Applying a Modified Discrimination Model to Enhance Defense and Sensor Systems Security

BY
Dr. Buddy H. Jeun, Ph.D (Engineering)
And
John Younker, M.S (Engineering)

Sensor Fusion Technology, LLC
4522 Village Springs Run
Dunwoody, GA 30338
mail: jeunb@bellsouth.net
telephone 678-662-9556

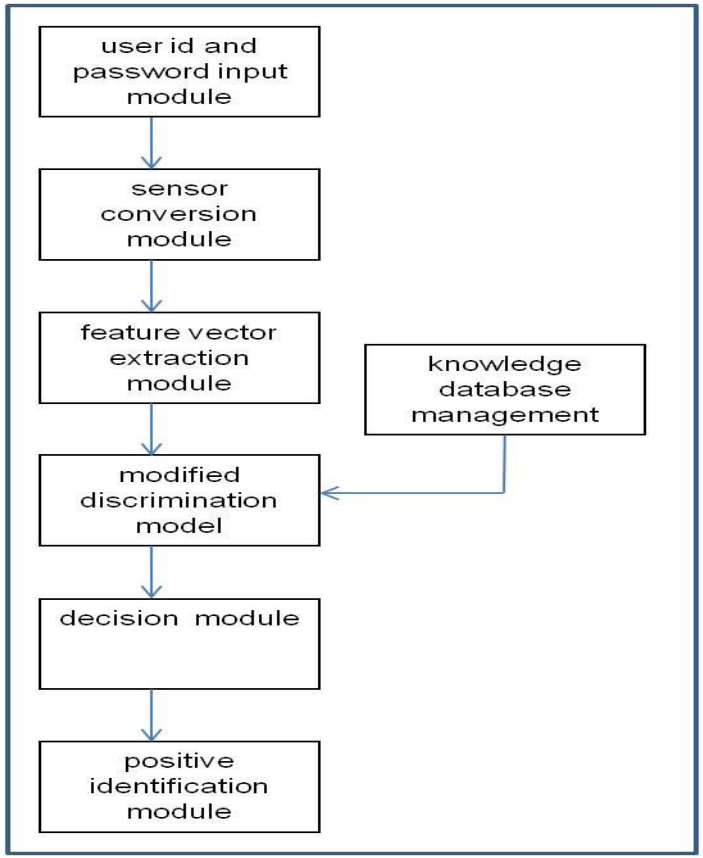
Contents

Introduction

- The primary objective of this paper to explore the application of a modified discrimination model to enhance defense and sensor systems security.
- The current method of protecting sensitive information is using user id and password as used in personal computers, institutional, governmental, and national defense systems. Correct user id and password is required to access secured information. However, user id and password can be falsified and hacked.
- The proposed application provides a technical solution to protect the user id and password method by employing a modified discrimination model to provide a positive verification of user id and password.

Introduction (continued)

The proposed architecture consists of seven modules:



Traditional Discrimination Model

- The traditional discrimination model has been known for a long time in the fields of science, engineering, and sociology. IBM's SPSS and UCLA's BMPD are commercial computer software packages that use the discrimination model for analysis.
- The Bayesian conditional probability theory is one of the classical discrimination models. Mathematically, the Bayesian probability model for the discrimination and classification application can be expressed as following:
- $$P\left(\frac{T_k}{S}\right) = \frac{\left\{P\left(\frac{S}{T_K}\right) \times P(T_K)\right\}}{\left\{\sum_{i=1}^{n} \left[P\left(\frac{S}{T_i}\right) \times P(T_i)\right]\right\}}$$

Traditional Discrimination Model (continued)

where:

$\boldsymbol{P}\left(\frac{T_k}{S}\right)$ is the probability of T_k given that T_k is in S

$$\boldsymbol{P}\left(\frac{S}{T_k}\right) = \frac{1}{(n \times \sqrt{2\pi})} \times e^{-\{(T-Y)\times(T-Y)^T\}}$$

$\boldsymbol{T} = (t1, t2, t3, \ldots tn)$ as feature vector in S

$\boldsymbol{Y} = (y1, y2, y3, \ldots yn)$ as feature vector in S

$\boldsymbol{P}(T_k)$ is the probability density function of T_k

In general, the Bayesian model produces a very accurate result for the unit variate normal assumption. However, when the feature vector is not of multi-variate distribution, the probability function becomes unknown. Therefore, the probability estimation will be complicated, and the result will not be useful. For that reason, the traditional discrimination model needs to be modified for this application.

Modified Discrimination Model

- Mathematical The expression for the modified discrimination model we are proposing is as follows:

$D(X,Y) = \{ (X – Y) * (X – Y)^T \}$ [Buddy H. Jeun, 1980]

Where

X = (x1, x2, x3, ……xn) is a feature vector

Y = (y1, y2, y3, ……yn) is a feature vector

- D(X, Y) is the distance between feature vector X and the feature vector Y.
- The properties of the model are:

Mathematically, the above properties are expressed as:

$0 \leq D (X, Y) \leq Delta$

where Delta is a positive number.

Modified Discrimination Model (continued)

- The possible decisions of the model are:
  - (1) If D(X,Y) = 0, then feature vector of X is equal to the feature vector of Y
  - For example suppose:
  - feature vector of X = (1, 1, 1, 1, 1, 1, 1, 1)
  - feature vector of Y = (1, 1, 1, 1, 1, 1, 1, 1)
  - Then, since $D(X, Y) = (X – Y)*(X – Y)^T = 0$, the feature vector of X is equal to the feature of Y
  - (2) Otherwise, if D(X,Y) is greater than zero, then the feature vector of X is not equal to the feature vector of Y.

Knowledge DataBase of User IDs and Passwords

- The knowledge database contains the true reference information for positive identification and classification of all system users.
- The user id can be the true full name, or social security number, or e-mail address. The password can be a series of decimal digits with special characters.

Knowledge DataBase of User IDs and Passwords (continued)

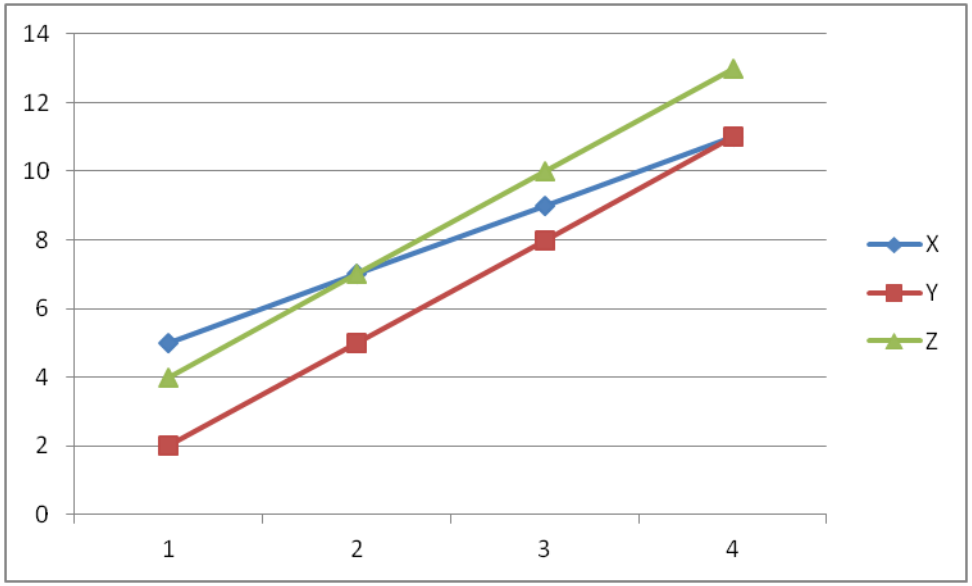
- example, suppose the knowledge database contains user id and password for three authorized persons X, Y and Z.
- The feature vector for X,Y and Z are:

X = {0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1}

Y = { 0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1 }

Z = { 0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1)

Knowledge DataBase of User IDs and Passwords (continued)



Comparison to Multi-Sensor Correlation Model

- Given feature vectors X and Y, the correlation coefficient of X and Y, can be expressed as follows: [Buddy H. Jeun, Alan Whittaker, 2002]

- $R (X, Y) = \{ (X \bullet Y) / ((X \bullet X) – (X \bullet Y) + (Y \bullet Y)) \}$

Where :

R (X, Y) is the correlation coefficient of X and Y

X = (x1, x2, x3,…….xn) as the feature vector X

Y = (y1, y2, y3,…….yn) as the feature vector of Y

X●X is the dot product of the feature vector of X and itself

X●Y is the dot product of the feature vector of X and the feature vector of Y

Y●Y is the dot product of the feature vector of Y and itself

Comparison to Multi-Sensor Correlation Model (continued)

- The properties of the multi-sensor correlation model are:

  R (X, Y) is **greater** than or equal to -1 and

  R (X, Y) is **less** than or equal to 1,

  That is mathematically:

  $$-1 \leq R(X,Y) \leq 1$$

Comparison to Multi-Sensor Correlation Model (continued)

- The decision rules of the multi-sensor correlation model are mathematically shown by two feature vectors X and Y as follows:

  X = (x1, x2, x3,……xn)

  Y = (y1, y2, y3,……yn)

  Now, if:

  R (X, Y) is **equal** to **1** then X is most likely **equal** to Y

  R (X, Y) is **less than 1** then X is most likely **not equal** to Y

Simulation and Verification of the Modified Discrimination Model

To demonstrate the application of the modified discrimination model, some simple mathematical simulated data is used in two cases considered as follows:

- **Case #1**

Consider person A whose feature vector of user id and password is the same as person X's feature vector stored in the Knowledge Database.

Mathematically:

A = (0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1)

X = (0, 1, 0, 1, 0, 1, 1, 1, 1, 0, 0, 1, 1, 0, 1, 1)

Y = (0, 0, 1, 0, 0, 1, 0, 1, 1, 0, 0, 0, 1, 0, 1, 1)

Z = (0, 1, 0, 0, 0, 1, 1, 1, 1, 0, 1, 0, 1, 1, 0, 1)

feature vector of A and feature vector of X, Y and Z into the modified discrimination model, we estimate the distances D(A, X), D(A, Y) and D(A, Z) as follows:

D(A, X) = 0

D(A, Y) = 5

D(A, Z) = 5

According to the decision rule of the modified discrimination model, D(A, X) is the smallest distance. Therefore, one can conclude that person A is positively identified as person X stored in the knowledge database. Person A should be allowed to access the secure information. Simulated data case #1 has demonstrated the power of positive identification.

Simulation and Verification of the modified Discrimination Model (continued)

- Verification for case #1
- Now the multi-sensor correlation model from the multi-sensor information fusion technology will be used to verify the accurate decision of modified discrimination model. This is done to prove that the feature vector of A is identically the same as the feature vector of X found in the secure knowledge database.

•Mathematically, the multi-sensor correlation model can be expressed as follows:

- $R (A, X) = \{ (A \bullet X) / ((A \bullet A) – (A \bullet X) + (X \bullet X)) \}$

Simulation and Verification of the Modified Discrimination Model (continued)

- Now substitute all of the dot products into the multi-sensor correlation model using the values given above we get:

(modified discrimination model) D (A, X) = 0

(multi-sensor correlation model) R (A, X) = 1

implying that the feature vector of A is positively identified as the feature vector of X.

Simulation and Verification of the Modified Discrimination Model (continued)

- **Case #2**

The feature vector of B as unknown person in terms of binary digits can be represented as follows:

$$B = (0, 1, 0, 0, 0, 1, 0, 1, 0, 1, 1, 0, 0, 1, 1, 1)$$

Using the feature vectors of B, X, Y, and Z and substituting into the modified discrimination model produces the following values:

$$D (B, X) = \{ (B – X) * (B – Y)^T \} = 9$$

$$D (B, Y) = \{ (B – Y) * (B – Y)^T \} = 8$$

$$D (B, Z) = \{ (B – Z) * (B – Z)^T \} = 12$$

Since none of D(B, X), D(B, Y) and D(B, Z) is zero, the feature vector of B does not pass the requirement set by the

modified discrimination model. Therefore, person B should **not** be allowed access to the secure information.

Simulation and Verification of the Modified Discrimination Model (continued)

•Verification for case #2

As in case #1, using the multi-sensor correlation model from multi-sensor information fusion technology we verify the accuracy of case #2 as follows:

$$R(B, X) = \{ (B \bullet X) / (B \bullet B) - (B \bullet X) + (X \bullet X) \} = 4$$

$$R(B, Y) = \{ (B \bullet Y) / (B \bullet B) - (B \bullet Y) + (Y \bullet Y) \} = 3$$

$$R(B, Z) = \{ (B \bullet Z) / (B \bullet B) - (B \bullet Z) + (Z \bullet Z) \} = 2$$

Since none of the coefficients of the feature vector of B and X, Y and Z is one, person B should be denied access to the secure information. This verifies the accuracy of the decision by the modified discrimination model for case #2.

Conclusions

- Simulated case #1, proves that the modified discrimination model can positively identify the secured person whose true user id and password is stored in the knowledge database. Therefore, those secured persons can be granted permission to access the secure information.

- Simulated case #2, proves that for those persons who are not registered in the knowledge database, the modified discrimination model can positively identify them and deny them access to the secure information.

- Since the modified discrimination model provides a powerful automatic system of positive identification, personal, institutional, governmental, and nationally secured information can be protected.

- Since the modified discrimination model converts the user id and password into a feature vector based on the multi-sensor information fusion technology, and the knowledge database stored and secured the feature vector of user id and password, therefore the modified discrimination model can minimize the risks from hacking.

- The modified discrimination model provides a computer algorithm that is easily implemented and embedded into the personal, institutional, governmental, and national security systems.

References

- [Geoffrey J. McLachlan, 1992], Discriminant Analysis and Statistical Pattern Recognition (Wiley series in probability and statistics)

- [Andrew R. Webb, Keith D. Copsey, 2011], Statistical Pattern Recognition, John Wiley & Son, New York

- [Buddy H. Jeun, 1980], The Design and Implementation of An Improved Multivariate Classification Scheme. Ph.D . Dissertation, Department of Electrical Engineering, University Of Missouri, Columbia, MO.

- [Buddy H. Jeun, Alan Whittaker, 2002] Multi-Sensor Information Technology Applied To The Development Of Smart Aircraft. 7th ICCRTS, The joint conference of the U.S. Department of Defense, and the Canadian Department of National Defense. Que'bec City, Canada.

- [Jeun, Younker, Hung, 2003] A Nuclear Plume Detection and Tracking Model For The Advanced Airborne Early Warning Surveillance Aircraft. 8th ICCRTS, Defense University, Washington, DC. June 17-19,2003.

- [K. Fukunaga, 1972] Introduction To Statistical Pattern Recognition, New York, Academic, 1972

- [P. A. Lacenbruch, 1975] Discrimination, Hefner Press