning a table is the ability to read a small subset of the table instead of the entire table. Queries which specify a restrictive condition on the partitioning column avoid full table scans. By defining a PPI on 'row effective timestamp' the report query performance was found to be four times faster and CPU savings about 33%.

Resource Usage: PPI vs. Non-PPI Table

- Query-1: Rows Returned: 53262
  - Non-PPI Table (01): Response Time: 8 sec: CPU=29
  - PPI Table (02): PPI defined on asof_src_dt; Response Time: 1 sec; CPU= 2 sec
- Query-2: Rows Returned: 424076
  - Non-PPI Table (07): Response Time: 25 sec: CPU=101
  - PPI Table (09): Response Time: 4 sec; CPU=33 sec

Report Name (SQL)	Total CPU	CPU Parallel Efficiency (%)	Total I/O	I/O Parallel Efficiency (%)	Total Peak Spool	Spool Parallel Efficiency (%)
v_ctrl_doc_captl_spnd_CURR_01.out	29	64	286,031	66	35,580,416	13
v_ctrl_doc_captl_spnd_CURR_02.out	2	15	2,122	33	35,580,416	13
v_ctrl_doc_captl_spnd_CURR_07.out	101	58	210,329	67	664,387,584	46
v_ctrl_doc_captl_spnd_CURR_09.out	33	49	11,108	68	455,382,016	32
	164		509,590		1,190,938,432	

Figure 10: Resource Usage: PPI vs. No PPI tables.

Figure 10 shows a comparison of query response time and computational resource savings between PPI and No-PPI queries. The first query was run to pull 53K rows, with no PPI defined. The response time was eight seconds and CPU consumption was 29 seconds in row one. The same query was run against the same table with PPI defined on row effective date. For the second run the response time was one second and resource consumption was two seconds per row two. The first two rows show the resource usage statistics. A second query was run to pull 424K rows, with no PPI defined. The response time was 25 seconds and resource consumption was 101 CPU seconds in row three. The same query was run against the same table with PPI defined on row effective date. This second run response time was four seconds and resource consumption was 33 seconds in row four.

There are many techniques to improve performance of data warehouse queries, ranging from commercial database indexes and query optimization. A number of indexing strategies have been proposed for data warehouses in literature and are heavily used in practice.

4 SQL Parallel Efficiency and DBMS Resource Usage

In a data warehouse where thousands of queries run by batch processes, analytical and ad-hoc queries and applications all run concurrently, the computing resources are the most precious resources. These computing resources need to be used very efficiently [33] to keep the data warehousing environment stable and running. The analytical community cannot tolerate long running queries or delayed results. Response time of queries is one of the most important indicators of data warehouse stability and its success. The knowledge workers lose confidence in the system if the enterprise data warehouse cannot return information within a reasonable time, especially when it comes to tactical decision making. Transaction latency expressed as a deadline is the most commonly used form of SLA [32], reflecting the user's expectation for the transaction to finish within a certain amount of specified time [41].

In order to ensure the data warehouse is stable, scalable, and queries run efficiently many organizations institute a governing body to oversee the operation and running of the data warehouses. They closely monitor the deployment of objects such as views, stored procedures and macros to make sure they perform efficiently in the data warehouse. In most cases all code that lands on data warehouses goes through a code review process to make sure they are optimized. As a cross-check the DBA (database administrators) team constantly monitors queries and load procedures to make sure the data warehouse is stable and running efficiently. Some things to watch for, to help with parallel efficiency:

4.1 Large Distribution Steps

Occasionally, there will be a job with a step which takes a lot of resources, just to get two tables ready for a join step. This happens (as mentioned earlier) when the two tables being joined do not have the same Primary Index (PI). Sometimes, one of the tables can be changed, so the PI's are the same. When the PI's are the same, the large redistribution step is eliminated, sometimes with very nice results. However, we need to be careful not to introduce too much non-uniqueness in the PI. Table 1 shows a real example of such results. The first row shows that almost all of the resources used on this job were spent in the redistribution step. However, once redistribution steps are eliminated most of there-source use disappeared (second row).

Table 1
Resource Saving Avoiding Row Redistribution

	Elapsed Time	I/O	CPU	Spool
before	00:05.1	00:00.0	31:12.0	510,038,016
after	00:00.1	00:00.0	09:36.0	0

4.2 Secondary Indexes

Sometimes we cannot just change the PI of a table for the purposes of helping a join step. There are many reasons for choosing the PI of a table. The first criteria should be Reporting Access Requirements. In these cases, it is possible to add a Secondary Index, the same as what we would have liked for a PI.

Table 2
Computing Resource Saving using Secondary Index

BEFORE					
LogDate	QryCount(*)	AvgElapsedTime	AvgCPUTime	AvgSpoolUsage	AvgIOCount
10/8/2012	717	00:02.2	8.70	503,352,026	88,810
AFTER					
LogDate	QryCount(*)	AvgElapsedTime	AvgCPUTime	AvgSpoolUsage	AvgIOCount
11/12/2012	717	00:01.1	0.27	1,639,843	13,663

In one such case, we had a lot of queries doing the same join, so we added a Secondary Index to a table. This helped the JOIN Condition find the applicable rows faster. This is bit different from an Index which helps find the rows being selected.

Our experiment shows that resource reduction for one day was 6,000 CPU seconds, spool space 350 gigabytes, and I/O reduction 50 million. Elapse time was 15 minutes.

4.3 Partitioning Rows which are Accessed Often

In this next case, we found that a lot of queries were asking for rows within a given date range, so we added Partitioning, in a way that reduced tables scans to a smaller set of data-blocks. Table 3 represents a set of queries for 1 day's activity.

Table 3
Using Partitioning to Avoid or Reduce Table-Scans

Log Date	UserName	Before Query Count	After Query Count	Before Avg CPU	After Avg CPU	CPU Saved
6/20/2011	MMBIETL1	14,037	15,535	0.25	0.23	286
6/20/2011	A341545	7,154	10,724	0.36	0.07	3111
6/20/2011	MMBIAP01	6,339	6,298	9.48	8.09	8771
6/20/2011	A130330	2,224	2,344	1.28	0.96	759
6/20/2011	MMBI001	2,115	2,314	46.93	43.06	8956
6/20/2011	A1887155	9	57	63.03	16.96	2626
						24510

4.4 Skewed Processing

Sometimes we chose a certain PI to help reporting, but allows too many rows to be stored on 1 or just a few AMPs (Units of Parallelism). We discussed Skew in a previous section. In this case, we can use a different PI to spread the data more evenly, then build a Secondary Index where we removed the first Primary Index.Here is a case where we did this. We use a bit more resource with a Secondary Index over a Primary, but we save a lot more than that by spreading the work out over the AMPS more evenly.

In Figure 11, we can see where we installed the change at 14:00 hours: There is a set of jobs which run every hour. We can see the reduction in resources for the next hours.

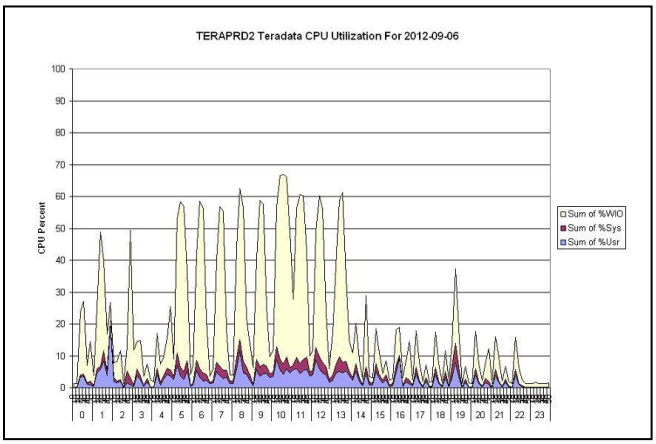


Figure 11: Resource Saving by Improving Parallel Efficiency.

4.5 Using Set versus Multi-Set Tables

SET tables do not allow duplicate rows – Multi-set tables do. The combination of a SET table, and a Non-Unique Primary Index, can dangerous if there is no other uniqueness constraint on the table (such as a Unique Secondary Index). If there is a Unique Secondary Index, the table does not need to "worry" about checking for duplicate rows (because the Index will be checking). If there is no other unique constraint on the table, when we have multiple rows with that same PI, as rows are inserted, the table needs to check all rows with the same PI to see if in fact the whole row is a duplicate. If it is a full row duplicate, it will not be allow to be inserted. This dupe-row-checking can get very expensive if there are a lot of rows with the same (non-unique) PI.

Table 4
Resource Savings – Set vs. Multi-Set tables

BEFORE LogDate	TotalIO	TotalCPU	StartTime	ElapsedTime	ImpactCPU	CpuSkew	SpoolUsage	Table being worked
4/8/2013	183,104.00	13.3	28:39.1	0:00:23.17	254.74	19.16	33,772,544	IMM_FCST_INSTALLATION_MV
4/8/2013	1,079,940.00	53.83	29:19.1	0:01:16.37	2,244.10	41.69	150,929,408	IMM_FCST_REQUEST_MV
4/8/2013	228,405.00	13.88	29:03.7	0:00:14.53	250.7	18.07	32,450,048	IMM_FCST_REMOVAL_MV
4/8/2013	943,791.00	36.86	30:48.0	0:00:55.51	1,520.78	41.26	73,801,728	IMM_FCST_ISSUE_MV
4/8/2013	190,319.00	10.06	30:36.2	0:00:10.95	234.86	23.35	41,681,920	IMM_FCST_RETURN_MV
	2,613,562.00		Totals	0:03:05.21	4505.18	Avg 43.53		
AFTER LogDate	**TotalIO**	**TotalCPU**	**StartTime**	**ElapsedTime**	**ImpactCPU**	**CpuSkew**	**SpoolUsage**	Table being worked
4/22/2013	123,034.00	3.76	27:30.7	0:00:02.57	7.49	1.99	21,797,376	IMM_FCST_RETURN_MV
4/22/2013	119,611.00	5.07	27:20.7	0:00:03.63	28.22	5.57	74,287,616	IMM_FCST_REMOVAL_MV
4/22/2013	347,427.00	8.3	27:25.1	0:00:04.86	30.96	3.73	113,038,848	IMM_FCST_REQUEST_MV
4/22/2013	312,617.00	8.71	27:34.2	0:00:04.99	18.58	2.13	117,903,360	IMM_FCST_ISSUE_MV
4/22/2013	176,712.00	6.66	27:15.3	0:00:04.55	47.66	7.15	89,515,008	IMM_FCST_INSTALLATION_MV
	1,077,901.00		totals	0:00:20.65	132.91	Avg 4.11		
Impove ments	**2X**			**9X**	**33X**	**10X**		

Let us suppose, there are 1000 rows with the same PI to be inserted. The second row inserted needs to only check 1 row for duplicates. The 100th row needs to check 99 rows. The 950th row needs to check 949 rows. The number of row checks would be 1+2+3+4+5…+999. Here are the results of a set of changes we made to a set of jobs which populated a few tables. We changed them from set to multi-set tables, to avoid duplicate row checking. And we changed the Primary Indexes to a column with less skewing. And we added a Secondary Index to replace the benefit we had with the old Primary Index.

5 Measuring SQL Performance

Performance statistics of SQL blocks in a stored procedure are shown in Figure 10. We show the score-card results of a stored procedure SQL blocks. The SQL's were written in such a way that they are in compliant with the parallel processing architecture of the underlying database system.

In Figure 12 we can see that each of the SQL's in a stored procedure passed in terms of CPU, I/O, and spool parallel efficiency. In the stored procedure SQL's were written in small code blocks. Each of them takes fewer CPU seconds; they run with parallel efficiency. We see all of the SQL blocks passed in score-carding.

SQL Scorecard

Project Name: Capital_DRV
Developer Name: Shameem Akhter
System used for Testing: TDEDW1

ETL		CPU Evaluation		I/O Evaluation		Spool Evaluation		EXPLAIN Evaluation		System Rating			
SQL-ID	Report Name (SQL)	Total CPU	CPU Parallel Efficiency (%)	Total I/O	I/O Parallel Efficiency (%)	Total Peak Spool	Spool Parallel Efficiency (%)	Statistics on all Joins & Filters?	Joins or Filters on Derived Attributes?	CPU : I/O Ratio	Resource Usage Rating	Parallel Efficiency Rating	Overall Score
1	pr_Ffact_grant_grir	4	62.56	22,064	94.97	34,054,656	64.87	YES	NO	0.18	PASS	PASS	PASS
2	pr_Ffact_grant_grir1	55	75.29	245,882	81.13	438,595,072	88.82	YES	NO	0.22	PASS	PASS	PASS
3	pr_Ffact_grant_grir2	3	54.77	22,488	85.49	503,808	85.42	YES	NO	0.11	PASS	PASS	PASS
4	pr_Ffact_grant_grir3	2	43.59	19,391	96.65	16,996,352	93.46	YES	NO	0.11	PASS	PASS	PASS
5	pr_Ffact_grant_grir4	2	57.75	15,148	87.18	4,651,008	81.57	YES	NO	0.13	PASS	PASS	PASS
6	pr_Ffact_grant_grir5	3	55.3	60,780	70.5	986,112	71.65	YES	NO	0.04	PASS	PASS	PASS
7	pr_Ffact_grant_grir6	4	61.53	63,890	71.18	986,112	71.65	YES	NO	0.06	PASS	PASS	PASS
8	pr_Ffact_grant_grir7	4	55.26	25,329	92.58	47,062,528	84.36	YES	NO	0.18	PASS	PASS	PASS
9	pr_Ffact_grant_grir8	2	65.89	15,855	86.02	1,014,784	71.19	YES	NO	0.13	PASS	PASS	PASS
10	pr_Ffact_grant_grir9	2	63.28	14,886	90.15	1,015,296	71.23	YES	NO	0.13	PASS	PASS	PASS
11	pr_Ffact_grant_grir10	3	58.56	27,569	96.05	17,804,288	93.12	YES	NO	0.11	PASS	PASS	PASS
12	pr_Ffact_grant_grir11	3	40.56	24,896	86.73	17,804,288	93.12	YES	NO	0.11	PASS	PASS	PASS
13	pr_Ffact_grant_grir12	3	51.19	66,406	72.66	1,973,760	83.66	YES	NO	0.04	PASS	PASS	PASS
14	pr_Ffact_grant_grir13	2	50.43	40,634	73.61	1,973,760	83.66	YES	NO	0.06	PASS	PASS	PASS
15	pr_Ffact_grant_grir14	2	65.49	15,162	88.73	1,015,296	93.89	YES	NO	0.13	PASS	PASS	PASS
16	pr_Ffact_grant_grir15	2	66.67	15,743	88.17	1,981,440	83.98	YES	NO		PASS	PASS	PASS
17	pr_Ffact_grant_grir16	2	49.53	14,415	84.36	1,981,440	83.98	YES	NO		PASS	PASS	PASS
18	pr_Ffact_grant_grir17	1	48.63	4,981	82.36	11,043,840	87.09	YES	NO		PASS	PASS	PASS
19	pr_Ffact_grant_grir18	2	45.94	8,798	91.65	8,102,912	83.26	YES	NO		PASS	PASS	PASS
20	pr_Ffact_grant_grir19	2	63.67	14,421	81.64	29,583,360	90.92	YES	NO		PASS	PASS	PASS
21	pr_Ffact_grant_grir20	2	36.89	7,740	94.85	19,288,064	94.56	YES	NO	0.31	PASS	PASS	PASS
		105		746,478		658,418,176		...	...	0.14	PASS		

Figure 12: Score-Card Results of SQL's of a Stored Procedure

6 Conclusion

In this article we provided an overview of a parallel processing DBMS architecture. We have highlighted as to what key aspects needs to be considered to take advantage of parallelism. We have provided an exhaustive list of techniques of SQL optimization. These techniques have been tested and implemented in a large production data warehouse system. In each of the optimization techniques we have provided computing resource savings as well query response time decrease statistics.

We proposed evaluating SQL queries using SQL scorecard tools. A scorecard process and performance optimization techniques will enable the SQL programmers to empower themselves in writing efficient queries without much dependence on database administrators. Currently, database administrators spend many hours inspecting various log files and queries [48]. Our proposed developer-centric SQL query optimization will help database administrators maintain a stable database system and its performance with much less effort. As part of our future research we intend to do explore optimization of queries in NoSQL database systems.

Acknowledgment

An earlier version of this article was presented at the IEEE 26th Canadian Conference of Electrical and Computer Engineering (CCECE 2013), Regina, Saskatchewan, Canada, May 5 - 8, 2013 [14].This article has been substantially revised and enhanced, based on comments received during conference presentation and comments received from anonymous reviewers of this journal. The author would like to thank the anonymous reviewers for their constructive comments provided on earlier drafts of this article.

References

[1] N. Rahman, "SQL Scorecard for Improved Stability and Performance of Data Warehouses," International Journal of Software Innovation (IJSI), Vol. 4, No. 3, pp. 22-37, July- September 2016.

[2] S. Akhter and N. Rahman, "Building a Customer Inquiry Database System," International Journal of Technology Diffusion (IJTD), Vol. 6, No. 2, pp. 59–76, April-June 2015.

[3] G. Allen and J. Parsons, "Is Query Reuse Potentially Harmful? Anchoring and Adjustment in Adapting Existing Database Queries," Information Systems Research, Vol. 21, No. 1, pp. 56-77, 2010.

[4] R. Armstrong, "When and Why to Put What Data Where," Teradata Corporation White Paper, 1-5, 2007.

[5] N. Rahman, "Saving DBMS Resources While Running Batch Cycles in Data Warehouses," International Journal of Technology Diffusion (IJTD), Vol. 1, No. 2, pp. 42–55, April-June, 2010.

[6] C.G. Corlatan, MM. Lazar, V. Luca and O.T. Petricica, “Query Optimization Techniques in Microsoft SQL Server,” Database Systems Journal, Vol. 5, No. 2, pp. 33-48., 2014.
[7] U. Dayal, H. Kuno, J.L. Wiener, K. Wilkinson, A. Ganapathi and S. Krompass, "Managing Operational Business Intelligence Workloads," ACM SIGOPS Operating Systems Review archive, Vol. 43, No. 1, pp. 92-98, 2009.
[8] S. Deepak, S.U. Kumar, M Durgesh and K.P. Bhupendra, "Query Processing and Optimization of Parallel Database System in Multi-processor Environments," In proceedings of the 2012 Sixth Asia Modelling Symposium (pp. 191-194).
[9] N. Rahman, "Measuring Performance for Data Warehouses - A Balanced Scorecard Approach," International Journal of Computer and Information Technology (IJCIT), Vol. 4, No. 1, pp. 1–7, January-March, 2013.
[10] E.W. Dempster, N.T. Tomov, M.H. Williams, H. Taylor, A Burger, P. Trinder and P. Broughton, "Modelling Parallel Oracle for Performance Prediction," Distributed and Parallel Databases, Vo. 13, No. 3, pp. 251–269, 2003.
[11] D. Dey, Z. Zhang and P. De, "Optimal Synchronization Policies for Data Warehouse," Information Journal on Computing, Vol. 18, No. 2, pp. 229-242, 2006.
[12] S. Elnaffar, P. Martin, B. Schiefer and S. Lightstone, "Is It DSS or OLTP: Automatically Identifying DBMS Workloads," Journal of Intelligent Information Systems, Vol. 30, No. 3, pp. 249–271, 2008.
[13] J. Evermann, "An Exploratory Study of Database Integration Processes," IEEE Transactions on Knowledge and Data Engineering, Vol. 20, No. 1, 2008.
[14] N. Rahman, "SQL Optimization in a Parallel Processing Database System," In proceedings of the IEEE 26th Canadian Conference of Electrical and Computer Engineering (CCECE 2013), Regina, Saskatchewan, Canada, May 5 - 8, 2013.
[15] A. Ferrández, A. Maté, J. Peral, J. Trujillo, E.D. Gregorio and M.-A. Aufaure, "A Framework for Enriching Data Warehouse Analysis with Question Answering systems," Journal of Intelligent Information Systems, Vol. 46, No. 1, pp. 61-82, 2016.
[16] S. Ganguly, W. Hasan and R. Krishnamurthy, "Query Optimization for Parallel Execution," In Proceedings of the 1992 ACM SIGMOD international conference on Management of data, San Diego, California, United States, pp. 9 – 18, 1992.
[17] García-García and C. Ordonez, "Extended Aggregations for Databases with Referential Integrity Issues," Data & Knowledge Engineering, Vol. 69, pp. 73–95, 2010.
[18] N. Rahman, "Temporal Data Update Methodologies for Data Warehousing," Journal of the Southern Association for Information Systems (JSAIS), Vol. 2, No. 1, pp. 25–41, 2014.
[19] T. Georgieva, "Discovering Branching and Fractional Dependencies in Databases," Data & Knowledge Engineering, Vol. 66, pp. 311-325, 2008.
[20] A. Ghazal, D. Seid, R. Bhashyam, A. Crolotte, M. Koppuravuri and G, Vinod, "Dynamic Plan Generation for Parameterized Queries," In Proceedings of SIGMOD’09, Providence, RI, USA.
[21] M. Golfarelli and S. Rizzi, "Data Warehouse Design: Modern Principles and Methodologies," McGraw-Hill Osborne Media; 1 edition, May 26, 2009.
[22] G. Hill and A. Ross, "Reducing Outer Joins," The VLDB Journal, Vol. 18, pp. 599–610, 2009.
[23] A. Karakasidis, P. Vassiliadis and E. Pitoura, "ETL Queues for Active Data Warehousing," In Proceedings of the 2nd International Workshop on Information Quality in Information Systems, IQIS 2005, Baltimore, MD, USA.
[24] N. Rahman and D. Rutz, "Building Data Warehouses Using Automation," International Journal of Intelligent Information Technologies (IJIIT), Vol. 11, No. 2, pp. 1–22, April-June, 2015.
[25] M.A. Kashem, A.S. Chowdhury, R. Deb and M. Jahan, "Query Optimization on Relational Databases for Supporting Top-k Query Processing Techniques," International Journal of Computer and Information Technology (IJCIT), Vol. 1, No. 1, pp. 53-58, 2010.
[26] A. Kerkad, L. Bellatreche, P. Richard, C. Ordonez and D. Geniet, "A Query Beehive Algorithm for Data Warehouse Buffer Management and Query Scheduling," International Journal of Data Warehousing and Mining, Vol. 10, No. 3, pp. 34–58, 2014.
[27] R. Kimball and M. Ross, "The Data Warehouse Toolkit: The Definitive Guide to Dimensional Modeling," 3rd edition, Wiley. 2013, Hoboken, New Jersey, USA.
[28] S. Krompass, H. Kuno, U. Dayal and A. Kemper, "Dynamic Workload Management for Very Large Data Warehouses – Juggling Feathers and Bowling Balls," In Proceedings of the 33rd international conference on Very large data bases, Vienna, Austria, 1105-1115, 2007.
[29] S. Krompass, H. Kuno, J.L. Wiener, K. Wilkinson, U. Dayal and A. Kemper, "Managing Long-Running Queries," In Proceedings of the 12th International Conference on Extending Database Technology: Advances in Database Technology, Saint Petersburg, Russia, pp. 132-143, 2009.
[30] D. Lee, S.K. Cha and A.H. Lee, "A Performance Anomaly Detection and Analysis Framework for DBMS Development," IEEE Transactions on Knowledge and Data Engineering, Vol. 24, No. 8, pp. 1345-1360, 2012.
[31] N. Rahman, D. Rutz, S. Akhter and F. Aldhaban, "Emerging Technologies in Business Intelligence and Advanced Analytics," ULAB Journal of Science and Engineering (JSE), Vol. 5, No. 1, pp. 7–17, November 2014.
[32] P. Leitner, J. Ferner, W. Hummer and S. Dustdar, "Data-driven and Automated Prediction of Service Level Agreement Violations in Service Compositions," Distrib Parallel Databases, Vol. 31, pp. 447-470, 2013.
[33] M. Mannino, S.N. Hong and I.J. Choi, "Efficiency Evaluation of Data Warehouse Operations," Decision Support Systems, Vol. 44, pp. 883-898, 2008.
[34] Y. Meng, P. Bird, P. Martin and W. Powley, "An Approach to Managing the Execution of Large SQL Queries," In proceedings of the 2007 Conference of the Center for Advanced Studies on Collaborative Research, Richmond Hill, Ontario, Canada (pp. 268-271).
[35] V. Narasayya, I. Menache, M. Singh, F. Li, M. Syamala and S. Chaudhuri, "Sharing Buffer Pool Memory in Multi-tenant Relational Data-

base-as-a-service," In proceedings of the VLDB Endowment, Vol. 8, No. 7, pp. 726–737, 2015.

[36] N. Rahman, "An Empirical Study of Data Warehouse Implementation Effectiveness," International Journal of Management Science and Engineering Management (IJMSEM), February 2016.

[37] T. Neumann, "Engineering High-Performance Database Engines," In proceedings of the VLDB Endowment, Vol. 7, No. 13, pp. 1734–1741, 2014.

[38] C. Ordonez and Z. Chen, "Horizontal Aggregations in SQL to Prepare Data Sets for Data Mining Analysis," IEEE Transactions on Knowledge and Data Engineering, Vol. 24, No. 4, pp. 678–691, 2012.

[39] R. Osman and W.J. Knottenbelt, "Database System Performance Evaluation Models: A survey, Performance Evaluation," Vol. 69, No. 10, pp. 471-493, 2012.

[40] W. Powley, P. Martin and P. Bird, "DBMS Workload Control Using Throttling: Experimental Insights," In proceedings of the 2008 conference of the center for advanced studies on collaborative research, Ontario, Canada.

[41] M.A. Sharaf and P.K. Chrysanthis, "Optimizing I/O-Intensive Transactions in Highly Interactive Applications," In Proceedings of the 35th SIGMOD international conference on Management of data, Providence, Rhode Island, USA, 785-798, 2009.

[42] F.D. Tria, E. Lefons, and F. Tangorra, "Benchmark for Approximate Query Answering Systems," Journal of Database Management, Vol. 26, No. 1, pp. 1–29, 2015.

[43] N. Rahman, N. Kumar and D. Rutz, "Managing Application Compatibility During ETL Tools and Environment Upgrades," Journal of Decision Systems (JDS), Vol. 25, No. 2, pp. 136-150. April-June 2016.

[44] D. VanderMeer, K. Dutta and A. Datta, "A Cost-based Database Request Distribution Technique for Online E-commerce Applications," MIS Quarterly, Vol. 36, No. 2, pp. 479-507, 2012.

[45] W. Weill, M. Subramani and M. Broadbent, "Building IT Infrastructure for Strategic Agility," MIT Sloan Management Review, 2002.

[46] B. Wixom and H. Watson, "The BI-Based Organization," International Journal of Business Intelligence Research, Vol. 1, No. 1, pp. 13-28, 2010.

[47] S. Wu, F. Li, S. Mehrotra and B.C. Ooi, "Query Optimization for Massively Parallel Data Processing," In proceedings of the 2nd ACM Symposium on Cloud Computing (SOCC '11), Cascais, Portugal, 2011.

[48] D.Y. Yoon, B. Mozafari and D. Brown, "DBSeer: Pain-Free Database Administration through Workload Intelligence," In proceedings of the VLDB Endowment, Vol. 8, No. 12, pp. 2036–2039, 2015.

[49] N. Rahman, "Enterprise Data Warehouse Governance Best Practices," International Journal of Knowledge-Based Organizations (IJKBO), Vol. 6, No. 2, pp. 21-37. April-June 2016.

Nayem Rahmanis a Senior Enterprise Application Developer in Cross Enterprise Systems IT, Intel Corporation. He has implemented several large projects using data warehousing and big data technologies for Intel's mission critical enterprise DSS platforms and solutions. He is currently working toward the Ph.D. degree in the Department of Engineering and Technology Management, Portland State University, USA.He holds an M.S. in Systems Science (Modeling & Simulation) from Portland State University, Oregon, USA and an MBA in Management Information Systems (MIS), Project Management, and Marketing from Wright State University, Ohio, USA. He has published more than 30 articles in peer-reviewed international jounrals and conference proceedings. He has four book chapters published. He has presented his creative work at many industry and academic conferences in USA and Canada.His principal research interests include Big Data Analytics, Big Data Technology Adoption, Data Mining for Business Intelligence, and Simulation-based Decision Support System (DSS).

Leonard Sutton is an Independent SQL Performance Tuning Consultant, with a focus on Massively Parallel Processing (MPP) systems. He is currently working at Boeing as an Independent Consultant.He is a certified Teradata Master. His work experience include 14 years at Nike Inc. (DB2 and Teradata), 11 years at Teradata Corporation as a Consultant helping many different Customers and 2 years with the Teradata Performance Tuning COE team. He has solid experience and knowledge about what DOES and DOES NOT work on Large Data Warehouse Systems using Parallel Processing Architecture.

Quality of ICT Facilities at the Tertiary Level Education in Bangladesh: Public versus Private University

H M Jahirul Haque and Faisal Mohammad Shahriar
University of Liberal Arts Bangladesh, Dhaka, Bangladesh
Email: jahirul.haque@ulab.edu.bd and faisal.shahriar@ulab.edu.bd

Abstract

The private and public universities of Bangladesh have taken significant initiatives in developing ICT facilities to ensure access to the state-of-the-art learning environment for their students. However, the impact of such initiatives is largely subjected to the studies on the perception of the very students using these facilities. The study focused on the students' perception about the quality of the ICT facilities available in public and private universities of Bangladesh in a comparative manner. Data for the study was collected from 265 students from 6 private and 3 public universities through a survey conducted on the perceived quality of ICT facilities in their respective institutions as expressed by 15 key indicators. Both descriptive and inferential analyses were used to test the hypotheses of the study. The study was conducted on a small scale and may be considered as a primer for future investigations. The outcomes of the study show that, perceived quality of the ICT facilities are mentionably different among the private and public unversitystudents, where private universities are performing better in 9 out of the 15 key indicators.

Keywords— ICT facilities, Private university, Public university, Quality.

© University of Liberal Arts Bangladesh
All rights reserved.

Manuscript received on 17 November 2016 and accepted for publication on 27 November 2016.

1 Introduction

QUALITY,as a widely used term to express the performance of any system and process is largely defined by its technical and functional dimensions [1], [2]. Quality in education is an issue that has been signified and studied for a long time. A good number of studies have been conduced to investigate the dimensions of quality in higher education after the inception of the digital era in the later part of the 20th century, a mentionable few of which may bethe works of Izquierdo[3], Madu & Kuei [4], Yorke [5], Hartman & Schmidt [6], and Ermer [7]. These studiescommonly identified the availability and usefulness of or ICTtoolsas significant determinants of quality in tertiary level education. Over the last two decades, ICT has become an indispensible component of education especially, at the tertiary level as a means of effective teaching and learning [8], [9], [10] and also adapting to globalization and societal requirements [11], [12]. As a result, ICT has significantly reshaped higher education through offering powerful learning environments and enhancing capabilities and skills of the students for cooperation, communication, problem solving and lifelong learning[13], [14], [15], [16], [17], [18], [19], [20], [21],[22].However, the effectiveness of ICT is subjected to the availability and maintenance of the equipmentsand accessories [23]. Successful implementation of ICT therefore, requires strong intervention and support from institutions [24]. In recognition to the significance of ICT for quality education, the private and public universities of Bangladesh have taken strong initiatives to develop ICT facilities especially, fortheir students[25],although the private universities claim to have performed better in this regard[26]. Better infrastructure of ICT and wider accessto ICT facilities have also been found to significantly contribute to the appeal of private universities among the prospective students in a good number of studies, for example Sabur [27], Lamagna [28], lamagna [29], Aminuzzaman [30]etc. This study is an initiative to investigate the quality ofICT facilities the private and public universities of Bangladesh from a students' point of view and empirically comparebetween performances of these two types of institutions in this regard. The outcomes of the study is based on the survey datafrom the students of6 private and 3 public universities regarding 15 key indicators namely, Number of available computers(Num_PC), Printing facilities(Print_Fac), Condition of ICT equipments(Equip_Con), Power back-up facilities(PB_Fac), Availability of necessary software(Soft_Avail), User-friendliness of provided software(Soft_Use), Frequency of software updates(Soft_Up), Effectiveness of antivirus software(Eff_AV), Sufficiency of internet facilities(Int_Avail), Speed of internet(Int_Sp), Data sharing facilities(Data_Sh), Availability of ICT support staffs(Staff_Avail), Service quality of ICT support staffs(Staff_Qual), Timely solution of ICT related problems(Time_Sol) and Overall quality of ICT facilities(Over_Qual)that are derived from aprevious study by Haque and Khan [31].

2 Objectives of the Study

The primary objective of the study is to investigate the perceived quality of ICT facilities among students of the private and public universities of Bangladesh. The other objective of the study is to determinle if the private universities are performing better than the public universities in terms of the quality of ICT facilities for their students and if yes, then in which key indicator(s). The objectives of the study are addressed through analyzing survey data with quantitative techniques.

3 Research Questions

In accordance to the objectives, the research questions of the study are as follows:

a) What is the present status of the key indicators of quality of the ICT facilities in the the private and public universities as perceived by the students?
b) Are the private universities perfoming better in terms of quality of ICT facilities than that of thepublic universities and if yes, then in which key indicator(s)?

4 Research Hypotheses

The research question (a) would be satisfied with descriptive statistics alone and thus, would not require any hypothesis. For the research question (b), the null hypothesis (H_0) is that, there is no difference between the value of population mean of the key indicators of quality of ICT facilities in the private and the same in the public universities (H_0: $\mu_{i\text{-Private}} = \mu_{i\text{-Public}}$, where '*i*' denotes the key indicators). As per the objective of the study, the alternative hypothesis (H_1) is that, the value population mean of the key indicators in the private universitiesare significantly higher than the same in the public universities (H_1: $\mu_{i\text{-Private}} > \mu_{i\text{-Public}}$).

5 Methodology

Snowball sampling technique has been adopted to build the two independent samples of the study, one from the private universities and the other from the public universities. The sample size for the private universities is 163, which includes students from 6 institutions namely Eastern University (EAU), International University of Business Agriculture and Technology (IUBAT), North-South University (NSU), Stamford University (STU), State University of Bangladesh (SUB), and University of Liberal Arts Bangladesh (ULAB). The sample size for the public universities on the other hand is 102, which includes students from 3 institutions namely, Bangladesh University of Engineering and Technology (BUET), Jagannath University (JNU), and University of Dhaka (DU). At present, there are 37 public and 91 private universities in Bangladesh. Therefore, to ensure equitability of data, the number of private and public universities taken under the study has been disproportionate. Data from the respondents were collected through questionnaire survey, a method which has been proven effective in similar studies [32], [33], [9]. The questionnaire recorded and measuredthe students' responses (ranging from strongly disagree to strongly agree) on a 5 point itemized rating scale namely, the Likertscaleto statements regarding to the 15 key indicators of the quality of ICT facilities of their respective institutions. Likert scale is found effective in similar studies on students [34], [35],[36],[37]. Personal interviewing technique has been followed during the survey.The timeline of the survey has been January to April 2016. XLSTAT, a computerized statistical program has been used to analyze the survey data. A part of the analyses is descriptive, which has been suggested for this type of studies [38]. Parametric tests, namely Independent samples Z-test (upper-tailed) and Independent samples t-test (upper-tailed) have been conducted to test the research hypothesis. Although the data are technically ordinal, Grace-Martin [39], Lubke & Muthen [40] and Glass *et* al. [41] have implied that data from Likert scale with at least 5 categories can be used for parametric tests in some situations where the differences between the ordinal categories are considered equal. Norman [42], Labovitz [43], and Kim [44] also implied that, data from Likert scale is significantly capable of yielding accurate outcomes as like as continuous data.

6 Descriptive Statistics

The the first sample of the study respondents of the study came from both public and private universities. The distribution of respondents in terms of universities and their types are shown in table 1:

TABLE 1
Institution Wise & Category Wise Distribution of the Respondents

Serial	Institution Wise Distribution			Category Wise Distribution		
	Name	Frequency	Percentage	Category	Frequency	Percentage
1.	EAU	36	13.6	Private	163	61.5
2.	IUBAT	01	0.4			
3.	NSU	36	13.6			
4.	STU	34	12.8			
5.	SUB	26	9.8			
6.	ULAB	30	11.3			
7.	BUET	32	12.1	Public	102	38.5
8.	DU	35	13.2			

9.	JNU	35	13.2			
-	**Total**	**265**	**100.0**	**Total**	**265**	**100.0**

The responses of the students regarding to the key indicators of the ICT facilities (or variables) according to the type of institution are shown in table 2 and a graphical projection of the mean scores of the responses from the two samples is given in the figure 1followingly:

TABLE 2
SUMMARY OF THE RESPONSES

Serial	Variables / Type of Institution	Mean (Private Univ.)	Mean (Public Univ.)	Difference (point)	Difference (percent)
1.	Number of PC (Num_PC)	3.45	3.25	0.20	4.0
2.	**Printing facilities (Print_Fac)**	**3.32**	**2.45**	**0.87**	**17.4**
3.	**Condition of ICT equipments (Equip_Con)**	**3.49**	**2.91**	**0.58**	**11.6**
4.	**Power back-up facilities (PB_Fac)**	**3.71**	**2.63**	**1.08**	**21.6**
5.	Availability of necessary software (Soft_Avail)	3.33	2.89	0.44	8.8
6.	**User-friendliness of provided software (Use_Soft)**	**3.61**	**2.95**	**0.66**	**13.2**
7.	**Frequency of software updates (Soft_Up)**	**3.29**	**2.51**	**0.78**	**15.6**
8.	**Effectiveness of antivirus software (Eff_AV)**	**3.17**	**2.48**	**0.69**	**13.8**
9.	Sufficiency of internet facility (Int_Avail)	3.56	3.28	0.28	5.6
10.	Speed of Internet (Int_Sp)	3.19	3.11	0.08	1.6
11.	Data sharing facilities (Data_Sh)	3.19	3.11	0.08	1.6
12.	Availability of ICT support staffs (Staff_Avail)	3.50	3.29	0.21	4.2
13.	Service quality of ICT support staffs (Staff_Qual)	3.55	3.45	0.10	2.0
14.	Timely solution of ICT related problem (Time_Sol)	3.52	3.32	0.20	4.0
15.	Overall quality of ICT facilities (Over_Qual)	3.58	3.19	0.39	7.8

Note: 1) The number of response for private and public university is 163 and 102 respectively
2) No missing value were found
3) The minimum and maximum response score is 1 and 5 respectively

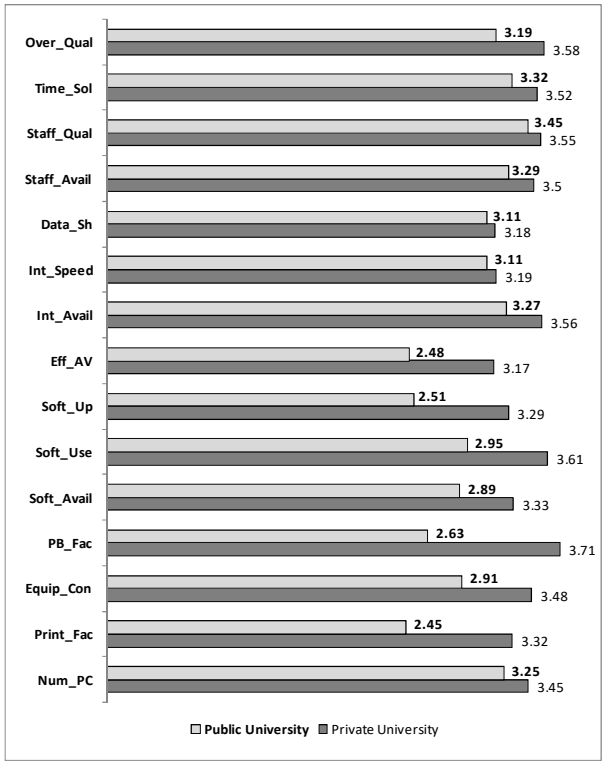


Figure 1: Mean score of responses on the key indicators of the quality of ICT Facilities from private and public universities.

From the descriptive data, it can be seen that the mean scores of responses from the students of private universities are higher than that of the public universitiesfor all the 15 key indicators. However, the contrast is significant for variables like facilities available at the universities namely, Printing facilities (Print_Fac), Condition of ICT equipments (Equip_Con), Power back-up facilities (PB_Fac), User-friendliness of provided software (Sof_Use), Frequency of software updates (Soft_Up), and Effectiveness of antivirus software (Eff_AV) where the scores of private universities exceed the same from public universities on a rangeof 0.58 to 1.08 on a 5 point scale (11.6 to 21.6 percent).

7 Parametric Test

7.1 Independent Samples Z-test

An 'Independent-Samples' Z-test (upper-tailed) has been conducted to test the null hypotheses (H_0) that the value of the population mean of the key indicators of the quality of ICT facilities in the private universities are equal to the samein the public universities with a precision level (α) of 0.05. The alternative hypothesis (H_1) was that the value of the population mean is higher for the private universities than the same in the public universities. The outcomes of the test are shown in the Table 3 below:

TABLE 3

Independent Samples Z-test (Upper-Tailed)

Variables	Difference between means	z (Observed)	|z| (Critical)	*p*-value (upper-tailed)	H_0
Num_PC	**0.203**	**1.070**	**1.645**	**0.142**	**Accepted**
Print_Fac	0.868	5.120	1.645	< 0.0001	Rejected
Equip_Con	0.573	3.441	1.645	0.000	Rejected
PB_Fac	1.078	6.266	1.645	< 0.0001	Rejected
Soft_Avail	0.439	2.674	1.645	0.004	Rejected
Use_Soft	0.656	3.705	1.645	0.000	Rejected
Soft_Up	0.779	4.649	1.645	< 0.0001	Rejected
Eff_AV	0.685	3.996	1.645	< 0.0001	Rejected
Int_Avail	0.284	1.659	1.645	0.049	Rejected
Int_Sp	**0.082**	**0.472**	**1.645**	**0.318**	**Accepted**
Data_Sh	**0.076**	**0.454**	**1.645**	**0.325**	**Accepted**
Staff_Avail	**0.203**	**1.188**	**1.645**	**0.117**	**Accepted**
Staff_Qual	**0.095**	**0.576**	**1.645**	**0.282**	**Accepted**
Time_Sol	**0.192**	**1.213**	**1.645**	**0.113**	**Accepted**
Over_Qual	0.390	2.218	1.645	0.013	Rejected

Note: 1) $H_0 : \mu_{i\text{-Private}} = \mu_{i\text{-Public}}$, $H_1 : \mu_{i\text{-Private}} > \mu_{i\text{-Public}}$

The outcomes of the Z-test show that, the *p*-values for the 6 out of 15 key indicators namely, number of available computers (Num_PC), speed of internet (Int_Sp), data sharing facilities (Data_Sh), availability of ICT support staffs (Staff_Avail), services quality of ICT support staffs (Staff_Qual) and timeliness of solutions of ICT related problems (Time_Sol) are not significant. Therefore, the null hypothesis is accepted for these key indicators, implying that there is no difference betweenthe values of population mean for these key indicators in the private and public universities. On the other hand,the *p*-values for printing facilities (Print_Fac), condition of ICT equipments (Equip_Con), power backup facilities (PB_Fac), availability of necessary software (Soft_Avail), user-friendliness of provided software (Soft_Use), frequency of software updates (Soft_Up), effectiveness of antivirus software (Eff_AV), sufficiency of internet facilities (Int_Avail) and overall quality of ICT facilities (Over_Qual) are significant. Therefore, the null hypotheses for these key indicators are rejected, implying that the values of the population mean for these key indicators are higher in the private universities than the same in the public universities.

7.2 Independent Samples *t*-test

An 'Independent-Samples'*t*-test (upper-tailed) hasalso been conducted to test the null hypotheses (H_0) that the mean scores of the 15 key indicators of the quality of ICT facilities of the private universities are equal to that of the public universities with a precision level (α) of 0.05. The alternative hypothesis (H_1) was that the mean scores are higher for the private universities than that of the public universities. The t-test was followed by a 'Fisher's F-test' for equality of the variances (two-tailed) to approximate the degrees of freedom (*df*) for the *t*-test. The outcome of the F-test isshown in the Table 4 below:

TABLE 4

FISHER'S F-TEST FOR EQUALITY OF VARIANCES (TWO-TAILED)

Variables	Ratio	F (Observed)	|F| (Critical)	*p*-value (one-tailed)	H_0	Equal variances assumed
Num_PC	**0.720**	**0.720**	**1.434**	**0.063**	**Accepted**	**Yes**
Print_Fac	**0.801**	**0.801**	**1.434**	**0.208**	**Accepted**	**Yes**
Equip_Con	0.681	0.681	1.434	0.029	Rejected	No
PB_Fac	**0.725**	**0.725**	**1.434**	**0.068**	**Accepted**	**Yes**
Soft_Avail	**0.898**	**0.898**	**1.434**	**0.538**	**Accepted**	**Yes**
Use_Soft	**0.742**	**0.742**	**1.434**	**0.091**	**Accepted**	**Yes**
Soft_Up	**0.946**	**0.946**	**1.434**	**0.747**	**Accepted**	**Yes**
Eff_AV	**1.100**	**1.100**	**1.434**	**0.608**	**Accepted**	**Yes**
Int_Avail	**0.783**	**0.783**	**1.434**	**0.166**	**Accepted**	**Yes**
Int_Sp	**0.908**	**0.908**	**1.434**	**0.582**	**Accepted**	**Yes**
Data_Sh	**1.098**	**1.098**	**1.434**	**0.615**	**Accepted**	**Yes**
Staff_Avail	**0.763**	**0.763**	**1.434**	**0.126**	**Accepted**	**Yes**
Staff_Qual	0.671	0.671	1.434	0.024	Rejected	No
Time_Sol	**0.847**	**0.847**	**1.434**	**0.346**	**Accepted**	**Yes**
Over_Qual	0.587	0.587	1.434	0.002	Rejected	No

Note: H_0: Ratio between the variances = 1; df1 = 162, df2 = 101; $\alpha = 0.05$

The outcomeof the F-test shows that equality of variances between the two samples of the study can be assumed for all except 3 of the 15 key indicators. The 3 key indicators assumed to have unequal variances are condition of ICT equipments (Equip_Con), services quality of ICT support staffs (Staff_Qual) andoverall quality of ICT facilities are significant (Over_Qual). The degrees of freedom (*df*) for these 3 indicators are 185, 183 and 174 respectively. The degrees of freedom (*df*) for all other variables are 263. The outcome of the *t*-test isshown in the Table 5 below:

TABLE 5
INDEPENDENT SAMPLES T-TEST (UPPER-TAILED)

Variables	*df*	Difference between means	t (Observed)	|t| (Critical)	*p*-value (upper-tailed)	H_0
Num_PC	**263**	**0.203**	**1.111**	**1.651**	**0.134**	**Accepted**
Print_Fac	263	0.868	5.252	1.651	< 0.0001	Rejected
Equip_Con	185	0.573	3.441	1.653	0.000	Rejected
PB_Fac	263	1.078	6.501	1.651	< 0.0001	Rejected
Soft_Avail	263	0.439	2.708	1.651	0.004	Rejected
Use_Soft	263	0.656	3.834	1.651	< 0.0001	Rejected
Soft_Up	263	0.779	4.678	1.651	< 0.0001	Rejected
Eff_AV	263	0.685	3.952	1.651	< 0.0001	Rejected
Int_Avail	263	0.284	1.706	1.651	0.045	Rejected
Int_Sp	**263**	**0.082**	**0.477**	**1.651**	**0.317**	**Accepted**
Data_Sh	**263**	**0.076**	**0.449**	**1.651**	**0.327**	**Accepted**
Staff_Avail	**263**	**0.203**	**1.226**	**1.651**	**0.111**	**Accepted**
Staff_Qual	**183**	**0.095**	**0.576**	**1.651**	**0.283**	**Accepted**
Time_Sol	**263**	**0.192**	**1.237**	**1.651**	**0.109**	**Accepted**
Over_Qual	174	0.390	2.218	1.651	0.014	Rejected

Note: 1) $H_0 : \mu_{i\text{-Private}} = \mu_{i\text{-Public}}$, $H_1 : \mu_{i\text{-Private}} > \mu_{i\text{-Public}}$

2) The degrees of freedom (*df*) for the test is approximated by the Welch-Satterthwaite formula as per the outcomes of the Fisher's F-test

The outcome of the *t*-test is completely similar to that of the Z-test described above. The *t*-test has accepted and rejected the hypotheses for the same variables as seen in the Z-test. Therefore, the interpretations of the Z-test are found to be the same in this

case. Thus it can be strongly inferred that, the private universities are showing better performance is ensuring quality for 9 of the 15 key indicators.

8 Implications and Conclusion

The study, although conducted on a limited scale, project that the perceived quality of the ICT facilities from the students' point of view in the private universities are higher than the same in the public universities on a number of key indicators. Nevertheless, the perceived quality on the the other key indicators are not significantly different for the private and public universities. Therefore, it is not conclusive that the private universities are absolutely out-performing the public universities in terms ofquality of the ICT facilities. Moreover, there are still significant scopes for both the private and public universities regarding the quality of ICT facilities even though the private universities may scorehigherin this regard. It is indeed an issue to consider that none of the variables in the study scored a mean value of 4 or higher in a 5-point scale, which shows a rather poor perception among the students about the capability and performance of the universities in providing state-of-the-art ICT facilities. One reason for this low level of perception might be that the students mostly use ICT facilities (i.e. computers, internet) at their own while they are at the university campus [45]. Therefore, they might not have definite attitudes towards the ICT facilities provided by their institutions.The outcome of the study is tentative in consideration of depth and rigor,and should be considered as a primer for further studies on this issue. It is therefore suggested that studies on larger sample with pre-screened respondents should be conducted to draw a true picture of the ICT scenario from the students' point of view.

References

[1] C. Gönroos, *Service Management and Marketing*, Massachusets: Lexington Books, 1990

[2] U. Lehtinen, and J.R. Lehtinen, "Two approaches to service quality dimensions", *The Service Industries Journal*, vol. 11, no. 3, pp. 287-303, 1991

[3] F.A. Izquierdo, "Quality-designed curricula," *European Journal of Engineering Education*, vol. 18, no. 4, pp. 339-344, 1993

[4] C.N. Madu, and C. Kuei, "Dimensions of Quality Teaching in Higher Education", *Total Quality Management*, vol. 4, no. 3, pp. 325-338, 1993

[5] M. Yorke, "Self-scrutiny of quality in higher education: a questionnaire," *Quality Assurance in Education*, vol. 3, no. 1, pp. 10-13, 1995

[6] D.E. Hartman, and S.L. Schmidt, "Understanding student/alumni satisfaction from a consumer's perspective," *Research in Higher Education*, Vol. 36 No. 2, pp. 197-217, 1995

[7] D.S. Ermer, "Using QFD Becomes an Educational Experience for Students and Faculty", *Quality Progress*, vol. 28, no.5, pp. 131-136, 1995

[8] M.A. Hossain, M.A. Salam, and F. Shilpi,"Readiness and Challenges of Using Information and Communications Technology (ICT) in Higher Education of Bangladesh",*The Online Journal of New Horizons in Education*, vol. 6, no. 1, pp. 123-132, 2016

[9] A.K.Bairagi, S.A.A. Rajon, and T. Roy,"Status and Role of ICT in Educational Institution to Build Digital Society in Bangladesh: Perspective of a Divisional City, Khulna",*International Journal of Advances in Engineering & Technology*, vol. 1, no. 4, pp. 374-383, 2011

[10] M. Dunmill, and A. Arslanagic, "ICT in Arts Education A Literature Review", *Report: the Te Puna Puoru National Centre for Research in Music Education and Sound Arts, University of Canterbury, New Zealand*, July 2006.
available: http://artsonline2.tki.org.nz/documents/ICTinArtsEducationALiteratureReview.pdf (accessed: 31 May 2016)

[11] M.S. Islam, and M.N. Islam, "Use of ICT in Libraries: An Empirical Study of Selected Libraries in Bangladesh", *Library Philosophy and Practice*(e-journal), paper 143, Aug. 2007, available: http://www.webpages.uidaho.edu/~mbolin/shariful.pdf (accessed: 27 May 2016)

[12] C. Blurton, "New Directions of ICT-Use in Education", *World Communication and Information Report, UNESCO*, 1999

[13] J. Pulkkinen, "Cultural globalization and integration of ICT in education"in K. Kumpulainen (Ed.), *Educational technology: Opportunities and challenges,* pp. 13–23, 2007, Oulu, Finland: University of Oulu

[14] D. Wood, "Theory, training, and technology: Part I.", *Education and Training*, vol. 37, no. 1, pp. 12–16, 1995

[15] M.Volman, and E.V. Eck, "Gender Equity and Information Technology in Education: The Second Decade," *Review of Educational Research,* vol. 71, no. 4, pp. 613–634, 2001

[16] T.Plomp, W.J. Pelgrum, and N. Law, "SITES2006—International comparative survey of pedagogical practices and ICT in education,"*Education and Information Technologies*, vol. 12, no. 2, pp. 83-92, 2007

[17] S. Mumtaz, "Factors Affecting Teachers' Use of Information and Communications Technology: A review of the Literature,"*Journal of Information Technology for Teacher Education*, vol. 9, no. 3, pp. 319-342, 2000

[18] C.R. Casal, "ICT for Education and Development,"*info*, vol. 9, no. 4, pp. 3-9, 2007

[19] R. Kozma, "National Policies that Connect ICT-Based Education Reform to Economic and Social Development,"*Human Technology*,vol. 1, no. 2, pp. 117-156, 2005

[20] I. Bhattacharya, and K. Sharma, "India in the Knowledge Economy – An Electronic Paradigm,"*International Journal of Educational Management*, vol. 21, no. 6, pp. 543-568, 2007

[21] K.Pajo, and C. Wallace, "Barriers to the uptake of web-based technology by university teachers",*Journal of Distance Education*, vol. 16, pp. 70–84, 2001

[22] J. Voogt, "Consequences of ICT for Aims, Contents, Processes and Environments of Learning"in J. van den Akker, W. Kuiper, & U. Hameyer (Eds.), *Curriculum landscapes and trends*, pp. 217–236, Dordrecht: Kluwer, 2003

[23] M.S.H. Khan, M. Hasan, and C.K. Clement, "Barriers to the Introduction of ICT into Education in Developing Countries: The Example of Bangladesh,"*International Journal of Instruction*, vol. 5, no. 2, pp. 61-80, 2012

[24] M.Cross, and F. Adam, "ICT Policies and Strategies in Higher Education in South Africa: National and Institutional Pathways,"*Higher Education Policy*, vol. 20, no. 1, pp. 73-95, 2007

[25] M. Roknuzzaman, "A Survey of Internet Access in a Large Public University in Bangladesh," *International Journal of Education and Development using ICT (e-journal)*, vol. 3, no. 2, 2006, available: http://ijedict.dec.uwi.edu/viewarticle.php?id=195&layout=html (accessed: 27 May 2016)

[26] M.R. Jewel, "An Empirical Study on the challenges and prospects of Private Universities in Higher Education in Bangladesh, "*Proc.International Conference on Tertiary Education (ICTERC 2013), Daffodil International University, Dhaka, Bangladesh,* pp. 117-123, Jan. 2013, available: http://dspace.library.daffodilvarsity.edu.bd:8080/bitstream/handle/123456789/879/AnEmpiricalStudyontheChallengesandProspectsofPrivateUniversitiesinHigherEducationinBangladesh.pdf?sequence=1 (accessed: 17 July 2016)

[27] M. A. Sabur, "Dhaka University verses private university: A comparative analysis of quality of education offered by the institutions", BSS dissertation, University of Dhaka, Bangladesh, 2004

[28] C.Z. Lamanga, *Strategic view of the development of Higher Education: Bangladesh AIUB perspective, Bangladesh Business Year Book*, Dhaka: A1 Publication, 2002

[29] C.Z. Lamanga, "Quality Assurance in Tertiary Education: Bangladesh Experience", *Seminar paper:The World Bank Learning Seminar, CIEP, France*, Jun. 2006

[30] S. Aminuzzaman, "Overview of quality assurance in the context of Bangladesh," *Workshop paper: American International University Bangladesh, Dhaka, Bangladesh*, 2007

[31] H.M.J. Haque, and M.H. Khan, "Quality of ICT Services at Higher Educational Institutions of Bangladesh,"*ULAB Journal of Science and Engineering*, vol. 1, no. 1, pp. 42-47, 2010

[32] M.S. Islam, and M.H. Fouji, "The Impact of ICT on Students' Performance: A Case Study of ASA University Bangladesh,"*ASA University Review*, vol. 4, no. 2, pp. 101-106, 2010

[33] N. Abdelaziz, J. Shahrir, and C.H. Leng, "Measuring Attitude Toward Computer and Internet Usage among Postgraduate Students in Malaysia," *The Turkish Online Journal of Educational Technology*, vol. 12, no. 2, pp. 200-216, 2013

[34] J. Dorup, "Experience and attitudes towards information technology among firstyear medical students in Denmark: Longitudinal questionnaire survey,"*Journal of Medical Internet Research,* vol. 6, no. 1:e-10, 2004, DOI: 10.2196/jmir.6.1.e10. Available: http://www.jmir.org/2004/1/e10/ (accessed: 31 May 2016)

[35] G. Gay, S. Mahon, D. Devonish,P.A. Alleyne, and P.G. Alleyne, "Perceptions of information and communication technology among undergraduate management students in Barbados",*International Journal of Education and Development using Information and Communication Technology,* vol. 2, no. 4, pp. 6-17, 2006

[36] J. Peeraer, and P.V. Petegem, "Factors Influencing Integration of ICT in Higher Education in Vietnam", *Proc. Global Learn 2010, Association for the Advancement of Computing in Education (AACE), Penang: Malaysia,* pp. 916-924, May 2010.
Available: http://www.vvob.be/vietnam/files/SubmissionGlobalLearnJP_v2.pdf (accessed: 28 May 2016)

[37] M. Drent, and M. Meelissen,"Which factors obstruct or stimulate teacher educators to use ICT innovatively?", *Computers & Education*,vol. 51, no. 1, pp. 187-199, 2008

[38] J. Cohen, "Multiple Regression as a General Data-Analytic System",*Psychological Bulletin*, vol. 70, no. 6, pp. 426-443, 1968

[39] K. Grace-Martin, "Can Likert Scale Data ever be Continuous?", available: http://www.theanalysisfactor.com/can-likert-scale-data-ever-be-continuous/, 2008. (accessed: 01 June 2016)

[40] G.H. Lubke, and B.O. Muthen, "Applying Multigroup Confirmatory Factor Models for Continuous Outcomes to Likert Scale Data Complicates Meaningful Group Comparisons", *Structural Equation Modeling,*vol. 11, pp. 514-534, 2004

[41] G.V.Glass, P.D. Peckham, and J.R. Sanders,"Consequences of failure to meet assumptions underlying the analyses of variance and covariance",*Review of Educational Research,* vol. 42, pp. 237-288, 1972

[42] G. Norman, "Likert Scales, Levels of Measurement and the "laws" of statistics",*Advances in health sciences education*, vol. 15, no. 5, pp. 625-632, 2010

[43] S. Labovitz,"The assignment of numbers to rank order categories",*American Sociological Review*, vol. 35, pp. 515-524, 1970

[44] Jae-On Kim, "Multivariate analysis of ordinal variables revisited",*American Journal of Sociology*, vol. 84, pp. 448-456, 1978

[45] G. Rabbani, and S. Chowdhury, "Quality of Higher Education in Bangladesh: Governance Framework and Quality Issues,"*Beykent University Journal of Social Sciences*, vol. 7, no. 1, pp. 78-91, 2014

APPENDIX

Survey Questionnaire

Quality of ICT Facilities at Tertiary Level Education of Bangladesh

Dear Respondent, Thank you very much for participating in the survey. The purpose of the study is purely academic. Your information and identity shall not be disclosed to anyone else.

Instruction: Please put Tick mark on the correct option. Write down appropriate data where applicable.

---

1. Name of your university:

2. Your study program and current study level (year/semester):

3. The number of computers available in the for the use of the students is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

4. Printing facilities available for students in the university is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

5. The condition of the ICT equipments (i.e. computers, printers, prjectors) in the university are satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

6. Power backup support (i.e. UPS) for computers in the university is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

7. Necessary softwares are widely available at the university for students' usage:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

8. Software provided by the univeristy for students' usage is user friendly and easy:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

9. Software provided by the univeristy are up-to-date and are regularly updated:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

10. Antivirus software provided by the univeristy is effective:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

11. Internet facilities for students at the university is sufficient:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

12. The speed of internet at the university is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

13. Data sharing facilities through internal network (i.e. LAN) at the university is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

14. Supporting staffs for IT related issues are available whenever necessary:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

15. The service quality of the IT staff(s) is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

16. IT related services and solutions are timely provided to students at the university:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

17. The overall quality of ICT realted services at the university is satisfactory:

Strongly Agree	*Agree*	*Neither Agree nor Disagree*	*Disagree*	*Strongly Disagree*

18. Please provide us your email address and cell phone number in case we need to contact you for further queries:

---------Thank You--------

H M Jahirul Haque received his PhD in computer engineering from Kharkov National University of Radio Electronics, Ukraine in July 2001 and Masters in Technical Sciences (Systems Engineering) from Cherkassy State Technological University, Ukraine in 1995. After serving the Leading University, Sylhet, Bangladesh as the Head of the Department of Computer Science and Engineering, he joined the University of Liberal Arts Bangladesh (ULAB) in 2004. Now he is a Professor in the Department of Computer Science and Engineering and holds the position of Dean of the School of Science and Engineering at the university. Additionally, he is the Pro Vice Chancellor of ULAB. Dr. Haque is currently devoted in education related research as he is dealing with academic and faculty development related issues at ULAB. He has worked with different reputed universities of UK, USA, Australia, Japan, Ukrain and Russia in different capacities.

Faisal Mohammad Shahriar received his BBA and MBA in Marketing from the University of Dhaka, Bangladesh. After serving the Leading Univeristy in Sylhet, Bangladesh and the State University of Bangladesh in Dhaka, Bangladesh as a faculty member, he is currently an Assistant Professor in the School of Business of the University of Liberal Arts Bangladesh (ULAB). He is also a PhD research fellow in the University of Dhaka. Mr. Shahriar has been pursuing research studies on the areas of tourism, social media & ICT and consumer behavior.