I.J.Modern Education and Computer Science, **2013, 2, 19-24**
Published Online February 2013 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmecs.2013.02.03

Designing a Universal Data-Oriented Random Number Generator

Rasoul Farjami Nezhad
Department of Computer Science, Islamic Azad University Tabriz Branch, Tabriz, Iran
Email: r.f.nejad@gmail.com

Mehdi Effatparvar
Department of Computer Science, Islamic Azad University Ardebil Branch, Ardebil, Iran
Email: effatparvar@iauardabil.ac.ir

Mohammad Rahimzadeh
Department of Computer Science, Islamic Azad University Germi Branch, Ardebil, Germi, Iran
Email: mgs_mohammad@engineer.com

Abstract — Data-oriented is new and applied theory which provides method that models the concepts with data structure. If the concept is modeled by using sufficient data in modeling, required inferences and calculations can be done fast with less complexity. Random variable was modeled with digital probability graph, by using Ahmad Fact and probability density function. Some data-oriented random generators have been presented based on data-oriented approach. In this paper a universal data-oriented random number generator is introduced which is able to generate random numbers with uniform, normal, exponential and chi-square distributions.

Index Terms— Probability Tree; Random variable; Data-oriented; Ahmad fact; Digit bank

I. INTRODUCTION

Random numbers are useful for a variety of purposes such as generating data encryption keys, simulation and modeling uncertain phenomena and for selecting samples from lager data set. Random numbers are modeled by using mathematical function, probability density function or probability mass function. For example the chi-square distribution with degrees of freedom is the distribution of a sum of the squares of independent standard random variables. It is one of the most widely used probability distributions in inferential statistics, e.g. in hypothesis testing, or in construction of confidence intervals.

So far, due to lack of storage memory and it's low-speed , the problems were solved with much less data lying complex algorithms. nowadays, growing memory technology causes to computing system to manage large amount of data very fast and easily. Data-oriented theory presents methods in which any concept can be modeled in terms of data structures. Thus in modern computing systems Concepts that are modeled with data-oriented theory can be recognized and processed more quickly. In theses models the questions can be answered by data processing or by fewer amounts of mathematical operations. The methods to answer the questions by using large amount of data, are called data–oriented methods [1].

This paper introduced a Universal Random Number Generator, URNG, with uniform, normal, exponential and chi-square distributions based on Data-Oriented theory. This model is compatible with modern computer's structure and it is able to generate random numbers with any distribution and utilize computers in statistical inference and probability with higher speed and productivity. The rest of this paper is organized as follow:

The related work of Data-Oriented modeling and basic definitions to outline URNG, presented in section two. The URNG is designed in section three, section four provides concluding remarks and future works.

II. BASIC CONCEPT

The basic structure of data-oriented modeling has been presented to handle structures like Probability Digraph, probability language, complete tree walk and n-complete tree walks [2]. In other words, requirement tools, definition and important mathematical theorems for these models have been presented in [3]. Recently some methods have been presented to model and generate numbers with a given distribution based on data-oriented approach. For example data–oriented models of uniform [4], normal [5], exponential [6] and chi Square random variables [1] have been presented. Each of these models generates random numbers with any given distribution separately, but URNG is suggested, This URNG have been able to generate random numbers with uniform, normal, exponential, and chi-square distributions.

Probability digraph, digital prodigraph, value of walk[1], value of leaf[2], have been defined in [1]. In this paper we use them and some other definitions are provided as follows:

Definition:

The probability density function of the continuous uniform distribution is defined by equation 1 [4]:

$$f(x) = \begin{cases} \frac{1}{b-a} \\ 0 \end{cases} \quad , a<x<b \qquad (1)$$

Otherwis0e.

The graph of the uniform distribution for a =2 and b=7 is shown in Fig. 1.

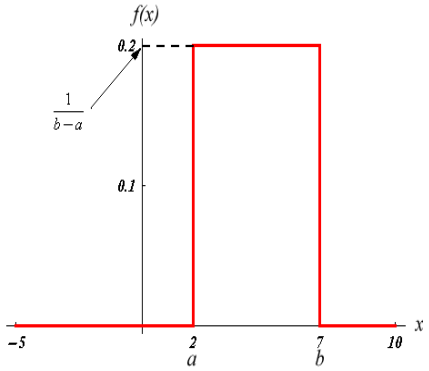

Fig. 1:The graph of the P.D.F. of uniform distribution [11].

Definition:

Let x be of a normal random variable with a mean of μ and variance δ^2, Then the probability density function is reached through equation 2 [5]:

$$\varphi(z) = n(x,\mu,\sigma) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \qquad (2)$$

The normal distribution is probably the most frequently used distribution. The graph of normal distribution is shown in Fig. 2.

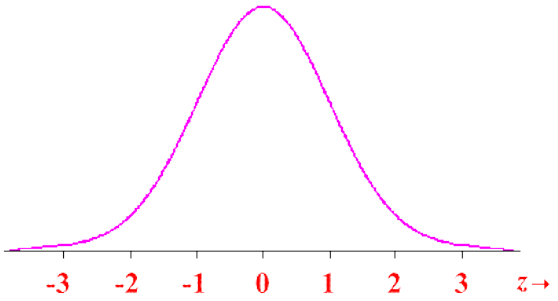

Fig. 2: The graph of the P.D.F. of normal distribution [11].

Definition:

In the probability theory and statistics, the exponential distribution is a class of continuous Probability distribution that describe the time between events in a Poisson process. In the formal notation, let x denote the waiting time until the first change occurs when observing a Poisson process, then it has an exponential distribution and its probability density function is defined by equation 3 [6]:

$$f(x) = \lambda e^{-\lambda x} \qquad , 0 \le x < \infty \qquad (3)$$

Where λ is the parameter of the distribution, called the rate parameter, In this paper we take $\lambda = 5$.

The graph of exponential distribution is shown in Fig. 3 [6].

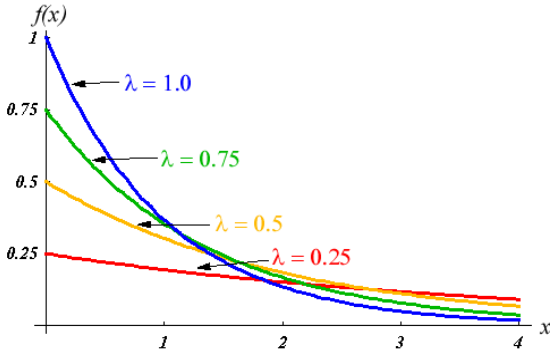

Fig. 3: The graph of the p.d.f. of exponential distribution [6].

Let xi, I=1,2,….,k be a independent and identically distributed Gaussian random variables with mean μ=0 and variance δ^2. A random variable z given by equation 4 is chi-square random variable with a k degree of freedom. it is an important continuous random variable [1].

$$z = \sum_{i=1}^{k} X_i^2 \qquad (4)$$

In probability theory and statistics, the chi-square distribution (or x^2 distribution) with a k degree of freedom is the distribution of a sum of the squares of k independent random variables. It is one of the most widely used probability distributions in inferential statistics, e.g. in hypothesis testing, or in construction of confidence intervals. The chi-square distribution is a special gamma distribution [1].

Definition:

Let z be a chi-square random variable with a k degrees of freedom, then the probability density function[3] is defined by equation 5 [1]:

$$f(z,k) = \frac{1}{2^{\frac{k}{2}} \times \Gamma(\frac{k}{2})} x^{\frac{k-2}{2}} e^{-\frac{k}{2}}, \quad z \ge 0. \qquad (5)$$

Where $\Gamma(\frac{k}{2})$ denotes the gamma function. The gamma function is defined by equation 6:

$$\Gamma(p) = \int_0^\infty t^{p-1} e^{-t}\, dt \ , \ p \ge 0. \qquad (6)$$

As the number of freedom is increased, the X^2 distribution become more symmetrical and when the

degrees of freedom become infinitely large, chi-square approaches normality [1]. The graph of chi-square distribution is shown in Fig. 4.

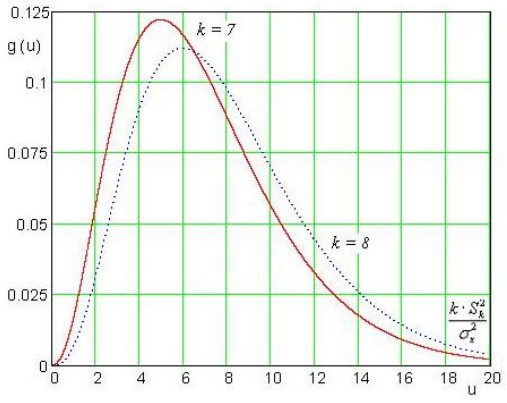

Fig. 4: The graph of the P.D.F. of normal distribution [7].

In this paper we assume k=3 and take the random variable z with 3 degrees of freedom. Then p.d.f. of z is equal to equation 7 [7]:

$$f(z) = \frac{1}{\sqrt{2\pi}} z^{\frac{1}{2}} e^{\frac{-z}{2}} \quad , \qquad z > 0 \tag{7}$$

Definition:

T is probability complete tree, if and only if the sum of the all adjacent weights of a vertex is either 1 or 0. Trees shown by Fig.1 and Fig.2 are probability complete trees. Note that the vertices have the total adjacency to edge weight of 0 are tree leaves [1].

Definition:

T is an n-complete probability tree if and only if:

1. It is a probability complete tree.
2. Each of vertices that have total adjacency to edge weight of 0 must be in depth n of the tree root.

All leaves must be in depth n, which is the depth of tree. Fig.1 shows a 2-complete probability tree, but the tree represented in Fig. 2 does not since it fails to satisfy the second part of this definition [1].

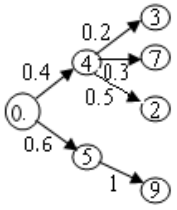

Fig. 5: A digital 2-complete probability tree

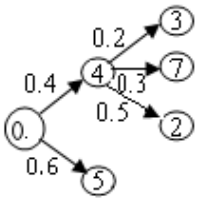

Fig. 6: A digital probabilistically complete tree

Definition:

Let t be a digital n-complete probability tree, and a_n be along with its leaf vertex on it. As shown in equation 8:

$$Vol_{an}=y=0.a_1a_2\ldots a_n \tag{8}$$

Then P_y is the probability of vol_{an}, if and only if have equation 9:

$$P_y=P_{O.a1}*P_{a1a2}*\ldots*P_{an\text{-}1an}=P_{O.a1}\Pi_{i=2}P_{aiai+1} \tag{9}$$

Data-oriented model of random variable z is presented by calculating the probability of digits by AF [11] and then these probabilities are used in probability complete tree as weights. AF is a fact that gives the probability of a random variable's digits by using its density function. In other words Ahmad fact describes the digits of a random variable are randomly distributed and specifies the probability of the kth digit is given of a determined random variable. Suppose that the random variable z has probability density function f(z). $p_{z_1^i}$ denotes the probability of the first digit of z to be i where i=0,1,2,3,…,9 in decimal base. $p_{z_1^i}$ is calculated by using AF as realation 10:

$$p_{z_1^i} = p(i < z < i+1) = \int_i^{i+1} f(z)dz \tag{10}$$

As we know, this relation gives the probability that the value of random variable z is in [i, i+1]. Considering AF, this is the probability that the first digit of z is equal i. For example $p_{z_1^2}$ shows the probability that first digit of z equal, 2. This notation is used to making digital 1-probability complete tree. By calculating second digit's probability, digital 2-probability complete tree can be made. For this purpose $p_{z_2^{ij}}$ is used. $p_{z_2^{ij}}$ shows the probability that the second digit of z equals j if the first digit has occurred i. $p_{z_2^{ij}}$ is computed from equation 11:

$$p_{z_2^{ij}} = p(i.j < z < i.j+1) = \int_{i.j}^{i.j+1} f(z)dz \tag{11}$$

For example $p_{z_2^{34}}$ is a probability that the z is in [3.4 , 3.5] and considering AF, this is the probability that the second digit of z to be 4 if the first has been equaled to 3. For calculating kth digit probability $p_{z_k^{ij\ldots m}}$, is used so probability of first, second,…, k-1th digits should be given.

Data-oriented theory is used to generate random numbers, In order to generate random number, some array, like digits bank is considered. In order to [1] for generate random number with two digits based on data-

oriented theory, $p_{z_1^i}, p_{z_2^{ij}}$ are calculated for i,j=0 to 9 that results for $p_{z_1^i}$ are shown in table 1. Here 11 digits banks are considered and form the value's obtained for $p_{z_1^i}$, the first two or three digits after decimal point are selected as the numbers of the iteration number for i are inserted in digits bank. For example if $p_{z_1^1} = 0.228$, so number 22 is selected which is considered the number of times that 1 is repeated and inserted in digits bank1. Table 2 shows the number of times a number appears in digit bank. Similarly, three digits after decimal point of $p_{z_2^{ij}}$ are selected and inserted in digits bank2…11. After recording these numbers in the digits banks, and randomly shuffling (or exchange or rotate) the content of them, one number is selected randomly from digits bank1, then depending on the selected numbers, one of the other ten digits banks is selected and froms this newly selected digits bank one element is chosen randomly too. In this method each digits bank has about 1000 elements with 4 bits for each of them that are 11000 elements and 5.5kb memory using in total. The architecture of this case is shown in Fig.8.

Table 1. Probability of fitrst digit

I	0	1	2	3	4	5	6	7	8	9
	0.186	0.228	0.163	0.127	0.091	0.055	0.038	0.027	0.016	0.0103

Table 2. Number of digits in digits bank1

I	0	1	2	3	4	5	6	7	8	9
Number of digits in digits bank1	18	23	16	13	9	6	4	3	2	1

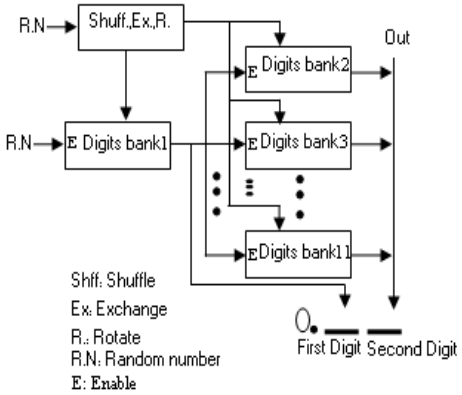

Fig. 7: Architecture of random number generator.

III. UNIVERSAL RANDOM NUMBER GENERATOR

Simulation result shows the random number generator with two digits has more precision, but needs more memory space [1] and each of these models generates random numbers with any given distributions separately, but we suggest universal random numbers generator with less memory space. In this section U.R.N.G is designed as follow:

From the values obtained for $p_{z_2^{ij}}$ as shown in table3, the first three digits after decimal point are selected as the numbers of the iteration number for ij in numbers bank. In this paper instate of 11 digits banks, one number bank for each distribution is considered and by calculating $p_{z_2^{ij}}$ for all i,j, distributions and inserting selection results in number banks, UDRNG can start, At first the type of distribution input by device and depending on the input distribution, one of the four numbers bank is selected, and by generating 2-bits digit randomly, content of selected numbers bank is shuffled or exchanged or rotated. Then from selected number bank, one element is chosen randomly.

Simulation results are shown in Fig.9 (a-d). With this method we can achieve a universal and cumulative random numbers.

The hardware structure of URNG is shown in Fig10, to implementation URNG in hardware structure we consider special memory or cache instate of numbers banks, decoder to select region of each distribution, 2-bit counter to select operation based on table3 for mix the content of cache.

Table 1: Number of digits in digits bank

Dist.	P\I	0	1	2	3	4	5	6	7	8	9
Uniform	$P_{u_1^i}$	$P_{u_1^0}$	$P_{u_1^1}$	$P_{u_1^2}$	$P_{u_1^3}$	$P_{u_1^4}$	$P_{u_1^5}$	$P_{u_1^6}$	$P_{u_1^7}$	$P_{u_1^8}$	$P_{u_1^9}$
Normal	$P_{N_1^i}$	$P_{N_1^0}$	$P_{N_1^1}$	$P_{N_1^2}$	$P_{N_1^3}$	$P_{N_1^4}$	$P_{N_1^5}$	$P_{N_1^6}$	$P_{N_1^7}$	$P_{N_1^8}$	$P_{N_1^9}$
Exponential	$P_{E_1^i}$	$P_{E_1^0}$	$P_{E_1^1}$	$P_{E_1^2}$	$P_{E_1^3}$	$P_{E_1^4}$	$P_{E_1^5}$	$P_{E_1^6}$	$P_{E_1^7}$	$P_{E_1^8}$	$P_{E_1^9}$
Chi-square	$P_{C_1^i}$	$P_{C_1^0}$	$P_{C_1^1}$	$P_{C_1^2}$	$P_{C_1^3}$	$P_{C_1^4}$	$P_{C_1^5}$	$P_{C_1^6}$	$P_{C_1^7}$	$P_{C_1^8}$	$P_{C_1^9}$

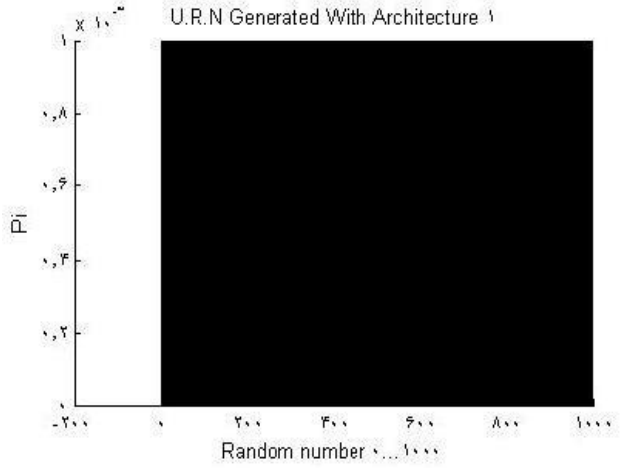

Fig.8.(a) UDRNG for normal distribution

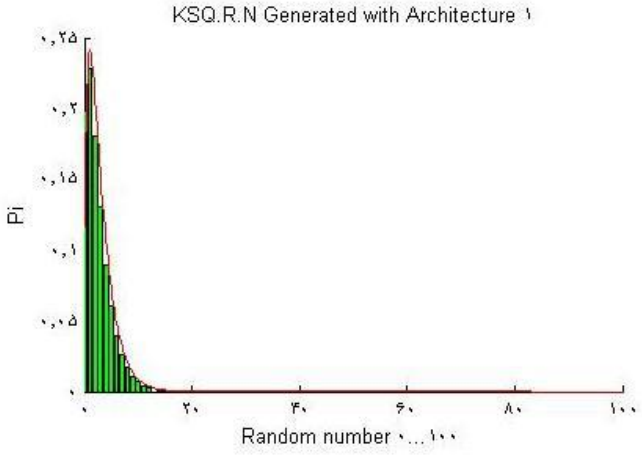

Fig.11.(d) UDRNG for exponential distribution

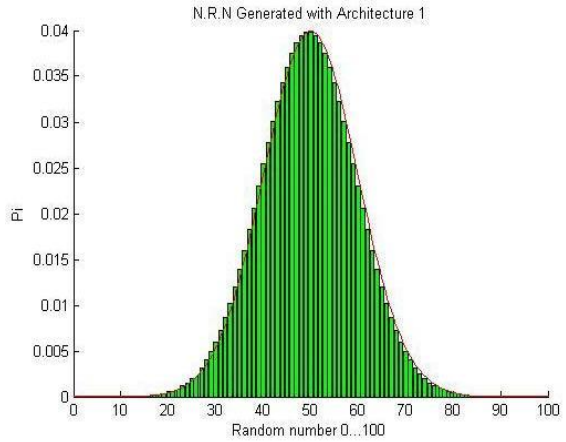

Fig.9.(b) UDRNG for uniform distribution

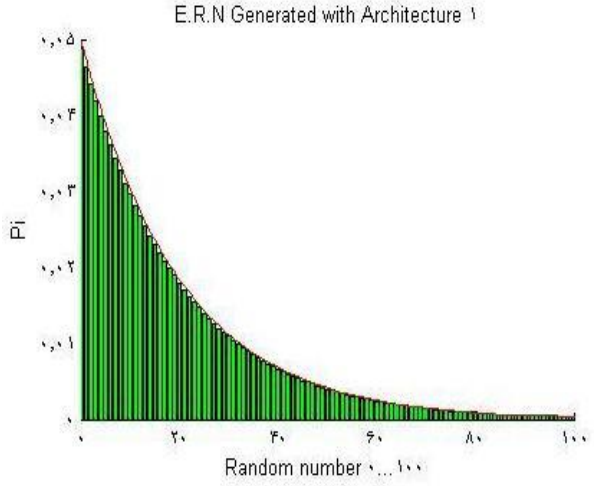

Fig.10.(c) UDRNG for chi-square distribution

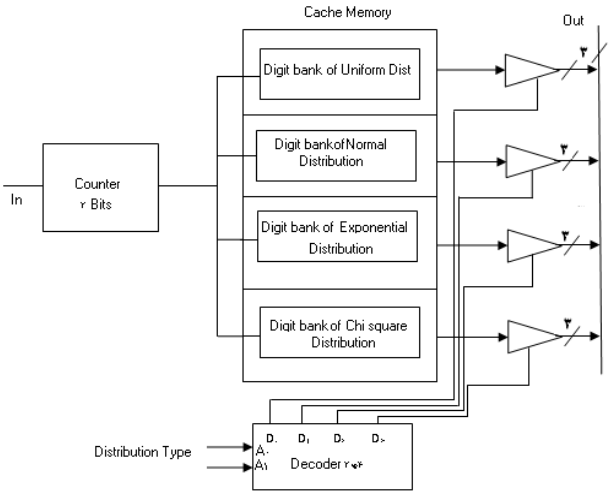

Fig.12: Hardware of universal random number generator

Table 4: Instruction set

operation	Op-code
Shuffle	00
Rotate to right	01
Exchange	10
Rotate to left	11

Now table 3 has some parameters in 2000 uniform numbers generated and 1000 numbers generated. Simulation is performed by MATLAB software and some parameter are shown in table3 that mean of square error (MSE) for uniform is achieved from 11 and for other distributions achieved from12.

$$MSE = \frac{1}{10}\sum_{i=0}^{9}\left(rf_i - \int_{100i}^{100(i+1)} f_s(s)ds\right)^2 \quad (12)$$

$$MSE = \frac{1}{100}\sum_{i=0}^{99}\left(rf_i - \int_{i}^{(i+1)} f_s(s)ds\right)^2 \quad (13)$$

where rf_i is the relative frequency of generated numbers in the interval [0.i,0.(i+1)] and

Table 5: Parameters table

Distribution	MSE	Memory cells used
Uniform	5.2e-5	1000*10bit=1.25kb
Normal	0.0253	1000*10bit=1.25kb
Exponential	0.0215	1000*10bit=1.25kb
Chi square	0.1219	1000*10bit=1.25kb

IV. CONCLUSION AND FURTHER WORKS

In this paper we presented a universal random number generator based on data-oriented to generate random number with uniform, normal, exponential and chi-square distributions. We believe that still we can decrease memory consumption by new techniques.

REFERENCES

[1] Habibzad-navin,A. ,E.S. Alikhani, M.Mirnia and S.Y. Torabi, "chi squre Random Variable Generator" India.2010 computational Intelligence and communication net works, PP.460-465.

[2] Habibzad-navin, A. and M.Mirnia, "Alternative views of the shortest path problem" in proceeding of the international Mathematical conference, Iran, 1999, PP.122-124.

[3] Habibizad-navin, A. and M.naghian Fesharaki and M.Teshnelab and M.Mirnia, "Probability graph and some of ios important structure". In proceeding of the 5th Seminar on probability and Statistic processes. Birjand, Iran, 2005, PP. 155-160.

[4] Habibizad-navin, A. and M.naghian Fesharaki and M.Teshnelab and M.Mirnia, "data – oriented modeling of uniform random variable: applied approach" In proceeding of World Academy of science, Engineering and Technology. Vienna, Austria, 2007, PP.382-385.

http://www.waset.org/pwaset/v21/v21-69.pdf

[5] Olfatkhah, R. and Habibzad-navin, A. and M.Mirnia, "Anovel model of normal random variable based on data –oriented theory", ICCEA. International conference on computer Engineering and applications, in press, 2010.

[6] Habibizad–navin , A. and R.Olfatkhah and M.Mirnia, "A Data – oriented model of exponential Random Varible", ICACC. International i conference of Advanced computer control china,2010,PP.603-607.

[7] Habibzad-navin, A. and N.Jafari Navimipour and M.Mirnia, "using labeled hyper multi digraph for Tabriz traffic modeling: data – oriented approach" Journal of Applied science, 9(15), ISSN: 1812-5654s, PP: 2808-2814, 2009

Rasoul farjaminezad is a Science Research in Computer Engineering at the Department of Computer Islamic Azad University-Tabriz Branch

