A Comparison of Data Fusion, Neural Network and Statistical Pattern Recognition Technologies to a Multi-Sensor Target ID and Classification Problem

Dr. Buddy H. Jeun, N. S. (Jay) Jayaraman
Lockheed Martin Aeronautical Systems
86 South Cobb Drive, Marietta, Georgia 30063

ABSTRACT

It has been widely known that data fusion, neural network and statistical pattern recognition technologies can be applied to target identification and classification problems. The main objective of this paper is to find out which of these techniques would be easy to use and provide acceptable results.

We had selected the "Multi-sensor Correlation Model" [1] from the field of data fusion technology. The concept of this model is based on the coefficient of similarity. For target identification problem, one have to estimate the coefficient of similarity between a known target (X) and the target (Y) to be identified. If the coefficient is closed to one , then it implied that target (Y) is the same as target (X), otherwise if the coefficient is close to zero, then it implied that target (Y) is not the same as target (X). It is mathematical simple and easy to implement.

The "Bayesian Model"[2] was selected from the field of statistical pattern recognition technology, This is a conditional probability model. For target identification problem, one have to calculate the posterior probability of a known target (X) given the target (Y) to one to be identified. If the conditional probability is close to one , then it implied that target (x) and target (Y) is the same, otherwise if it is close to zero, then it implied that target(X) and target(Y) is not the same. This model required multivariate normal assumption, probability density function, and apriori probability of the targets. It is not easy to apply.

The " Backpropagation Model"[3] was selected from the field of neural network technology, It is a three layered network; input, hidden and output layers. For target identification problem, one has to train the network with the known target (X), then apply the unknown target(Y) to the trained network as an input layer, if the output layer has a higher energy value , then the unknown target (Y) is being identified. This technique is very hard to implement, since it is necessary to find a collect number of hidden elements to set up a network for training. The network has to be well trained, if good result are to be expected.

We use two published [4] numerical data set, to created 150 targets. Each target has four distinct feature elements. We then applied the same data set to all of the three technologies and obtained the following results:

Data fusion technique :	94.7% correctly identified
Neural Network technique:	93.6% correctly identified.
Statistical Pattern Recognition technique:	94.6% correctly identified.

We conclude that Data Fusion technique is the winner for this particular application.

REPORT DOCUMENTATION PAGE	Form Approved OMB No. 0704-0188

1. REPORT DATE (DD-MM-YYYY)	2. REPORT TYPE	3. DATES COVERED (FROM - TO)
01-01-1998	Conference Proceedings	xx-xx-1998 to xx-xx-1998

4. TITLE AND SUBTITLE
A Comparison of Data Fusion, Neural Network and Statistical Pattern Recognition Technologies to a Multi-Sensor Target ID and Classification Problem
Unclassified

5a. CONTRACT NUMBER
5b. GRANT NUMBER
5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Jeun, Buddy H. ;
Jayaraman, N. S. ;

5d. PROJECT NUMBER
5e. TASK NUMBER
5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME AND ADDRESS
Lockheed Martin Aeronautical Systems
86 South Cobb Drive
Marietta, GA30063

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME AND ADDRESS
Director, CECOM RDEC
Night Vision and Electronic Sensors Directorate, Security Team
10221 Burbeck Road
Ft. Belvoir, VA22060-5806

10. SPONSOR/MONITOR'S ACRONYM(S)
11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
APUBLIC RELEASE
,

13. SUPPLEMENTARY NOTES
See Also ADM201041, 1998 IRIS Proceedings on CD-ROM.

14. ABSTRACT
It has been widely known that data fusion, neural network and statistical pattern recognition technologies can be applied to target identification and classification problems. The main objective of this paper is to find out which of these techniques would be easy to use and provide acceptable results. We had selected the ?Multi-sensor Correlation Model? [1] from the field of data fusion technology. The concept of this model is based on the coefficient of similarity. For target identification problem, one have to estimate the coefficient of similarity between a known target (X) and the target (Y) to be identified. If the coefficient is closed to one , then it implied that target (Y) is the same as target (X), otherwise if the coefficient is close to zero, then it implied that target (Y) is not the same as target (X). It is mathematical simple and easy to implement. The ?Bayesian Model?[2] was selected from the field of statistical pattern recognition technology, This is a conditional probability model. For target identification problem, one have to calculate the posterior probability of a known target (X) given the target (Y) to one to be identified. If the conditional probability is close to one , then it implied that target (x) and target (Y) is the same, otherwise if it is close to zero, then it implied that target(X) and target(Y) is not the same. This model required multivariate normal assumption, probability density function, and apriori probability of the targets. It is not easy to apply. The ? Backpropagation Model?[3] was selected from the field of neural network technology, It is a three layered network; input, hidden and output layers. For target identification problem, one has to train the network with the known target (X), then apply the unknown target(Y) to the trained network as an input layer, if the output layer has a higher energy value , then the unknown target (Y) is being identified. This technique is very hard to implement, since it is necessary to find a collect number of hidden elements to set up a network for training. The network has to be well trained, if good result are to be expected. We use two published [4] numerical data set, to created 150 targets. Each target has four distinct feature elements. We then applied the same data set to all of the three technologies and obtained the following results: Data fusion technique : 94.7% correctly identified Neural Network technique: 93.6% correctly identified. Statistical Pattern Recognition technique: 94.6% correctly identified. We conclude that Data Fusion technique is the winner for this particular application.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19. NAME OF RESPONSIBLE PERSON
			Public Release	16	Fenster, Lynn lfenster@dtic.mil
a. REPORT Unclassified	b. ABSTRACT Unclassified	c. THIS PAGE Unclassified			19b. TELEPHONE NUMBER International Area Code Area Code Telephone Number 703767-9007 DSN 427-9007

(A) INTRODUCTION :

Data fusion, neural network and statistical pattern recognition technologies have been considered as powerful techniques to solve the positive target identification and classification problems. The capability of positive target identification and classification will play an important role in the advanced avionics of the future fighter. “first see, first kill” will be the underlying design principle for the development of the war fighter aircraft.

The main objective of this paper is to evaluate these techniques by comparing the theoretical concept and application of these advanced technologies. This paper consists of five major parts; (1) Data fusion technology (2) statistical pattern recognition technology (3) neural network technology (4) knowledge data base and (5) Simulation.

In the first part, data fusion technology, Coefficient of Similarity model, its mathematical expression , properties and decision logic is introduced. In the second part, statistical pattern recognition technology, the Bayesian model , its mathematical expression, probability properties, and decision logic is introduced. In the third part, neural network technology, Backpropagation model, its three layer architecture model, mathematical equations associated with the BP model and its decision logic in application is discussed. In the fourth part, knowledge data base is the main topic of discussion. Lastly, the discussion is focused on the simulation of positive target identification and classification of targets.

(B) DATA FUSION TECHNOLOGY :

(1) The Coefficient of Similarity Model (CSM)

One of the simple models of Data Fusion technology is the Coefficient of Similarity Model. The CSM can be used to measure the relationship between two target feature vectors. Mathematically, the Coefficient of Similarity Model can be expressed as below:

$$R_{XY} = X \bullet Y / (X \bullet X - X \bullet Y + Y \bullet Y) \qquad [1]$$

where $X = \{ X1, X2, \ldots\ldots\ldots\ldots.Xk \}$

$Y = \{ Y1, Y2, \ldots\ldots\ldots\ldots..Yk \}$

$X \bullet X = \Sigma(Xi \,.\, Xi)$

$X \bullet Y = \Sigma(Xi \,.\, Yi)$

$Y \bullet Y = \Sigma(Yi \,.\, Yi)$

X and Y are target feature vectors

(2) Properties of the Coefficient of Similarity Model

(a) Show that $Rxy = 1.0$ if $X=Y$

proof : Since $R_{XY} = X \bullet Y / (X \bullet X - X \bullet Y + Y \bullet Y)$

and $X = Y$ then

$X \bullet X = \Sigma(Xi \,.\, Xi)$

$X \bullet Y = \Sigma(Xi.Xi)$

$Y \bullet Y = \Sigma(Xi \,.\, Xi)$

by substitution, we have:

$Rxy = \Sigma(Xi \,.\, Xi)/\{\Sigma(Xi \,.\, Xi) + \Sigma(Xi \,.\, Xi) - \Sigma(Xi \,.\, Xi)\}$

$= \Sigma(Xi \,.\, Xi)/\Sigma(Xi \,.\, Xi)$

$= 1.0$

therefore $Rxy = 1.0$ for $X=Y$

(b) Show that $Rxy = 0.0$ for $X=0$ and $Y =/0$

proof : since $Rxy= X \bullet Y/(X \bullet X + Y \bullet Y - X \bullet Y)$

and $X = \{ 0., 0., \ldots\ldots.0. \}$

and $Y = \{ Y1, Y2, \ldots\ldots\ldots\ldots..Yp \}$

by substitution, we have:

$X \bullet X = \Sigma(0.\ 0) = 0.0$

$X \bullet Y = \Sigma(0.\ yi) = 0.0$

$Y \bullet Y = \Sigma(Yi \,.\, Yi)=k=/0$, k is a non zero constant

that is $Rxy = 0/(0 - 0 + k) = 0 / k = 0$

(c) show that $Rxy =0$ for $X =/0$ and $Y = 0$

proof : since $Rxy = X \bullet Y/(X \bullet X + Y \bullet Y - X \bullet Y)$

and $X = \{ x1, x2, \ldots..xk\}$

and $Y = \{ 0, 0, \ldots\ldots\ldots.0 \}$

$X \bullet X = \Sigma(xi.xi) = k=/0$, k is a non zero constant

$X \bullet Y = \Sigma(xi.0) = 0$

$Y \bullet Y = \Sigma(0.0) = 0$

by substitution, we have:

i.e. $Rxy = 0/(k - 0 + 0) = 0/k= 0.0$

(d) show that $Rxy = 0$ for $X = 0$ and $Y = 0$

proof : since $Rxy = X \bullet Y / (X \bullet X + Y \bullet Y - X \bullet Y)$

and $X = \{ 0, 0, \ldots\ldots\ldots 0 \}$

and $Y = \{ 0, 0, \ldots\ldots\ldots 0 \}$

$X \bullet X = \Sigma(0.0) = 0$

$X \bullet Y = \Sigma(0.0) = 0$

$Y \bullet Y = \Sigma(0.0) = 0$

by substitution, we have:

i.e. $Rxy = 0/(0 - 0 + 0) = 0/0 = 0.0$ (Le Hospital's rule)

(e) Show that $0.0 < Rxy < 1.0$ for $X > 0.0$ & $Y > 0.0$

proof : since $Rxy = X \bullet Y / [X \bullet X + Y \bullet Y - X \bullet Y]$

and $X \bullet X = \sum xi.xi$

and $Y \bullet Y = \sum yi.yi$

and $X \bullet Y = \sum xi.yi$

and $X \bullet X + Y \bullet Y - X \bullet Y$

$= \sum(xi^2 - 2\ xi\ yi + xiyi^2 + yi^2)$

$= \sum(xi - yi)^2 + \sum xiyi$

by substitution, we have:

$Rxy = \sum xiyi / (\sum(xi - yi)^2 + \sum xiyi)$

since $\sum(xi - yi)^2 > 0$ for all i

that is $Rxy >= 0$ and $Rxy < 1$ for $X > 0$ and $Y > 0$

(3) Decision

For a given target feature vector $X = \{ x1, x2, x3, \ldots\ldots\ldots\ldots xp\}$
and another target feature vector $Y = \{ y1, y2, y3, \ldots\ldots\ldots\ldots yp\}$

(a) If $Rxy \rightarrow 1.0$ then
Target X is positively identified as target Y

(b) If $Rxy \rightarrow 0.0$ then
Target X is not the same as target Y

(c) If $Rxy = 0.5$ then
No decision can be make on target X and target Y

To extract a well defined target feature vector from the multi sensor data is not as easy as we think, It requires some expert knowledge on the target to be identified.

Feature extraction is the major problem for data fusion, statistical pattern recognition and neural network technologies in the field of target identification.

(4) Examples:

In order to show the capabilities of the Coefficient of Similarity Model (CSM) , we purposely make up couple of numerical examples as below:

Example #1

Suppose , for a given target X from one sensor and another target Y from another sensor and target X and target Y have the target feature vectors as follows:

$$X = \{ 0., 0., 0., 0., 0., 0. \}$$
$$Y = \{ 1., 1., 1., 1., 1., 1. \}$$

by applying the Coefficient of Similarity Model (CSM), we have:

$$X \bullet X = \Sigma(Xi \,.\, Xi) = 0.0$$
$$X \bullet Y = \Sigma(Xi \,.\, Yi) = 0.0$$
$$Y \bullet Y = \Sigma(Yi \,.\, Yi) = 6.0$$

and by substitution, we have:

$$R_{XY} = X \bullet Y \,/\, (X \bullet X - X \bullet Y + Y \bullet Y) = 0.0$$

one can conclude that target X and target Y are not the same target.

Example #2

Suppose , for a given target X from one sensor and another target Y from another sensor and target X and target Y have the target feature vectors as below:

$$X = \{ 1., 1., 1., 1., 1., 1. \}$$
$$Y = \{ 1., 1., 1., 1., 1., 1. \}$$

by applying the Coefficient of Similarity Model (CSM), we have:

$$X \bullet X = \Sigma(Xi \,.\, Xi) = 6.0$$
$$X \bullet Y = \Sigma(Xi \,.\, Yi) = 6.0$$
$$Y \bullet Y = \Sigma(Yi \,.\, Yi) = 6.0$$

and by substitution, we have:

$$\begin{aligned} R_{XY} &= X \bullet Y \,/\, (X \bullet X - X \bullet Y + Y \bullet Y) \\ &= 6.0 \,/(6.0 - 6.0 + 6.0) \\ &= 1.0 \end{aligned}$$

one can conclude that target X and target Y are the same target

From the above simple simulated examples, one can see that the Coefficient of Similarity Model is simple and easy to integrate to any avionics software for solving the positive target identification problems.

(C) STATISTICAL PATTERN RECOGNITION TECHNOLOGY:

(1) Mathematical expression for the Bayesian Model

**Let Xi = (x1,x2, ……….xn) represent the unknown target feature vector
and T1, T2, T3,…. Ti be the targets
and Yi =(y1,y2,…….yn) be the target feature vector from
a knowledge data base.
question is:**

$$Xi \in Tj\ ?$$

to which target Tj, the unknown target Xi belongs?

**According to the Bayesian conditional probability theory [2],
the target probability can be formulated as below:**

$$Pr(Tk/Xi) = Pr(Tk) * Pr(Xi/Tk) / \sum\{ Pr(Tj)*Pr(Xi /Tj)$$

**where Pr(Tk) is the apriori probability of target Tk
Pr(Xi/Tj) is the probability of target Xi given
it is target Tj
Pr(Tk/Xi) is the posterior probability of target
Tk containing target Xi
and Xi is assumed as a multivariate normal distributed random variable,
and the target probability density function can be expressed as following:**

$$Pr(Xi/Tj) = (1.0/\{n\sqrt{2}* \prod\}*\{\Sigma j\})* \{ \exp(- 0.5* (Yj-Xi)^{T} * \Sigma j^{-1} * (Yj-Xi) \} \quad [3]$$

**where Σj^{-1} is the inverse of the covariance matrix for target Tj
$(Yj-Xi)^{T}$ is the transpose vector of (Yj-Xi)
Pr(Xi/Tj) is the multivariate normal density function
of target Xi given it is belong to target Tj**

(2) Properties of the Bayesian model:

<1> Pr(Tj) >= 0.0 for all j=1,2,3…….n
<2> $\sum$\{Pr(Tj/Xi)\} = 1.0 for all j= 1,2,3…n

(3) Decision rule:

**<1> If Pr(Tk/Xi) = Max{ Pr(Tj/Xi) } for all j =1,2,…..n
then Xi belongs to target Tk
that is unknown target Xi is positively identified.**
**<2> If Pr(Tk/Xi) =/ max{ Pr(Tj/Xi) } for all j=1,2,3…n
then $Xi \notin Tj$
that is unknown target Xi is not the same as target Tj**

(4) Equivalence decision:

The test statistic for a given unknown target can be expressed as follows:

$$Dj(Xi) = (Yj-Xi)^T * \Sigma j^{-1} * (Yj-Xi) \qquad [5]$$

where Xi = (x1,x2,x3,………xn) as the unknown target feature vector.
Yj= (y1,y2,y3………..yn) as the target feature vector from the knowledge data base.
Σj^{-1} is the inverse of the covariance matrix of target Tj.

<1> If Dk(Xi) = Min { Dj(Xi) } for all j=1,2,3,…n
then $Xi \in Tj$
that is, the unknown target Xi is positively identified as target Tj.

<2> If Dk(Xi) =/ Min { Dj(Xi) } for all j=1,2,3,…n
then $Xi \notin Tj$
That is, the unknown target Xi cannot identify as target Tj.

(5) Limitation of the Bayesian Model:

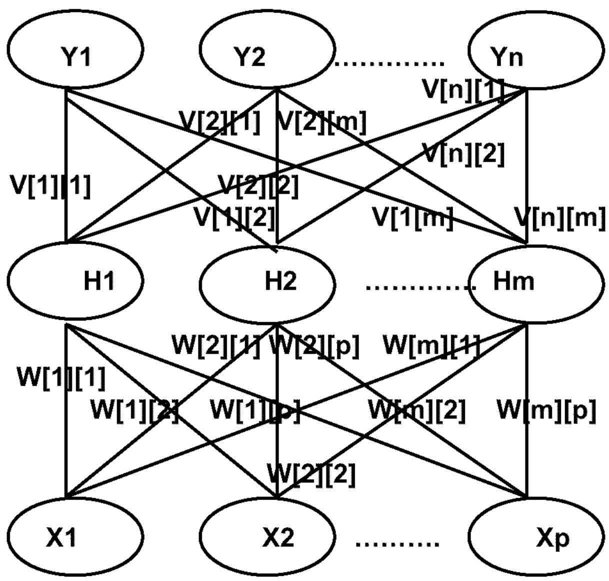
<1> Multivariate Normality assumption may not be true for all real time problems.
<2> Apriori probability of the unknown target is unknown most of the time.
<3> Feature element s in the target state vector is not easy to extract for the real time target.
<4> estimation of target probability is not easy, because the target probability density function is unknown most the time.

(D) NEURAL NETWORK MODEL :

<1> Architecture of the Backpropagation model:

Backpropagation (BP) model is one of the multi-layer perceptron (MLP) models and BP model is the work horse of Neural Network technology in field of object recognition and classification. BP model has three layers; input layer, hidden layer, and output layer. Input information is mapped to the output layer through the activation energy at the hidden layer, the errors of mapping are transmitted back to the input layer, the mapping is complete when the total error approaches approximately zero.

The architecture of the three layer Backpropagation model is represented below:



Where W[m][p] is the weight matrix between the input layer & the hidden layer.

V[n][m] is the weight matrix between the hidden layer & the output layer.

Xi is the feature element at the input layer.

Hi is the activation energy at the hidden layer and mathematically can be expressed as below:

$$H[i] = 1.0/\,[1.0 + \exp(-\,\mathrm{sum}[i])] \qquad [3]$$

Where $\mathrm{sum}[i] = \sum(W[i][j]*X[j])$

$i=1,2,\ldots m;\ j=1,2,\ldots p$

Yi is the target probability at the output layer and mathematically can be expressed as below:

$$Y[i] = \exp(-\mathrm{sum}[i]) \,/\, \{ \textstyle\sum \exp(-\,\mathrm{sum}[k]) ;k=1,2,..n\}$$

Where $\mathrm{sum}[i] = \sum(V[i][j] *H[j])$

$i=1,2,\ldots n;\ j=1,2,\ldots m$

<2> **Properties of the output function in the BP model**

(a) Y[i] >= 0.0 for all i.
(b) ΣY[i] = 1.0 for all i.

a, Show that Y[i] >= 0.0 for i=1,2,……n
Proof : since Y[i] = exp(-sum[i]) / Σ(exp(-sum[j]))
and sum[i] >=0.0 for i=1,2,….n
that is Y[i] >= 0.0 for i=1,2,…n

b, Show that Σ Y[i] = 1.0
Proof : since Y[i] = exp(-sum[i]) / Σ(exp(sum[j])) ------------(A)
By equation(A), we have:
for i= 1, Y[1] = exp(-sum[1]) / Σ(exp(-sum[j])) --------(1)
i= 2, Y[2] = exp(-sum[2]) / Σ(exp(-sum[j])) --------(2)
…………………………………………………..
i= n, Y[n] =exp(-sum[n]) / Σ(exp(-sum[j])) --------(n)
and
Σ Y[i] = Y[1] + Y[2] + Y[3] ……..Y[n] ------------(B)
Substituting equation(1)----(n) into equation (B), we have:
Σ Y[i] = exp(-sum[1]) / Σ(exp(-sum[j])) +
exp(-sum[2]) / Σ(exp(-sum[j])) +
-------------------------------------- +
exp(-sum[n]) / Σ(exp(-sum[j]))
= { exp(-sum[1]) + exp(-sum[2]) +..+ exp(-sum[n])}
/ Σ(exp(-sum[j])
= Σ(exp(-sum[j]) / Σ(exp(-sum[j])
= 1. 0
Therefore Σ Y[i] = 1.0

Since Y[i] satisfies the equations (a) and (b) , it is implied that Y[i] is a proper target probability density function at the output layer of the three layer Backpropagation (BP) model.

(3) Application

There are two different network models associated with the BP model for target Identification and classification problems. These are:

a. Training BP model

b. operational BP model

The training BP model and the operational BP model both have three layers, same number of elements in the input layer, hidden layer and output layers, that is, both have the same architecture but have different algorithm and different objective. The training model tries to teach the network to recognize the certain target, and estimate the optimal weight matrices for the network , that is, calculate the weight matrix between the input layer and the hidden layer, and the weight matrix between the hidden layer and the output layer. The objective of the operational BP model is to carry out the act of recognizing the unknown target with the weight matrices from the training BP model, sometimes, these weight matrices are referred as the memories of the network. Our experience has indicated that it takes a long time to train a network, but takes no time at all for the operational BP model to process the targets to be identified.

The equations associated with the training model are listed below:

<1> H[i] = 1.0 / { 1.0 - exp(- sum[i]}
and sum[i] = ∑W[i][j]*X[j] ; i = 1,2,…m, j=1,2,…p
H[j] is the activation energy at the hidden layer.

<2> Probability function at the output layer:
Y[i] = exp(-sum[i]) / ∑(exp(- sum[j])
where i=1,2,…n; j=1,2,…n

<3> Error at the output layer:
Delta[i] = { d[i] - Y[i] } * Y[i] * { 1.0 - Y[i]}
where d[i] is the desired probability vector
Y[i] is the estimated probability vector, i= 1,2,…..n
Delta[i] is the error vector at the output layer.

<4> Weight matrix between the hidden layer and the output layer:
$V[i][j]_{k+1} = V[i][j]_k + Nu * Delta[j] * Y[j]$
$+ Zeta * \{ V[i][j]_k - V[i][j]_{k-1} \}$
Where Nu is the learning rate
Zeta is the smoothing factor
V[i][j] is the weight matrix between the hidden layer and the output layer.

<5> Error at the hidden layer:
Beta[j] = H[j] * (1.0 - H[j]) * sum[j]
where sum[j] = ∑(Delta[i] * W[i][j]; i = 1,2,…p j =1,2,..m
Equations (1) ….(5) are the basic concepts to build the algorithm for the training network. The network is considered to be learned when the total error of the network at the output layer of the network approaches approximately zero.

The equations associated with the operational network are listed below:

<1> Activation energy at the hidden layer
<2> Weight matrix between the input layer and the hidden layer
<3> Weight matrix between the hidden layer and the output layer
<4> Probability function at the output layer

These equations listed above are the same as the equations in the training BP model.

The procedure to apply the operational BP model are expressed as follows:

<1> Apply the optimal weight matrices from the training BP model to initialize the weight matrices in the operational BP model.
<2> Apply the feature elements to the input layer of the operational BP model.
<3> Estimate the activation energy at the hidden layer.
<4> Calculate the probability vector at the output layer.
<5> The target with the maximum probability at the output layer is the target to be identified.

The rate of success depends on the optimal weight matrices from the training BP model. If the weight matrices or memories are optimal that is, the weight matrices are from the fully learned network, then the operational network will have a very good result.

(E) KNOWLEDGE DATA BASE :

(U) A knowledge data base was constructed from two published data set. From the original data set, we make some changes in the object name and feature elements, mainly to enhance our simulation purpose. The modified data set becomes the knowledge data base, and the six targets with feature elements are listed below:

Target #1

x1	x2	x3	x4
5.1 4.9 ---- 5.0	3.5 3.0 ---- 3.3	1.4 1.4 ----- 1.4	0.2 0.2 ---- 0.2

Target #2

x1	x2	x3	x4
7.0 6.4 ---- 5.7	3.2 3.2 ----- 2.8	4.7 4.5 ---- 4.1	1.4 1.5 ---- 1.3

Target #3

x1	x2	x3	x4
6.3 5.8 ---- 5.9	3.3 2.7 ---- 3.0	6.0 5.1 ---- 5.1	2.5 1.9 ----- 1.8

Target # 1 , Target #2 and Target #3 are constructed from the following reference:
Kendal, M. G. and Stuart, A.
“ ADVANCE THEORY OF STATISTICS”
VOL, Page 318
HAFNER, NEW YORK, 1966

(E) KNOWLEDGE DATA BASE : (CONTINUED)

Target #4

x1	x2	x3	x4
9.2 10.7 ---- 10.5	2.0 3.0 ---- 4.0	27.0 24.0 ----- 30.0	3.0 0.0 ------- 3.0

Target #5

x1	x2	x3	x4
14.4 14.7 ---- 4.7	4.0 7.0 ----- 3.0	34.0 33.0 ---- 29.0	4.0 2.0 --- 4.0

Target #6

x1	x2	x3	x4
18.6 17.4 ---- 19.4	7.0 6.0 ---- 9.0	42.0 47.0 ---- 51.0	3.0 5.0 ----- 1.0

Target # 4 , Target #5 and Target #6 are constructed from the following reference:
Tull, D. S. and Green, P. E.
“ RESEARCH FOR MARKETING DECISION”
Page 523-524
PRINTI_HALL, 1975

(F) Simulations :

The main objective of our simulation is to identify the simulated target from a knowledge data base by applying the data fusion technology, neural network technology and the the statistical pattern recognition technology.

We had built a knowledge data base, which contained six Targets, and each target with four feature elements; target #1, target #2 and target #3 each have fifty target feature vectors, target #4, target #5 and target #6 each have ten target feature vectors, for a total of 150 target feature vectors in the knowledge data base.

(1) Application of Data fusion technology:

a, Get one target feature vector (X) as the input vector.

b, pass X to the Coefficient of similarity model, from which the coefficient of similarity between (X) and all six targets are estimated.

c, The unknown target feature vector (X) will be identified as the target that has the maximum value of coefficient of similarity.

Repeat steps a, b and c until all the target feature vectors are Completely identified.

(2) Application of Statistical Pattern recognition technology:

a, Get one target feature vector (X) as the input vector.

b, Pass to the Bayesian model , from which the posterior probability of X with all six targets in the knowledge data base are estimated.

c, The unknown target feature vector (X) will be identified as the target that has the maximum value of posterior probability.

Repeat steps a, b, and c until all the target feature vectors are completely identified.

(3) Application of neural network technology:

a, Get one target from the knowledge data base as the input vector , pass to the training BP model, obtain the optimum weight matrices as the memories for that trained target

b, Pass the optimum weight matrices or memories of the trained target to initialized the operational BP model.

c, now get one target feature vector (X) as the input vector to the operational BP model, from which the activation energy and the output probabilities for all trained targets are estimated.

d, unknown target feature vector (X) will be identified as the target that has the maximum value of output probability from the operational BP model.

Repeat steps a, b, c, and d until all target feature vectors are positively identified.

(G) CONCLUSIONS :

(1) Theoretically, the Coefficient of Similarity Model is the most simple which compares the Back propagation model and the Bayesian model.

(2) The difficulty with the Bayesian model is the multivariate normality assumption, and the apriori probability estimation. Also the feature vectors of the target are difficult to extract from the real time multi-sensor data.

(3) The difficulty with the neural network model are:
  (a) Exact architecture for the target identification problem; the number of elements in the input layer are the number of elements of the feature vector, the number of elements in the output layer are the number of target to be trained , but the number of elements in hidden layer are very difficult to determine.
  (b) The information for the target to be trained are not easy to obtain in real time.

(4) The result of simulations are:

Data Fusion technique:	94.7% correctly identified.
Neural Network technique:	93.6% correctly identified.
Statistical Pattern recognition technique:	94.6% correctly identified

We conclude that Data Fusion technique is the winner for this particular application.

(5) For the real time application, more tests for the Coefficient of Similarity model with real time multi-sensor data are needed.

ACKNOWLEDGEMENT:

The authors wish to thank Mr. Roy Lecroy VP of Engineering, Mr. Al Whittaker Director of Avionics and Software, and Mr. Chuck Smith Jr. Manager Dept. 73-68, for their continued support and encouragement for the work described in this paper.

(H) REFERENCES:

[1] Jeun, Buddy H.,
"A MULTI-SENSOR INFORMATION FUCTION MODEL"
in the CISC-97 JOINT SERVICE COMBAT IDENTIFICATION SYSTEM CONFERENCE TECHNICAL PROCEDING ,
Volume I, Page 483
1997

[2] David L. Hall,
" MATHEMATICAL TECHNIQUES IN MULTISENSOR DATA FUSION"
Artech House,
1992

[3] Rumelhart, D.E., MCClelland, Jame L.,
" PARALLEL DISTRIBUTED PROCESSING: EXPLORATION IN MICROSTRUCTURE OF COGNITION", VOL. I , MIT PRESS
1968.

[4] (a) Kendal, M. G. and Stuart, A.
" ADVANCE THEORY OF STATISTICS"
VOL, Page 318
HAFNER, NEW YORK, 1966

(b) Tull, D. S. and Green, P. E.
" RESEARCH FOR MARKETING DECISION"
Page 523-524
PRINTI_HALL, 1975

[5] Jeun, Buddy H.,
Ph. D. DISERTATION, " AN IMPROVED MULTI-VARIETE CLASSIFICATION SCHAME",
COLLEGE OF ENGINEERING, UNIVERSITY OF MISSOURI,
1979 .