Nonparametric Regression with Trapezoidal Fuzzy Data

T. Razzaghnia
Department of Statistics, Roudehen Branch, Islamic Azad University, Roudehen - Iran.
Corresponding Author
E-mail: razaghnia@riau.ac.ir

S. Danesh
Department of Statistics, Science and Research Branch, Islamic Azad University,
Tehran - Iran.
E-mail: s.danesh@srbiau@riau.ac.ir

Abstract- This paper is an investigation into nonparametric fuzzy regression with crisp input and asymmetric trapezoidal fuzzy output. It analyzes the a nonparametric techniques in statistics, namely local linear smoothing (L-L-S) with trapezoidal fuzzy data to obtain the best smoothing parameters. In addition, it makes an analysis on one real-world datasets and calculates the goodness of fit to illustrate the application of the proposed method.

Key Words- *Nonparametric Fuzzy Regression, Trapezoidal Fuzzy Numbers, Local Linear Smoothing (L-L-S).*

*****

I. INTRODUCTION

Since the fuzzy regression was introduced by Tanaka et al.[1], several fuzzy regression approaches have been proposed, including the mathematical programming based methods [1], least squares based methods [2] and other methods [3]. In many real-world problems, it may be unrealistic to predetermine a fuzzy parametric regression relationship especially for a large dataset with a complicated underlying variation trend. Along this line of consideration, some other approaches have been developed to handle the fuzzy regression problems without predefining a specific form of the underlying regression relationship. For instance, Ishibushi and Tanaka [4] have suggested several fuzzy nonparametric regression methods by using the traditional back propagation networks. Also, statistical nonparametric smoothing techniques have achieved significant development in recent years [5]. These smoothing techniques are especially useful to handle the nonparametric regression problems and therefore there may be other promising tools for developing fuzzy nonparametric regression. In this aspect, Cheng and Lee [3] have extended the k-nearest neighbor (K-NN) and kernel smoothing (K-S) methods for the context of fuzzy nonparametric regression. In Wang et al. [6], the local linear smoothing method, the special case of the local polynomial smoothing technique, is fuzzified to handle the fuzzy nonparametric regression with crisp input and LR fuzzy output based on the distance measure proposed by Diamond [7]. Farnoosh et al. [8] used ridge estimation in nonparametric regression with triangular fuzzy data.

In this paper, we propose to fuzzify and analyze the three nonparametric regression techniques in statistical regression, namely local linear smoothing (L-L-S), the K- nearest neighbor smoothing (K-NN) and the kernel smoothing techniques (K-S) with trapezoidal fuzzy data.

II. PRELIMINARIES

A fuzzy number $\tilde{A}$ is a convex normalized fuzzy subset of the real line **R** with an upper semi-continuous membership function of bounded support [7].

Definition 2.1. An asymmetric trapezoidal fuzzy number $\tilde{A}$, denoted by $\tilde{A}=(a^{(1)},a^{(2)},a^{(3)},a^{(4)})$ is defined as:

$$\tilde{A}(x)=\begin{cases} L\left(\dfrac{a^{(2)}-x}{a^{(2)}-a^{(1)}}\right) & x<a^{(2)} \\ 1 & a^{(2)}\le x\le a^{(3)} \\ R\left(\dfrac{x-a^{(3)}}{a^{(4)}-a^{(3)}}\right) & x>a^{(3)} \end{cases}$$

where $a^{(1)},a^{(2)}a^{(3)},a^{(4)}$ are four parameters of the asymmetric trapezoidal fuzzy number.

Definition 2.2. Suppose that $\tilde{A}=(a^{(1)},a^{(2)},a^{(3)},a^{(4)})$ and $\tilde{B}=(b^{(1)},b^{(2)},b^{(3)},b^{(4)})$ are two trapezoidal fuzzy numbers. Diamond distance between $\tilde{A}$ and $\tilde{B}$ can be expressed as:

$$d^2(\tilde{A},\tilde{B})=(a^{(1)}-b^{(1)})^2+(a^{(2)}-b^{(2)})^2 + (a^{(3)}-b^{(3)})^2+(a^{(4)}-b^{(4)})^2$$

 ISSN: 2321-8169
3826 - 3831

This distance measures the closeness between two trapezoidal fuzzy membership functions when $d^2(\tilde{A},\tilde{B})=0$.

It means that the membership functions of $\tilde{A}$ and $\tilde{B}$ are equal.

Let $F=\left\{\tilde{Y}:\tilde{Y}=(y^{(1)},y^{(2)},y^{(3)},y^{(4)})\right\}$ be a set of all trapezoidal fuzzy numbers. The following univariate fuzzy nonparametric regression model is considered by $Y=F(x)\{+\}\varepsilon$. In this model, X is a crisp independent variable (input) and Y is a symmetric trapezoidal fuzzy dependent variable (output). ε is an error term, and $\{+\}$ is an operator whose definition depends on the fuzzy ranking method used.

In this paper, for the nonparametric regression techniques, K-N-N and K-S are based on the concept of local averaging. In other words, the estimated value of the regression surface at point k_0 is the weighted average of the responses of the observations in the neighborhood of k_0.

Definition 2.3. Let $K_i, i=1,2,\ldots,n$ where the index is in ascending order, then the smoothing function based on local averaging can be represented as:

$$S(K=K_i)=\underset{i-k\le j\le i+k}{AVE}(Y_j)=$$

$$\underset{i-k\le j\le i+k}{AVE}(y_j^{(1)},y_j^{(2)},y_j^{(3)},y_j^{(4)})$$

where AVE denotes the mean, median or any weighted average.

III. Smoothing methods for trapezoidal fuzzy numbers

The basic idea of smoothing is that if a function f is fairly smooth, then the observations made at and near x should contain information about value of x. Thus, it should be possible to use local averaging of the data x to construct an estimator for $F(x)$ which is called the smoother. There are several smoothing techniques. We proposed K-nearest neighbor smoothing (K-NN), kernel-smoothing (K-S) and local linear smoothing (L-L-S) methods for trapezoidal variable in this section.

In the following discussion, asymmetric trapezoidal fuzzy numbers are applied as asymmetric trapezoidal membership functions for deriving nonparametric regression model based on the smoothing parameters.

These models are considered univariate fuzzy nonparametric regression model as:

$$\tilde{Y}=F(x)\{+\}\varepsilon= \qquad (1)$$

$$\left(Y^{(1)}(x),Y^{(2)}(x),Y^{(3)}(x),Y^{(4)}(x)\right)\{+\}\varepsilon$$

where Y is a trapezoidal fuzzy dependent variable as output. x is a crisp independent variable as input , $x\in\square$, and x domain is assumed to be D. $F(x)$ is a mapping $D\to F$. The definition of the smoothing method for trapezoidal fuzzy variables is as follows:

- Local linear smoothing method (L-L-S)

In the following discussion, Razzaghnia et al. [9] proposed the first linear regression analysis with trapezoidal coefficients. Asymmetric trapezoidal fuzzy numbers are applied as asymmetric trapezoidal membership functions for deriving bivariate regression model. A univariate regression model can be expressed as:

$$\hat{\tilde{Y}}_i=\tilde{A}_0+\tilde{A}_1X_i=\left(a_0^{(1)},a_0^{(2)},a_0^{(3)},a_0^{(4)}\right)$$

$$+\left(a_1^{(1)},a_1^{(2)},a_1^{(3)},a_1^{(4)}\right)\mathrm{X_i} \qquad (2)$$

This model can be rewritten as

$$\hat{\tilde{Y}}_i=(a_0^{(1)}+a_1^{(1)}X_i,a_0^{(2)}+a_1^{(2)}\mathrm{X_i},a_0^{(3)}$$

$$+a_1^{(3)}\mathrm{X_i},a_0^{(4)}+a_1^{(4)}\mathrm{X_i})$$

where $i=1,\ldots,n$ and n is the sample size.

and $\tilde{Y}_i=\left(Y_i^{(1)},Y_i^{(2)},Y_i^{(3)},Y_i^{(4)}\right)$ is an observed value for $i=1,\ldots,n$. So $\hat{\tilde{Y}}_{i.L}$ and $\hat{\tilde{Y}}_{i.R}$ are the left bound and right bound of the predicted $\hat{\tilde{Y}}_i$ at membership h level. Also $\tilde{Y}_{i,L}$ and $\tilde{Y}_{i,R}$ are left bound and right bounds of observed $\tilde{Y}_i$ at membership h level.

Thereupon,

 ISSN: 2321-8169
3826 - 3831

$$\hat{\tilde{Y}}_{i,L} = ha_0^{(2)} + ha_1^{(2)}X_i + (1-h)a_0^{(1)} + (1-h)a_1^{(1)}X_i$$

$$\hat{\tilde{Y}}_{i,R} = ha_0^{(3)} + ha_1^{(3)}X_i + (1-h)a_0^{(4)} + (1-h)a_1^{(4)}X_i$$

$$\tilde{Y}_{i,L} = hY_i^{(2)} + (1-h)Y_i^{(1)}$$

$$\tilde{Y}_{i,R} = hY_i^{(3)} + (1-h)Y_i^{(4)}.$$

Let $(X_i, \tilde{Y}_i)$ be a sample of the observed crisp inputs and trapezoidal fuzzy outputs with underlying fuzzy regression function of model (2).

$F(x)$ is estimated at any $x \in D$ based on $(x_i, \tilde{Y}_i)$ for $i = 1,\ldots,n.$ When the local linear smoothing technique is used, we shall estimate $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ for each $x \in D$ by using the distance proposed by Diamond [7] as a measure of the fit (Definition 2.2).
This distance is used to fit the fuzzy nonparametric model (1).

Let $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ have continuous derivatives in the domain $x \in D$. Then for a given $x_0 \in D$ and Taylors expansion, $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ can be locally approximated in neighborhood of x_0, respectively by the following linear functions:

$$Y^{(1)}(x) \square \hat{Y}^{(1)}(x) = Y^{(1)}(x_0) + Y'^{(1)}(x_0)(x - x_0) \quad (3)$$

$$Y^{(2)}(x) \square \hat{Y}^{(2)}(x) = Y^{(2)}(x_0) + Y'^{(2)}(x_0)(x - x_0) \quad (4)$$

$$Y^{(3)}(x) \square \hat{Y}^{(3)}(x) = Y^{(3)}(x_0) + Y'^{(3)}(x_0)(x - x_0) \quad (5)$$

$$Y^{(4)}(x) \square \hat{Y}^{(4)}(x) = Y^{(4)}(x_0) + Y'^{(4)}(x_0)(x - x_0) \quad (6)$$

where $Y'^{(1)}(x_0), Y'^{(2)}(x_0), Y'^{(3)}(x_0)$ and $Y'^{(4)}(x_0)$ are respectively, the derivatives of $Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x)$ and $Y^{(4)}(x)$ based on Diamond distance (Definition 2.2) and the local linear smoothing method is estimated at x_0,

$$F(x_0) = (Y^{(1)}(x_0), Y^{(2)}(x_0), Y^{(3)}(x_0), Y^{(4)}(x_0))$$

by minimizing

$$\sum_{i=1}^{n} d^2\left(\tilde{Y}_i, \hat{\tilde{Y}}_i\right) = \sum_{i=1}^{n} d^2\left(\left(Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}\right), \left(\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}\right)\right) K_h(|x_i - x_0|) \quad (7)$$

With respect to $Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}$ and $\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}$ for the given kernel k(.) and smoothing parameter h, where

$$K_h(|x_i - x_0|) = k\left(\frac{|x_i - x_0|}{h}{h}\right)$$ for

$i = 1,\ldots,n$ are a sequence of weights at x_0 .Two commonly used kernel functions are parabolic shape functions:

$$k_1(x) = \begin{cases} 0.75(1-x^2) & if\ |x| \le 1 \\ 0 & otherwise \end{cases}$$

and Gaussian function:

ISSN: 2321-8169
3826 - 3831

$$k_2(x) = (2\pi)^{-1/2} \exp(\frac{-x^2}{2})$$

Also, by substituting (3), (4), (5) and (6) at (7), the following can be obtained

$$\sum_{i=1}^{n} d^2\left(\tilde{Y}_i, \hat{\tilde{Y}}_i\right) =$$

$$\sum_{i=1}^{n} d^2(\left(Y_i^{(1)}, Y_i^{(2)}, Y_i^{(3)}, Y_i^{(4)}\right),$$

$$\left(\hat{Y}_i^{(1)}, \hat{Y}_i^{(2)}, \hat{Y}_i^{(3)}, \hat{Y}_i^{(4)}\right)) K_h(|x_i - x_0|)$$

$$= \sum_{i=1}^{n} (Y_i^{(1)} - Y^{(1)}(x_0) - Y'^{(1)}(x_0)$$

$$(x_i - x_0))^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n} (Y_i^{(2)} - Y^{(2)}(x_0) - Y'^{(2)}(x_0)$$

$$(x_i - x_0))^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n} (Y_i^{(3)} - Y^{(3)}(x_0) - Y'^{(3)}(x_0)$$

$$(x_i - x_0))^2 K_h(|x_i - x_0|)$$

$$+ \sum_{i=1}^{n} \left(Y_i^{(4)} - Y^{(4)}(x_0) - Y'^{(4)}(x_0)(x_i - x_0)\right)^2$$

$$K_h(|x_i - x_0|) \quad (8)$$

By solving this weighted least-squares problem, the following can be obtained

$$Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x), Y^{(4)}(x),$$

$$Y'^{(1)}(x), Y'^{(2)}(x), Y'^{(3)}(x), Y'^{(4)}(x)$$

at x_0. So the estimation $F(x)$ at x_0 is:

$$\hat{\tilde{Y}}(x_0) = (\hat{Y}^{(1)}(x_0), \hat{Y}^{(2)}(x_0),$$

$$\hat{Y}^{(3)}(x_0), \hat{Y}^{(4)}(x_0)) \quad .$$

Equation (8) has eight unknown parameters

$$Y^{(1)}(x), Y^{(2)}(x), Y^{(3)}(x), Y^{(4)}(x),$$

$$Y'^{(1)}(x_0), Y'^{(2)}(x_0), Y'^{(3)}(x_0), Y'^{(4)}(x_0)$$

to derive a formula for the unknown parameters nonparametric regression based on minimizing this distance, the derivatives (8) with respect to the eight unknown parameters need to be derived, set to zero and solve the eight unknown parameters.

According to the principle of the weighted least-squares and utilizing matrix notations, we can obtain

$$(\hat{Y}^{(1)}(x), \hat{Y}'^{(1)}(x))^T = (X^T(x_0) \quad (9)$$

$$W(x_0; h) X(x_0))^{-1} X^T(x_0) W(x_0; h) \tilde{Y}^{(1)}$$

$$(\hat{Y}^{(2)}(x), \hat{Y}'^{(2)}(x))^T = (X^T(x_0) \quad (10)$$

$$W(x_0; h) X(x_0))^{-1} X^T(x_0) W(x_0; h) \tilde{Y}^{(2)}$$

$$(\hat{Y}^{(3)}(x), \hat{Y}'^{(3)}(x))^T = (X^T(x_0) \quad (11)$$

$$W(x_0; h) X(x_0))^{-1} X^T(x_0) W(x_0; h) \tilde{Y}^{(3)}$$

$$(\hat{Y}^{(4)}(x), \hat{Y}'^{(4)}(x))^T = (X^T(x_0) \quad (12)$$

$$W(x_0; h) X(x_0))^{-1} X^T(x_0) W(x_0; h) \tilde{Y}^{(4)}$$

where

$$X(x_0) = \begin{pmatrix} & \\ 1 & x_1 - x_0 \\ 1 & x_2 - x_0 \\ \vdots & \vdots \\ 1 & x_n - x_0 \end{pmatrix}, \tilde{Y}^{(1)} = \begin{pmatrix} Y_1^{(1)} \\ Y_2^{(1)} \\ \vdots \\ Y_n^{(1)} \end{pmatrix},$$

$$\tilde{Y}^{(2)} = \begin{pmatrix} Y_1^{(2)} \\ Y_2^{(2)} \\ \vdots \\ Y_n^{(2)} \end{pmatrix}, \tilde{Y}^{(3)} = \begin{pmatrix} Y_1^{(3)} \\ Y_2^{(3)} \\ \vdots \\ Y_n^{(3)} \end{pmatrix},$$

$$\tilde{Y}^{(4)} = \begin{pmatrix} Y_1^{(4)} \\ Y_2^{(4)} \\ \vdots \\ Y_n^{(4)} \end{pmatrix}$$

 ISSN: 2321-8169
3826 - 3831

$$W\left(x_0;h\right)=\text{Diag}(K_h\left(\left|x_1-x_0\right|\right),$$

and $K_h\left(\left|x_2-x_0\right|\right),\ldots,K_h\left(\left|x_n-x_0\right|\right))$

is a $n\times n$ diagonal matrix with its diagonal elements being $K_h(\left|x_i-x_o\right|)$ for $i=1,\ldots,n$ and symbol T is transpose of a matrix. If we suppose $e_1=(1,0)^T$ and

$$H\left(x_0;h\right)=(X^T\left(x_0\right)W\left(x_0;h\right)$$

$$X\left(x_0\right))^{-1}X^T\left(x_0\right)W\left(x_0;h\right)$$

The estimate of $F(x)$ at x_0 is

$$\hat{\tilde{Y}}\left(x\right)=(\hat{Y}^{(1)}\left(x_0\right),\hat{Y}^{(2)}\left(x_0\right),$$
$$\hat{Y}^{(3)}\left(x_0\right),\hat{Y}^{(4)}\left(x_0\right))$$

$$=(e_1^T H\left(x_0;h\right)\tilde{Y}^{(1)},e_1^T H\left(x_0;h\right)\tilde{Y}^{(2)},$$
$$e_1^T H\left(x_0;h\right)\tilde{Y}^{(3)},e_1^T H\left(x_0;h\right)\tilde{Y}^{(4)}) \qquad (13)$$

- Smoothing parameters selection

The most important aspect for averaging techniques and local linear smoothing method is selecting the size of neighborhood to average k and parameter h. There are different methods for selecting parameter h such as the cross-validation method, and generalized cross validation which are used to obtain parameter h. Let

$$\hat{\tilde{Y}}\left(x_i,h\right)=(\hat{Y}^{(1)}\left(x_i,h\right),\hat{Y}^{(2)}\left(x_i,h\right),$$
$$\hat{Y}^{(3)}\left(x_i,h\right),\hat{Y}^{(4)}\left(x_i,h\right))$$

The fuzzified cross-validation procedure (CV) for selecting parameter h local linear smoothing method based on Diamond distance is defined as:

$$CV\left(h\right)=\frac{1}{n}\sum_{i=1}^{n}d^2\left(\tilde{Y}_i,\hat{\tilde{Y}}_i\right)=$$

$$\frac{1}{n}\sum_{i=1}^{n}((Y_i^{(1)}-\hat{Y}_i^{(1)})^2+$$

$$(Y_i^{(2)}-\hat{Y}_i^{(2)})^2+(Y_i^{(3)}-\hat{Y}_i^{(3)})^2$$
$$+((Y_i^{(4)}-\hat{Y}_i^{(4)})^2) \qquad (16)$$

as its minimization gives the h optimal value.

$$CV\left(h_0\right)=\min_{h>o}CV\left(h\right)$$

In fact, we may compute $CV\left(h\right)$ for a series of value of h to search for h.

So selected optimal value of h by the $CV\left(h\right)$ nearly depends on the degree of smoothness of Y_{iL} and Y_{iR} .

Large value of h leads to lack-of-fit and small value of h makes over-fit.

IV. Numerical Example

In this section, there are an example in which the input is a crisp number and the output is a trapezoidal fuzzy number. We estimate the values by using three smoothing methods. Then these methods can be compared with each other and for this purpose, their GOF and their charts are used.

Example : This example is a generated dataset in the same way as that in Cheng and Lee [3]. The following function is considered $f\left(x\right)=\frac{x^2}{5}+2e^{\frac{x}{10}}$

So x_i is uniformly generated within the interval [0, 1] and i=1,…,100,

$$\tilde{Y_i}=\left(\mathrm{Y_i^{(1)}},\mathrm{Y_i^{(2)}},\mathrm{Y_i^{(3)}},\mathrm{Y_i^{(4)}}\right)=$$

$$(y_i-e_i,y_i+\frac{1}{3}e_i,y_i+\frac{2}{3}e_i,y_i+e_i)',$$

So

$$y_i=f\left(X_i\right)+rand[-0.5,0.5]\text{ and}$$
$$e_i=1/4f\left(X_i\right)+rand[0,1].$$

Local Linear smoothing method is applied to the fitting model. So Gauss and Parabolic shape kernel are used to produce the weight sequence for local linear smoothing Table 3 shows smoothing parameter selected by cross-validation procedure results from different methods. Figures 4, 5 and 6 show the results of three methods. These results can be compared using figure 3 and table 3. Like the previous example, L-L-S method is better than K-NN, and K- S methods. In table 3, GOF of L-L-S method is lower than K-NN, K- S methods.

ISSN: 2321-8169
3826 - 3831

Table 1

The obtained results of different methods for sample 2

method	kernel	Smoothing parameter	GOF
LLS	Gauss	0.43	0.0045
	Parabolic shape	1.2	0.0046

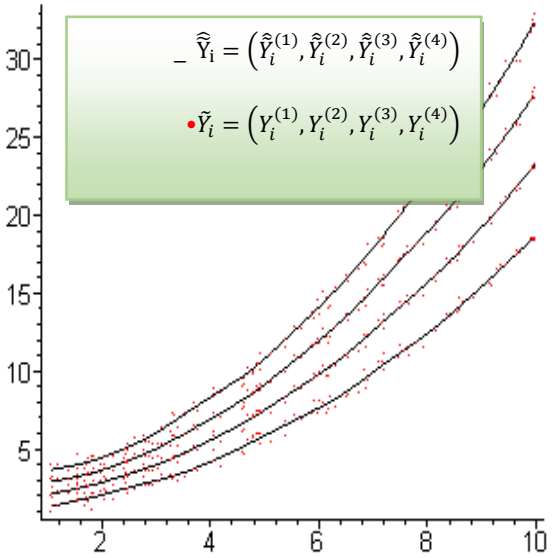

Figure1: Obtained results by L-L-S method with Gausian kernel for h=0.43

REFERENCES

[1] H. Tanaka, S. Uejima, K. Asia, "Linear regression analysis with fuzzy model", IEEE Transactions on Systems, Man, and Cybernetics 12 ,1982, pp 903-907.

[2] P. T. Chang, E. S. Lee, A generalized fuzzy weighted least-squares regression, Fuzzy Sets and Systems 82, (1996) 289-298.

[3] C. B. Cheng, E. S. Lee, "Nonparametric fuzzy regression K-NN and Kernel Smoothing techniques", Computers and Mathematics with Applications 38 ,1999, pp 239-251

[4] H. Ishibushi, H. Tanaka," Fuzzy regression analysis using neural networks", Fuzzy Sets and Systems 50 ,1992, pp 257-265.

[5] W. Hardle, "Applied Nonparametric Regression", Cambridge University Press, New York, 1990.

[6] N.Wang, W.X. Zhang and C.L Mei, "Fuzzy nonparametric regression based on local linear smoothing technique", Information Sciences 177 ,2007, pp 3882-3900.

[7] P. Diamond, Fuzzy least squares, Information Sciences 46 ,1988, pp 141-157.

[8] R. Farnoosh, J. Ghasemian and o. Solaymani Fard, A modification on ridge estimation for fuzzy nonparametric regression, Iraninan Journal of Fuzzy systems 9 ,2012, pp 75-88.

[9] T. Razzaghnia, E. Pasha, E. Khorram, A. Razzaghnia, "Fuzzy linear regression analysis with trapezoidal coefficients", First Joint Congress On Fuzzy And Intelligent Systems 2007, Aug. 29-31, Mashhad, Iran.