Risk Forecast Using Hidden Markov Models

Charles Pak
Nova Southeastern University
GSCIS, 3301 College Avenue,
Fort Lauderdale, FL 33314
cpak@nova.edu

Dr. James Cannady
Nova Southeastern University
GSCIS, 3301 College Avenue,
Fort Lauderdale, FL 33314
cannady@nova.edu

ABSTRACT

Today's fast moving technologies create innovative ideas, products, and services, but they also bring with them new security risks. The gap between new technologies and the security needed to keep them from opening up new risks in information systems (ISs) can be difficult to close completely. Changes in ISs are inevitable because computing environments, intentionally or unintentionally, are always changing. These changes bring with them vulnerabilities on new or existing ISs, which cause security states to move between mitigated, vulnerable, and compromised states. In previous work, we introduced the near real-time risk assessment using hidden Markov models (HMMs). This paper applies that theory to a prototype MatLab™ environment.

Categories and Subject Descriptors

H.4.3 Communications Applications; C.2.6 Internetworking

General Terms

Management, Algorithm, Security, and Theory

Keywords

Risk Assessment, Risk Analysis, Risk Management, Hidden Markov Models, MatLab™, Simulink®, Viterbi Algorithm.

1. INTRODUCTION

Dynamic ISs must be managed in order to survive in today's risk-filled computing environment. The first step in managing risk is to assess risk on mission critical assets. There are many different risk assessment methodologies available to assess risk. However, a risk assessment to quantitatively predict future risk is scarce. HMMs have been used in many scientific, manufacturing, and medical fields to model existing environments and to predict the probabilities of future events [2]. Applying a HMM to risk assessment to assess the risk levels of mission critical assets and to investigate the security state changes for each asset is a novel approach. Security states on mission critical assets change due to internal or external threats. By using a HMM on an organizational asset, the future probability of risk on a particular mission critical asset could be predicted. The application of HMMs has proven very practical in many scientific fields because of their probabilistic accuracy between two stochastic probability distributions: transition, and emission. The transition probability distribution represents the hidden transition probability that changes from one state to another and the emission probability distribution represents the observation probability. These doubly stochastic probability distributions represent the characteristics of HMMs.

Risk is defined as the probability of security incident and its impact [5]. Because each organizational asset can change its security state from mitigate, vulnerable, or compromised, these security states can represent the transition probability distribution. In an organizational asset, a security incident can cause the asset to malfunction or become inoperable, which represents the observable consequence of security incident—the HMM emission probability distribution. Therefore, applying the HMM to organizational assets can elicit a theory based, quantitative risk assessment.

2. ASSET PROTECTION

Knowledge of the kinds of assets an organization has and an understanding of internal and external threats would likely minimize an organization's security expenditures and losses. The near real-time risk assessment methodology provides a proactive risk assessment using an organization's most current IS information. In a dynamically changing IS environment, the organizational asset priorities change to meet business objectives. A risk assessment that can accurately reflect the current state of rapidly changing assets, technologies, and computing environments is critical to an organization's mission and will serve as a lifeline in today's fast moving digital age [5].

The following sections analyze security states on organizational mission critical assets—a database, a domain name system (DNS), a domain controller (DC), a web server, a file server, and an e-mail server—in order to assess risks. Each organization will have a different set of asset lists to assess, depending on the size of the organization. Some assets may change their security states dynamically while others may remain static. However, assessing mission critical assets in near real-time will help stakeholders understand the current level of security measures and plan a mitigation procedure. As protecting organizational assets is one of the highest priorities of mission objectives, near real-time risk assessment can add valuable security measures toward achieving an organization's total security objectives.

3. ASSET RISKS

3.1 Database Security States in a Threat Environment

In previous work [3, 4], we introduced the near real-time risk assessment methodology using HMMs. To model this methodology, MatLab™ Simulink® [1] was employed (Figure 1). Figure 1 illustrates a HMM automatic risk assessment based on inputs of two stochastic probability distribution element sets from the database server used in the model: a transition probability distribution matrix and an emission probability distribution matrix. Table 1 displays the database server's risk assessment matrix inputs. Both HMM inputs (transition and emission probability distribution matrices) were multiplied in a matrix calculation and produced an array of security states (mitigated, vulnerable, and compromised). This array of security states was then combined to produce the total risk level at a specific time interval. Thus, Figure 1 shows a HMM where two stochastic probability distribution elements were multiplied to produce HMM products. The transition probability distribution matrix represents the security incident probability; the emission probability distribution matrix represents the security impacts or consequences caused by the incident.

Both Figure 2 and Figure 3 show inputs to a HMM. Figure 2 is a transition probability distribution input, and Figure 3 is an emission probability distribution input. Both of these stochastic probability distribution matrix elements form the HMM's attributes in near real-time risk assessment.

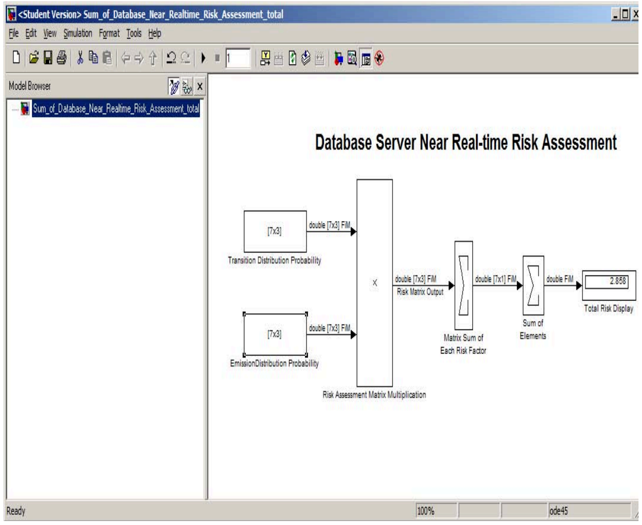
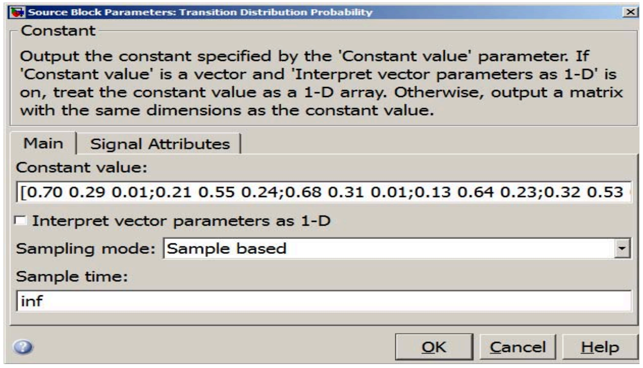

<Student Version> Sum_of_Database_Near_Realtime_Risk_Assessment_total
File Edit View Simulation Format Tools Help
1
Model Browser
Sum_of_Database_Near_Realtime_Risk_Assessment_total
Database Server Near Real-time Risk Assessment
[7x3]
double [7x3] FiM
Transition Distribution Probability
[7x3]
double [7x3] FiM
EmissionDistribution Probability
×
Risk Assessment Matrix Multiplication
double [7x3] FiM
Risk Matrix Output
Matrix Sum of Each Risk Factor
double [7x1] FiM
Sum of Elements
double FiM
2.858
Total Risk Display
Ready
100%
ode45

Figure 1. Near Real-time Risk Assessment on Database

As Figures 2 and 3 show, each matrix element forms a three-security state—mitigated, vulnerable, and compromised—probability ratio totaling one. Therefore, each set of matrix elements totals one at any single time interval. As time elapses, these matrix data elements vary in response to the threat environment, and thereby produce a dynamically changing risk assessment that reflects the changing security states at each monitoring time interval. Figure 2 illustrates a transition probability distribution input to the database server.


Source Block Parameters: Transition Distribution Probability

Constant

Output the constant specified by the 'Constant value' parameter. If 'Constant value' is a vector and 'Interpret vector parameters as 1-D' is on, treat the constant value as a 1-D array. Otherwise, output a matrix with the same dimensions as the constant value.

Main | Signal Attributes

Constant value:

[0.70 0.29 0.01;0.21 0.55 0.24;0.68 0.31 0.01;0.13 0.64 0.23;0.32 0.53

□ Interpret vector parameters as 1-D

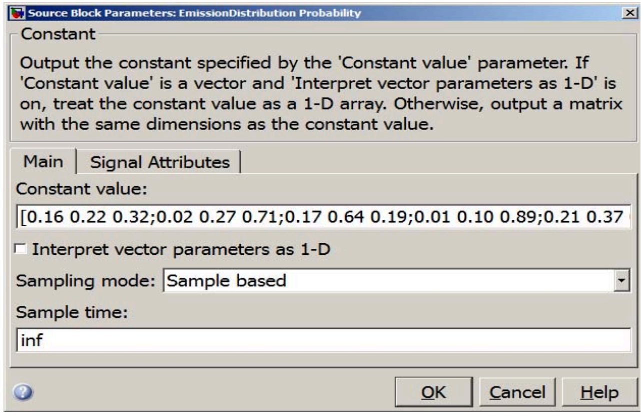
Sampling mode: Sample based

Sample time:

inf

Figure 2. Database Server Transition Probability Distribution Inputs to HMM

Similarly, Figure 3 displays an emission probability distribution in a three-security state matrix format. Each matrix element group represents mitigated, vulnerable, and compromised security states. In a HMM, the transition probability distribution represents security state changes for a mission critical asset. The emission probability distribution matrix data elements represent observed security states of the asset. In this HMM, both of these probability distribution inputs are multiplied in a matrix format to produce a near real-time risk assessment.


Source Block Parameters: EmissionDistribution Probability

Constant

Output the constant specified by the 'Constant value' parameter. If 'Constant value' is a vector and 'Interpret vector parameters as 1-D' is on, treat the constant value as a 1-D array. Otherwise, output a matrix with the same dimensions as the constant value.

Main | Signal Attributes

Constant value:

[0.16 0.22 0.32;0.02 0.27 0.71;0.17 0.64 0.19;0.01 0.10 0.89;0.21 0.37

□ Interpret vector parameters as 1-D

Sampling mode: Sample based

Sample time:

inf

Figure 3. Database Server Emission Probability Distribution Inputs to HMM

Table 1 lists two required HMM matrix input probability distributions and their resulting HMM matrix products. Table 1 also includes the resulting total risk assessment from the near real-time risk assessment design. The total risk is produced by combining each of the time interval matrix security state elements.

Table 1. Database Server Risk Assessment Matrix

	Database Server Risk Assessment Matrix
Transition Probability	[0.70 0.29 0.01;0.21 0.55 0.24;0.68 0.31 0.01;0.13 0.64 0.23;0.32 0.53 0.15;0.00 0.01 0.99;0.25 0.63 0.12]
Emission Probability	[0.16 0.22 0.32;0.02 0.27 0.71;0.17 0.64 0.19;0.01 0.10 0.89;0.21 0.37 0.42;0.00 0.01 0.99;0.13 0.64 0.23]
HMM Matrix Product	[0.1120 0.0638 0.0032;0.0042 0.1485 0.1704;0.1156 0.1984 0.0019;0.0013 0.0640 0.2047;0.0672 0.1961 0.0630;0.0000 0.0001 0.9801;0.0325 0.4032 0.0276]
Risk Assessment	[0.179;0.3231;0.3159;0.27;0.3263;0.9802;0.4633]
Total Risk	[2.8578]

The most critical asset in the simulation was the customer credit card database. The compromised security state peaks at 30 seconds (Figure 4). However, the database server is unaware of the security risk ramifications until the password is cracked and the system compromised. Figure 4 shows a security compromise at 30 seconds with a high-risk level. The customer credit card database is a mission critical asset, and necessary for any organization to conduct its business. In this simulation, the database's security state indicates the three security states at each threat incident time interval. However, it is not until 30 seconds that the attacker's intent is revealed, at which time the customer credit card database is breached. Other preparatory actions, such as port scanning, enumeration, and password cracking have led the attacker to this database breach.

Figure 4 clearly shows a peak in the compromised risk level at 30 seconds, while mitigated and vulnerable states during this time interval show relatively low risk levels.

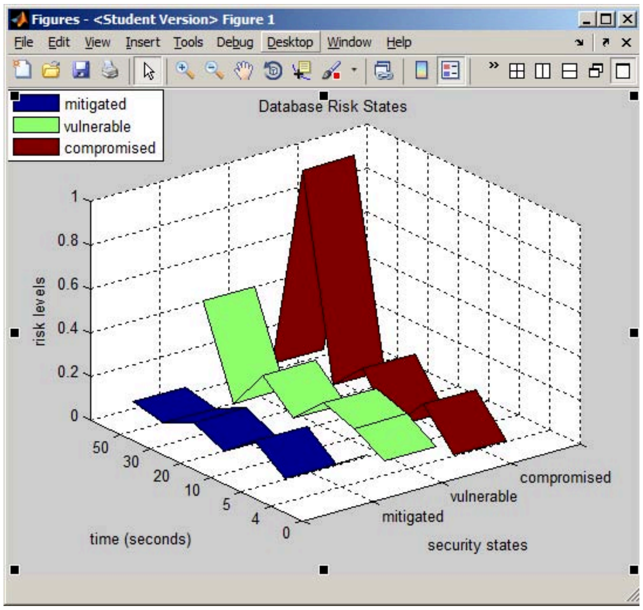


Figure 4. Database Server Security State Transition Trends

3.2 DNS Security States in a Threat Environment

The DNS server provides an integral service in resolving computer names. Its security states in a threat environment have been modeled, and Figure 5 shows the near real-time risk assessment on a DNS using HMMs. The server's mitigated, vulnerable, and compromised security states reflect induced threats at each time interval. This model used Simulink® blocks to represent the stochastic probability distribution matrix calculations.

The first two transition and emission probability distribution inputs have been multiplied in a matrix format to produce an array of matrix data elements which are combined to produce the risk assessment. Table 2 displays the probability distribution matrix inputs and the resulting HMM matrix products and risk assessments. The resulting HMM matrix products and risk assessment were produced in real-time as the two probability matrix inputs were inserted.

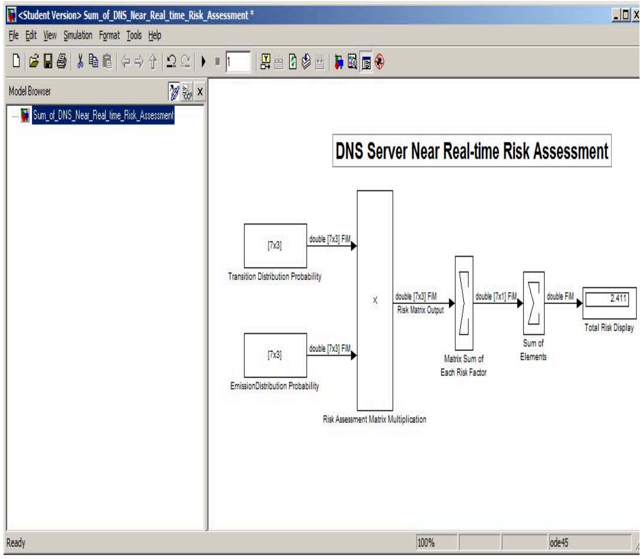

<Student Version> Sum_of_DNS_Near_Real_time_Risk_Assessment *
File Edit View Simulation Format Tools Help
Model Browser
Sum_of_DNS_Near_Real_time_Risk_Assessment
DNS Server Near Real-time Risk Assessment
[7x3]
double [7x3] FiM
Transition Distribution Probability
[7x3]
double [7x3] FiM
EmissionDistribution Probability
X
Risk Assessment Matrix Multiplication
double [7x3] FiM
Risk Matrix Output
Matrix Sum of
Each Risk Factor
double [7x1] FiM
Sum of
Elements
double FiM
2.411
Total Risk Display
Ready
100%
ode45

Figure 5. Near Real-time Risk Assessment on DNS

Table 2. DNS Server Risk Assessment Matrix

	DNS Server Risk Assessment Matrix
Transition Probability	[0.68 0.31 0.01;0.16 0.67 0.17;0.55 0.33 0.12;0.16 0.67 0.17;0.15 0.74 0.11;0.01 0.22 0.77;0.04 0.75 0.21]
Emission Probability	[0.23 0.35 0.42;0.07 0.24 0.69;0.23 0.40 0.37;0.01 0.14 0.85;0.11 0.20 0.69;0.00 0.04 0.96;0.01 0.21 0.78]
HMM Matrix Product	[0.1564 0.1085 0.0042;0.0112 0.1608 0.1173;0.1265 0.1320 0.0444;0.0016 0.0938 0.1445;0.0165 0.1480 0.0759;0.0000 0.0088 0.7392;0.0004 0.1575 0.1638]
Risk Assessment	[0.2691;0.2893;0.3029;0.2399;0.2404;0.748;0.3217]
Total Risk	[2.4113]

Figure 6 is a graph illustrating each security state matrix value. Each threat agent causes a fluctuation of security states at 4, 10, and 30 seconds in response to the threats in the network for which the DNS is providing services.

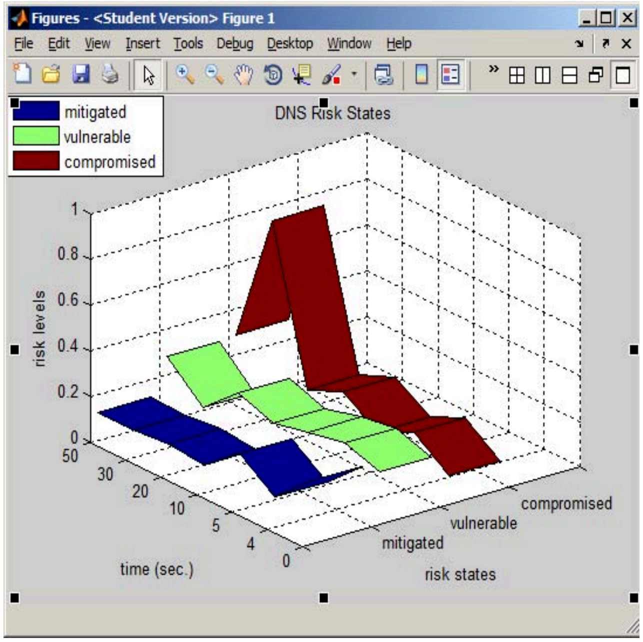


Figure 6. DNS Security State Transition Trends

Figure 6 shows the impact of each security state in comparison to other security states. The graph illustrates at what time the organization experienced the most risky DNS operation. Because each risk level is a matrix element, each element's values are added up to show the relative probability distribution to the accumulated total risk levels at each time interval.

Figure 6 also shows the trend of security state changes over time. Each ribbon represents a security state as a series of undulating states in time. The flow of each security state is compared to its security level. These data have been collected using a HMM in the MatLab™ modeling environment used in Figure 5. The Figure 5 near real-time risk assessment model was designed to calculate each asset's HMM values in a time series, and this data was used in Figure 6. The model in Figure 5 produced an automatic HMM output based on stochastic values of transition and emission probability distribution matrix inputs.

3.3 Domain Controller (DC) Security States in a Threat Environment

Figure 7 models the near real-time risk assessment on a simulated DC. The model is exactly identical to the DNS model; however, the required stochastic probability distribution inputs of this model are different. The two stochastic probability distribution matrix inputs used in Figure 7 are unique to this particular DC, and the stochastic probability distribution inputs will vary depending on the security environment this asset is in. The stochastic probability matrix inputs from Table 3 are inserted into the Figure 7 model to produce the near real-time risk assessment. The resulting HMM matrix product is shown in Table 3 along with its risk assessment.

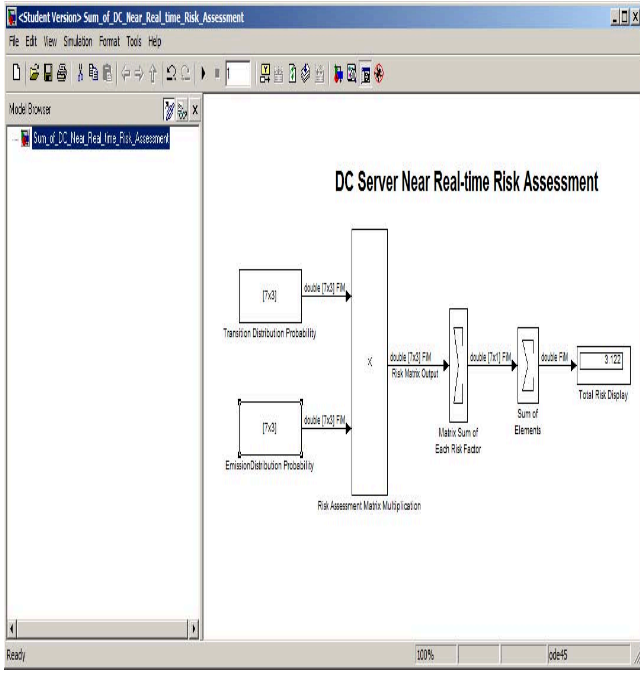

<Student Version> Sum_of_DC_Near_Real_time_Risk_Assessment
File Edit View Simulation Format Tools Help
Model Browser
Sum_of_DC_Near_Real_time_Risk_Assessment
DC Server Near Real-time Risk Assessment
[7x3]
double [7x3] FiM
Transition Distribution Probability
[7x3]
double [7x3] FiM
EmissionDistribution Probability
X
Risk Assessment Matrix Multiplication
double [7x3] FiM
Risk Matrix Output
Matrix Sum of
Each Risk Factor
double [7x1] FiM
Sum of
Elements
double FiM
3.122
Total Risk Display
Ready
100%
ode45

Figure 7. Near Real-time Risk Assessment on DC

Figure 8 illustrates the security state transitions of a DC. DCs maintain network accounts and credentials and must be contacted by client computers for user validation. There is, therefore, a constant client and server interaction. Figure 8 displays a similar pattern of risk levels as the DNS encountered at 4, 5, 10, 20, and 30 seconds. However, at 10 seconds, a noticeable compromised security state suggests a password crack. The database breach at 30 seconds was carried out with a compromised domain administrator password, and the authenticating domain controller is in a highly compromised security state. There is a noticeable trend of security state changes at each time interval. At 4 seconds, a highly vulnerable security state led to a compromised security state at 10 seconds. As the password crack completed its exploitation, the compromised security state lowered its security state and the vulnerable security state arose. As Figure 8 illustrates, the password crack was realized at 30 seconds. As the compromised security state arose at 30 seconds, mitigated and vulnerable states lowered their security states in the HMM probability.

Table 3. DC Server Risk Assessment Matrix

	DC Server Risk Assessment Matrix Using HMM
Transition Probability	[0.72 0.27 0.01;0.04 0.75 0.21;0.50 0.39 0.11;0.04 0.27 0.69;0.11 0.70 0.19;0.02 0.14 0.84;0.18 0.61 0.21]
Emission Probability	[0.25 0.30 0.45;0.11 0.34 0.55;0.26 0.41 0.33;0.02 0.04 0.94;0.32 0.33 0.35;0.00 0.02 0.98;0.05 0.32 0.63]
HMM Matrix Product	[0.1800 0.0810 0.0045;0.0044 0.2550 0.1155;0.1300 0.1599 0.0363;0.0008 0.0108 0.6486;0.0352 0.2310 0.0665;0.0000 0.0028 0.8282;0.0090 0.1952 0.1323]
Risk Assessment	[0.2655;0.3749;0.3262;0.6602;0.3327;0.826;0.3365]
Total Risk	[3.122]

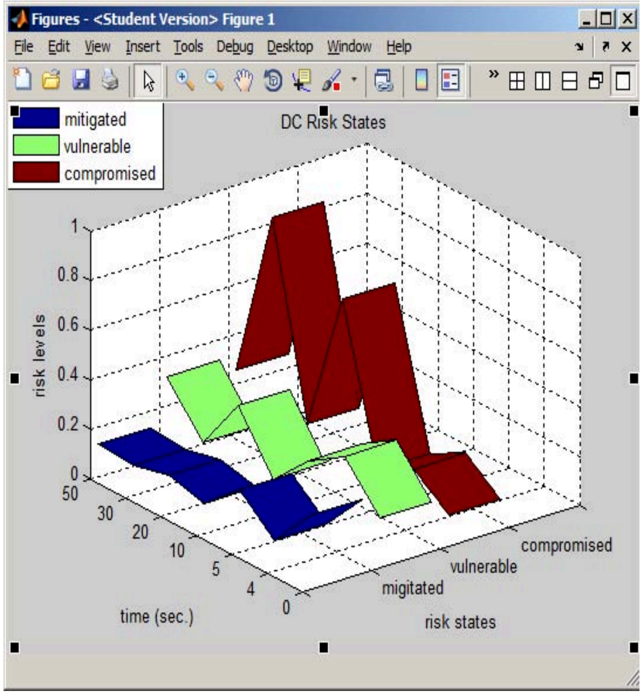


Figure 8. DC Security State Transition Trends

Figure 8 displays a similar pattern of security state changes as that of the DNS. However, because of the nature of DCs, the password crack incident triggered a compromised security state peak at 10 seconds. This suggests the password exploitation was against a domain administrator account maintained by the domain controller, and the impact of the threat was high on this asset. Thus, Figure 8 illustrates a DC security state trend that is consistent with the security states of the DNS server. The two compromised security state peaks at 10 and 30 seconds reflect the password crack and database breach exploitations in the network domain. Figure 8 shows the impacts of these security states in their proportional scale and security trends in time. One can easily see that the compromised security state has taken the major role due to the password exploitation. Each security state matrix value has been calculated using the HMM on each asset as recorded in Table 3.

3.4 Web Server Security States in a Threat Environment

Figure 10 illustrates the security state transitions on a simulated web server. The web server security state levels fluctuate noticeably at 4, 10, and 30 seconds. The web server might have provided the front-end web interfaces to customers, and might, thereby, be responsible for the attacker's access to the customer database at 30 seconds. The web server security state level displays a high risk level during the customer credit card database breach at 30 seconds. Although there is no direct exploitation on the web server, the threat level is raised in the internal network and the vulnerable security states are high until the database breach at 30 seconds at which time the compromised security state takes over the probability ratio.

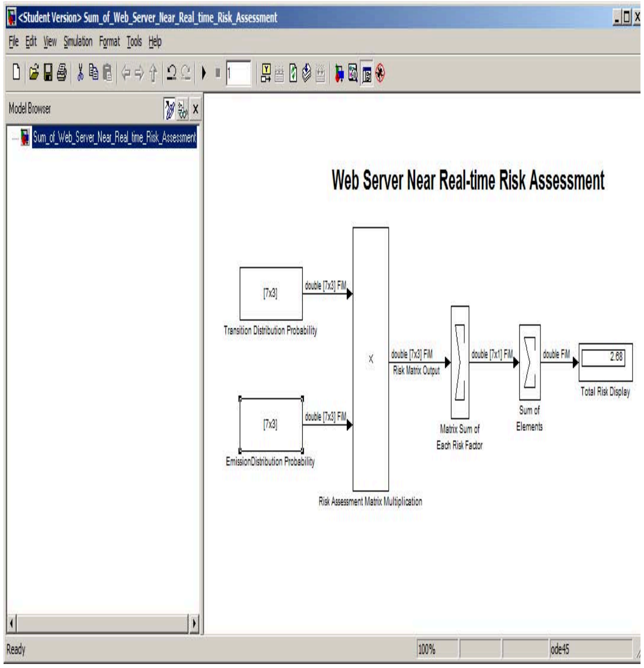

<Student Version> Sum_of_Web_Server_Near_Real_time_Risk_Assessment
File Edit View Simulation Format Tools Help
Model Browser
Sum_of_Web_Server_Near_Real_time_Risk_Assessment
Web Server Near Real-time Risk Assessment
[7x3]
double [7x3] FiM
Transition Distribution Probability
[7x3]
double [7x3] FiM
EmissionDistribution Probability
X
Risk Assessment Matrix Multiplication
double [7x3] FiM
Risk Matrix Output
Matrix Sum of Each Risk Factor
double [7x1] FiM
Sum of Elements
double FiM
2.68
Total Risk Display
Ready
100%
ode45

Figure 9. Total Near Real-time Risk Assessment on Web Server

Due to open port connections on the web server, its vulnerable state is relatively high throughout the assessment time frame. As Figure 10 shows, the vulnerable state displays proportionally high risk levels until the compromised state at 30 seconds at which time the compromised state raised its risk level.

Table 4 provides the relevant matrix data elements for the risk assessment in Figure 9. When plotted, these data elements form a similar pattern as that found in Figure 10. Figure 10 illustrates how a vulnerable state can lead to a compromised state given enough opportunity for the attacker to exploit the vulnerability. Several attempts may have been made until a vulnerable spectrum was found, at which time the final exploitation carried out the attack. As illustrated in Figure 10, a high vulnerable spectrum was displayed throughout the time intervals until the final exploitation at 30 seconds.

Table 4. Web Server Risk Assessment Matrix

	Web Server Risk Assessment Matrix Using HMMs
Transition Probability	[0.56 0.42 0.02;0.04 0.72 0.24;0.54 0.45 0.01;0.01 0.86 0.13;0.33 0.66 0.01;0.01 0.06 0.93;0.12 0.67 0.21]
Emission Probability	[0.23 0.33 0.44;0.19 0.39 0.42;0.14 0.36 0.52;0.10 0.24 0.66;0.11 0.35 0.54;0.01 0.11 0.88;0.10 0.39 0.51]
HMM Matrix Product	[0.1288 0.1386 0.0088;0.0076 0.2808 0.1008;0.0756 0.1620 0.0052;0.0010 0.2064 0.0858;0.0363 0.2310 0.0054;0.0001 0.0066 0.8184;0.0120 0.2613 0.1071]
Risk Assessment	[0.2762;0.3892;0.2428;0.2932;0.2727;0.8251;0.3804]
Total Risk	[2.4138]

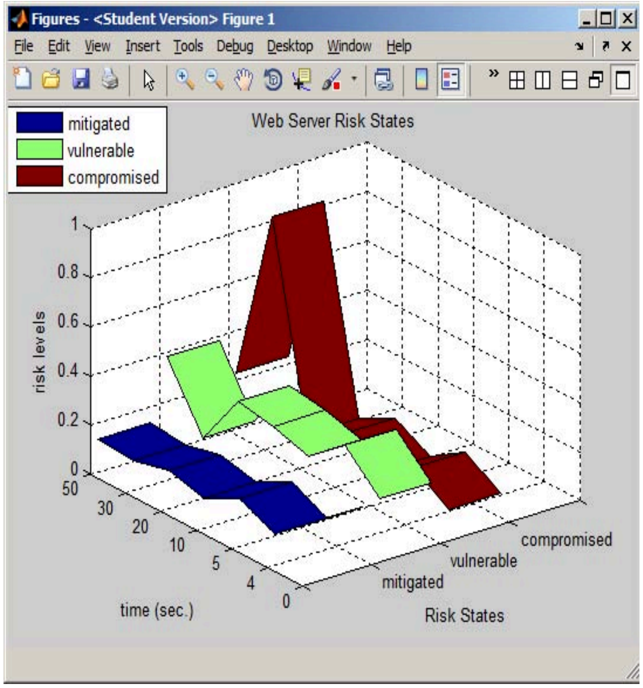


Figure 10. Web Security State Transition Trends

In Figure 10, the three security states on the web server lay out the landscape of risk states to show the security trends on the web server as time elapses. The web server could be the first target for the attacker to exploit due to its open ports on the server.

3.5 File Server Security States in a Threat Environment

File servers have relatively safe security states as they are passive servers providing file access to users, and are typically installed on internal networks. A file server may not be the first target of attacks; however, being a member server in a network, mitigated, vulnerable, and compromised security states follow a pattern similar to that of the web server. Figure 12 shows the file server's security states in response to the threats created on the network. The file server indicates low security activity although the security states fluctuate in response to the induced threats. Risk levels on each security state remain relatively low on this particular server.

Figure 11 displays the file server near real-time risk assessment model, which produces a real-time risk assessment based on the HMM probability inputs of a transition probability and an emission probability. The relevant matrix data elements are recorded in Table 5 along with the produced risk assessment.

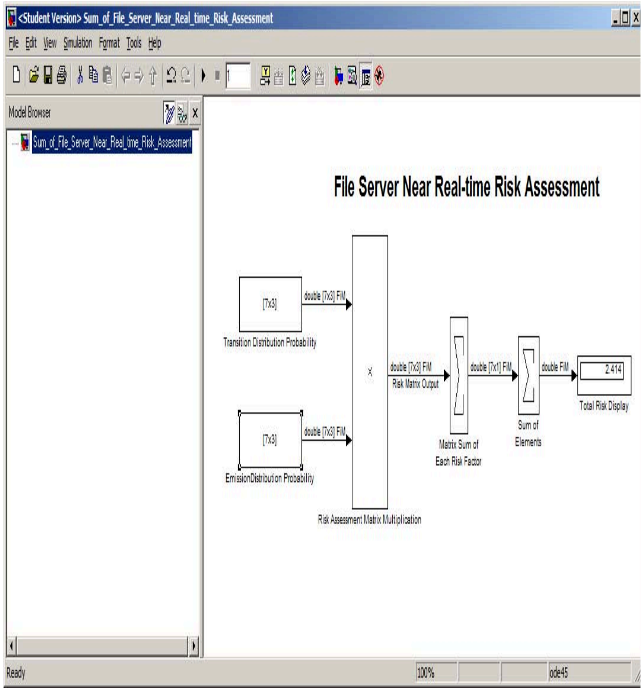

<Student Version> Sum_of_File_Server_Near_Real_time_Risk_Assessment
File Edit View Simulation Format Tools Help
Model Browser
Sum_of_File_Server_Near_Real_time_Risk_Assessment
File Server Near Real-time Risk Assessment
[7x3]
double [7x3] FIM
Transition Distribution Probability
[7x3]
double [7x3] FIM
EmissionDistribution Probability
X
double [7x3] FIM
Risk Matrix Output
Risk Assessment Matrix Multiplication
double [7x1] FIM
Matrix Sum of Each Risk Factor
double FIM
Sum of Elements
2.414
Total Risk Display
Ready
100%
ode45

Figure 11. Near Real-time Risk Assessment on File Server using HMMs

Table 5. File Server Risk Assessment Matrix

	File Server Risk Assessment Matrix Using HMMs
Transition Probability	[0.86 0.13 0.01;0.14 0.52 0.34;0.75 0.24 0.01;0.39 0.47 0.14;0.44 0.55 0.01;0.18 0.38 0.44;0.20 0.66 0.14]
Emission Probability	[0.31 0.34 0.35;0.08 0.22 0.70;0.31 0.62 0.17;0.03 0.31 0.64;0.39 0.31 0.30;0.02 0.26 0.72;0.24 0.36 0.40]
HMM Matrix Product	[0.2666 0.0442 0.0035;0.0112 0.1144 0.2380;0.2325 0.1488 0.0017;0.0117 0.1457 0.0896;0.1716 0.1705 0.0030;0.0036 0.0988 0.3168;0.0480 0.2376 0.0560]
Risk Assessment	[0.3143;0.3636;0.383;0.247;0.3451;0.4192;0.3416]
Total Risk	[2.4138]

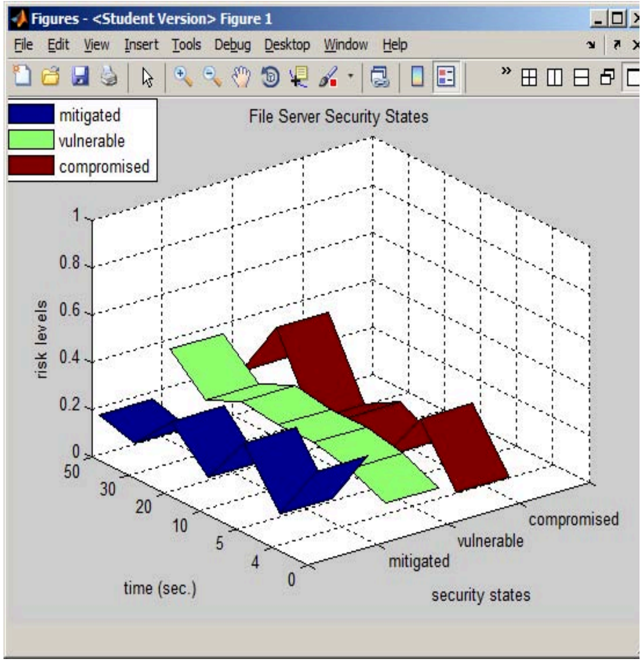


Figure 12. File Server Security State Transition Trends

Figure 12 shows the file server security state transition pattern. The level of security criticality is lower than the other observed assets. Notably, Figure 12 shows a similar pattern of security states as Figure 10, but with lower security risk levels. The security states of mitigated, vulnerable, and compromised remain relatively calm as time elapses although each threat incident shows higher risk levels at 4 and 30 seconds.

3.6 E-mail Server Security States in a Threat Environment

The security trends in Figure 14 show a similar pattern of security changes in the e-mail server as those found in previously discussed assets. The e-mail server interacts with other e-mail servers to deliver messages. The receiving e-mail server listens to a port 25 connection in order to be ready to establish a TCP session for simple mail transport protocol (SMTP) messages. Because the server must maintain open ports in order to send and receive e-mail messages, threats to e-mail servers are as prominent as those to web servers. Thus, Figure 14 shows a pattern of security states similar to those experienced by the web server.

Figure 13 illustrates a similar risk assessment model that produces a near real-time risk assessment based on two stochastic probability distribution data elements. The resulting risk assessment has been tabulated in Table 6, which produces the graph in Figure 14.

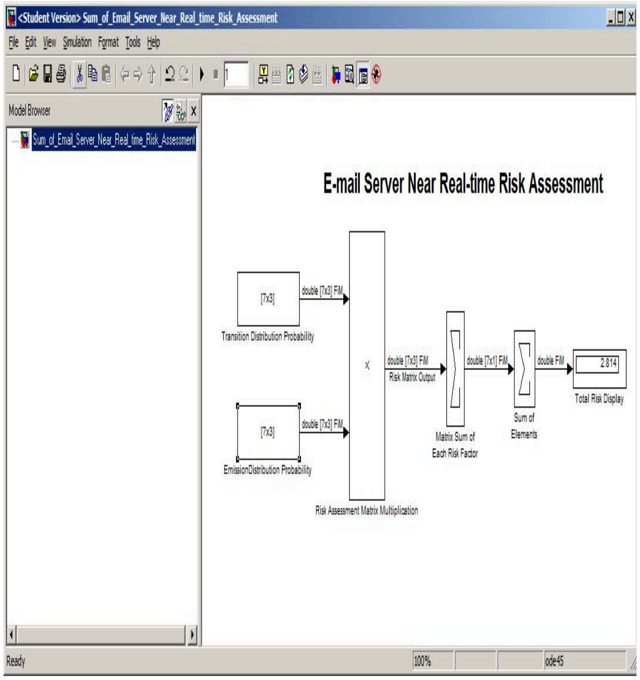

<Student Version> Sum_of_Email_Server_Near_Real_time_Risk_Assessment
File Edit View Simulation Format Tools Help
Model Browser
Sum_of_Email_Server_Near_Real_time_Risk_Assessment
E-mail Server Near Real-time Risk Assessment
[7x3]
double [7x3] FiM
Transition Distribution Probability
X
double [7x3] FiM
Risk Matrix Output
double [7x1] FiM
double FiM
2.814
Total Risk Display
[7x3]
double [7x3] FiM
EmissionDistribution Probability
Matrix Sum of
Each Risk Factor
Sum of
Elements
Risk Assessment Matrix Multiplication
Ready
100%
ode45

Figure 13. Near Real-time Risk Assessment on E-mail Server using HMMs

Table 6. E-mail Server Risk Assessment Matrix

	E-mail Server Risk Assessment Matrix Using HMMs
Transition Probability	[0.60 0.33 0.07;0.04 0.69 0.27;0.54 0.45 0.01;0.24 0.59 0.17;0.23 0.66 0.11;0.02 0.14 0.84;0.02 0.87 0.11]
Emission Probability	[0.22 0.36 0.42;0.09 0.17 0.74;0.38 0.45 0.17;0.06 0.09 0.85;0.26 0.35 0.39;0.01 0.10 0.89;0.10 0.52 0.38]
HMM Matrix Product	[0.1320 0.1180 0.0294;0.0036 0.1173 0.1998;0.2052 0.2025 0.0017;0.0144 0.0531 0.1445;0.0598 0.2310 0.0429;0.0002 0.0140 0.7476;0.0020 0.4524 0.0418]
Risk Assessment	[0.2802;0.3207;0.4094;0.212;0.3337;0.7618;0.4962]
Total Risk	[2.814]

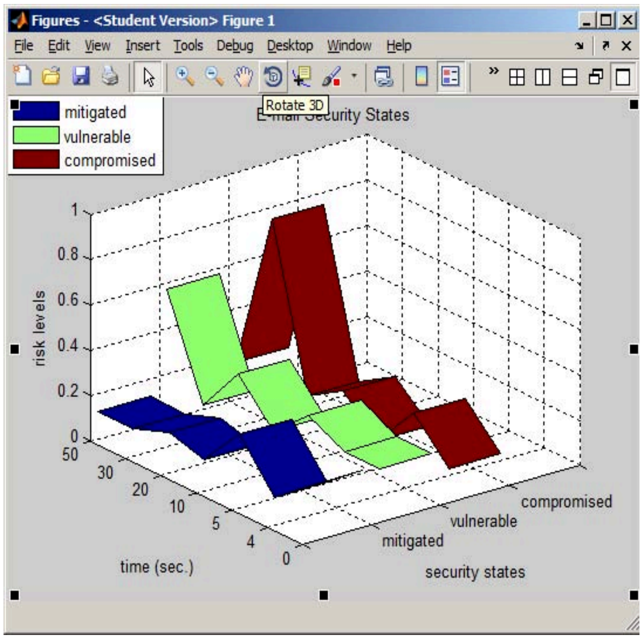


Figure 14. E-mail Server Security State Transition Trends

4. RISK ASSESSMENT ON ASSETS

4.1 Threat Analysis on Organizational Assets

Different threats cause different security states. Depending on the attacker's motives and targets, organizational assets experience different threats. Port scanning, enumeration, and password cracks can lead to an attacker's ultimate goal of a customer database breach. Further, organizational assets may experience different levels of security states over time. Close monitoring and an understanding of organizational asset security state changes as the threat event unfolds can lead to an accurate near real-time risk assessment.

When an administrator password is compromised, the entire network is at risk. Figure 12 illustrates risk transitions on the file server caused by an administrator account being compromised. An intruder who impersonates an administrator cannot only damage valuable data on the file server but in the entire network. For this reason, the risk level indicates a high level of risk during the password policy enumeration, the administrator password crack, and the customer credit card database breach.

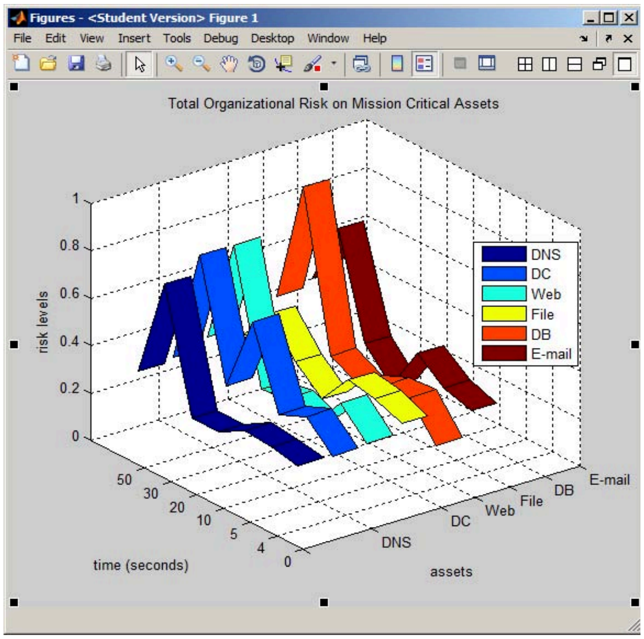


Figure 15. Organizational Risk Levels on Mission Critical Assets

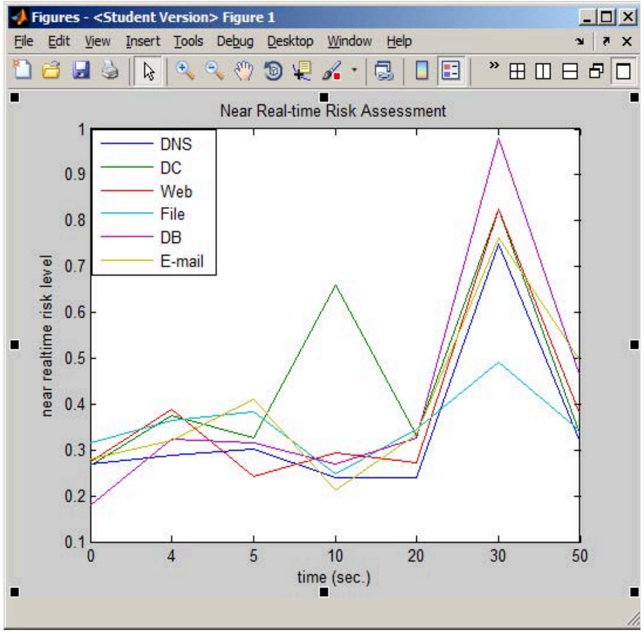


Figure 16. Risk Changes on Mission Critical Assets

Figure 4 illustrates a security state transition on the database server. High levels of risk activity at 4, 10, and 30 seconds indicate an active penetration toward the customer credit card database. The database server has the highest risk levels of all the organizational assets being assessed. While the database server has shown high security risk levels during the network penetration time intervals, Figure 14 shows that the e-mail server has a similar pattern. The e-mail server is one of the mostly utilized application servers in an organization. It leaves its TCP port connection open at all times in order to send and receive e-mail messages. Its total risk level is relatively high during the network penetration period and indicates high risk levels at 4, 10, and 30 seconds.

As system administrators strive to remove vulnerabilities and mitigate the risk on each network resource, management could review security state changes and provide stakeholders with a picture of the security state of each mission critical asset in near real-time.

The organizational mission critical asset risk assessment is best illustrated by Figure 15. The near real-time risk assessment presents a quantitative risk assessment on each asset over time, and shows the relative risk levels to each mission critical asset. The risk assessment data table has been plotted in Figure 16 line graph. Figure 15 and 16 provide a comprehensive picture of risk for the organization to monitor, and to help management decide on the best mitigation plan.

4.2 Near Real-time Risk Assessment on Database

The prototype simulation lab environment was created to model a near real-time risk assessment in MatLab™ (Figure 17). Two HMM stochastic probability matrices were inserted into the HMM simulation blocks which produced a matrix product of two HMM stochastic probability distributions. The resulting matrix multiplication produced an array of raw matrix elements. The numbers for each matrix element were added together to produce a series of HMM security risk factors through a time interval. The risk assessment detailed in Figure 1 is a series of risk assessment matrix elements shown over seven discrete time intervals (0, 4, 5, 10, 20, 30, and 50 seconds). The first right-side inner block in Figure 17 displays the sum of risk matrix elements at 0 seconds. The second block of the risk assessment shows the sum of risk matrix elements at 4 seconds. There are seven time intervals in this risk assessment process. Each time interval produces a risk assessment value in the HMM as shown in the last blocks of Figure 17. Figure 17 thus shows a near real-time risk assessment of all seven time intervals for each induced threat to the database asset.

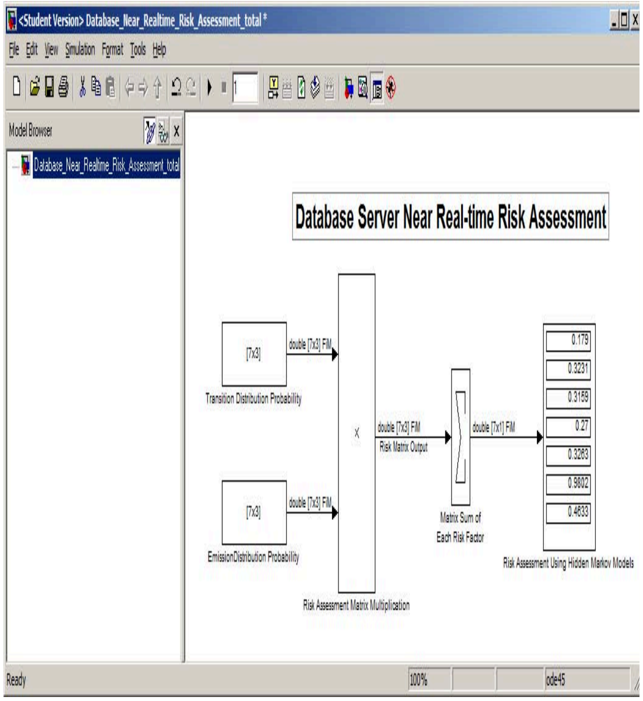

<Student Version> Database_Near_Realtime_Risk_Assessment_total *
File Edit View Simulation Format Tools Help
Model Browser
Database_Near_Realtime_Risk_Assessment_total
Database Server Near Real-time Risk Assessment
[7x3]
double [7x3] FIM
Transition Distribution Probability
[7x3]
double [7x3] FIM
EmissionDistribution Probability
X
Risk Assessment Matrix Multiplication
double [7x3] FIM
Risk Matrix Output
Matrix Sum of
Each Risk Factor
double [7x1] FIM
0.179
0.3231
0.3159
0.27
0.3263
0.9802
0.4633
Risk Assessment Using Hidden Markov Models
Ready
100%
ode45

Figure 17. Validation of Risk Assessment on Database

As a doubly stochastic probability product model in a HMM, the near real-time risk assessment accepts two stochastic probabilities: one as a transitional probability distribution and another as an emission probability distribution as defined by the HMM. Two stochastic probability matrices are inserted into the risk assessment matrix blocks in a MatLab™ simulation environment. The risk assessment output matrix elements are further calculated to produce the final near real-time risk assessment on the database as displayed in Figure 1. Thus, Figure 1 illustrates the near real-time risk assessment on the database, and Figure 17 shows that the assessment took place over a series of seven time intervals. This model was then used for the remaining mission critical assets to produce the entire organizational risk assessment as displayed in Figure 15 and 16.

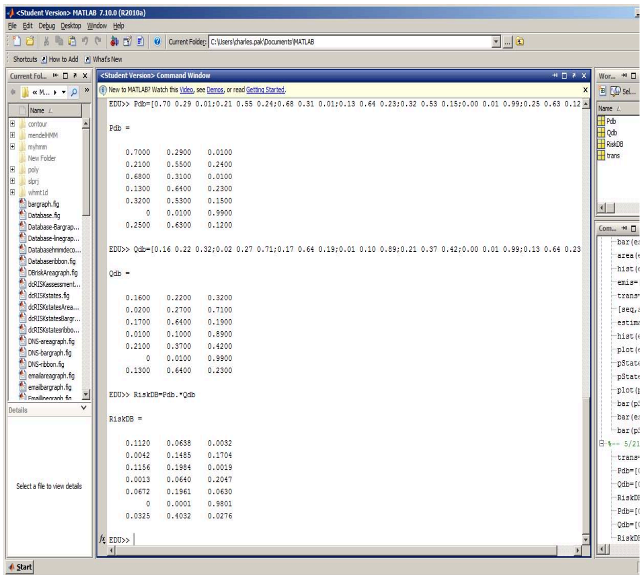
4.3 Validation of Risk Assessment on Database

Each transition probability distribution matrix input was inserted into the HMM in a 7x3 matrix format. Each emission probability distribution matrix was also entered into the HMM block in a 7x3 matrix format. Once the two HMM probability distribution matrices were entered in the HMM, a stochastic probability distribution output was produced as a sum of each risk factor matrix element value, thus creating the risk assessment value at each time interval as a HMM.

Although the simulation produced an automatic near real-time risk assessment, a manual risk assessment using HMMs was also conducted to validate the near real-time risk assessment. The following scripts were produced in a MatLab™ programming environment. Several programming variables were created in the scripts for this validation. The Pdb variable stores the HMM transition probability matrix elements, and the Qdb variable stores the HMM emission probability matrix elements. In terms of risk assessment, Pdb represents the probability of a security incident on a database asset, and Qdb represents the probability of security impacts or costs caused by the incident. The product of these two matrices produced the near real-time risk assessment in a matrix format that included all associated time intervals. Each row element of the matrix format included the probability distribution of mitigated, vulnerable, and compromised over time. Each row matrix elements were then combined to produce the ultimate risk level during any specific time interval. The following scripts represent the actual near real-time calculation in manual mode to validate the above MatLab™ simulation design. Figure 18 is the actual MatLab™ scripting window that displays the risk assessment calculation and validates the automatic risk assessment process.

```
EDU>> Pdb=[0.70 0.29 0.01;0.21 0.55 0.24;0.68 0.31 0.01;0.13
0.64 0.23;0.32 0.53 0.15;0.00 0.01 0.99;0.25 0.63 0.12]

Pdb =

    0.7000    0.2900    0.0100
    0.2100    0.5500    0.2400
    0.6800    0.3100    0.0100
    0.1300    0.6400    0.2300
    0.3200    0.5300    0.1500
         0    0.0100    0.9900
    0.2500    0.6300    0.1200
```

```
EDU>> Qdb=[0.16 0.22 0.32;0.02 0.27 0.71;0.17 0.64 0.19;0.01
0.10 0.89;0.21 0.37 0.42;0.00 0.01 0.99;0.13 0.64 0.23]

Qdb =

    0.1600    0.2200    0.3200
    0.0200    0.2700    0.7100
    0.1700    0.6400    0.1900
    0.0100    0.1000    0.8900
    0.2100    0.3700    0.4200
         0    0.0100    0.9900
    0.1300    0.6400    0.2300

EDU>> RiskDB=Pdb.*Qdb

RiskDB =

    0.1120    0.0638    0.0032
    0.0042    0.1485    0.1704
    0.1156    0.1984    0.0019
    0.0013    0.0640    0.2047
    0.0672    0.1961    0.0630
         0    0.0001    0.9801
    0.0325    0.4032    0.0276
```



```
EDU>> Pdb=[0.70 0.29 0.01;0.21 0.55 0.24;0.68 0.31 0.01;0.13 0.64 0.23;0.32 0.53 0.15;0.00 0.01 0.99;0.25 0.63 0.12

Pdb =

    0.7000    0.2900    0.0100
    0.2100    0.5500    0.2400
    0.6800    0.3100    0.0100
    0.1300    0.6400    0.2300
    0.3200    0.5300    0.1500
         0    0.0100    0.9900
    0.2500    0.6300    0.1200

EDU>> Qdb=[0.16 0.22 0.32;0.02 0.27 0.71;0.17 0.64 0.19;0.01 0.10 0.89;0.21 0.37 0.42;0.00 0.01 0.99;0.13 0.64 0.23

Qdb =

    0.1600    0.2200    0.3200
    0.0200    0.2700    0.7100
    0.1700    0.6400    0.1900
    0.0100    0.1000    0.8900
    0.2100    0.3700    0.4200
         0    0.0100    0.9900
    0.1300    0.6400    0.2300

EDU>> RiskDB=Pdb.*Qdb

RiskDB =

    0.1120    0.0638    0.0032
    0.0042    0.1485    0.1704
    0.1156    0.1984    0.0019
    0.0013    0.0640    0.2047
    0.0672    0.1961    0.0630
         0    0.0001    0.9801
    0.0325    0.4032    0.0276

fx EDU>>
```

Figure 18. Validation of Risk Assessment on Database

4.4 MatLab™ Modeling Environment

The sum of risk matrix elements in Figure 18 validates the MatLab™ modeling prototype results. Other matrix element-wise calculations were validated in a similar manner. As illustrated above, the MatLab™ prototype produced the near real-time risk assessment matrix elements using a HMM. Each asset's security states were calculated, and their aggregated risk assessment was produced in the model. The MatLab™ model was efficient in calculating the HMM matrix elements to produce the resulting risk assessment. MatLab™ was especially useful in calculating the matrix element-wise products that combined the probability of a security incident and the impact of such an incident as a doubly stochastic probability distribution calculation. As each set of doubly stochastic probability distribution ratios were inserted into the HMM, the product of these two probability matrices was produced in an element-wise matrix calculation. To produce the risk assessment, these raw matrix elements are added together to correctly arrange the matrix element calculation that produced the overall risk assessment. Both manual calculation and the automatic modeling produced identical results and validated the near real-time risk assessment.

Transition and emission probability distributions were input matrix elements for the HMM model. As the HMM matrix products were calculated, further analysis was required to produce the near real-time risk assessment as displayed in Table 1. The HMM matrix product included three security state probability distributions (mitigated, vulnerable, and compromised) at each time interval. Seven time intervals were introduced into the simulation environment, and seven sets of matrix elements were produced. As designed, the near real-time risk assessment produced the near real-time risk assessment of seven sets of matrix elements. These were then combined to produce the total risk assessment (Figure 1). Table 1 shows the outputs of each stage of the near real-time risk assessment. As illustrated above, the HMMs can be applied to risk assessments to calculate the risk factors of organizational assets. Each dynamically changing security state of assets are characterized by the HMM's transition probability distribution, and the observable security states by the HMM's emission probability distribution. The matrix product of these two doubly stochastic probability distributions produced the ultimate near real-time risk assessment. As HMMs have been applied to many scientific fields, the HMMs have proven very practical and efficient in assessing risks.

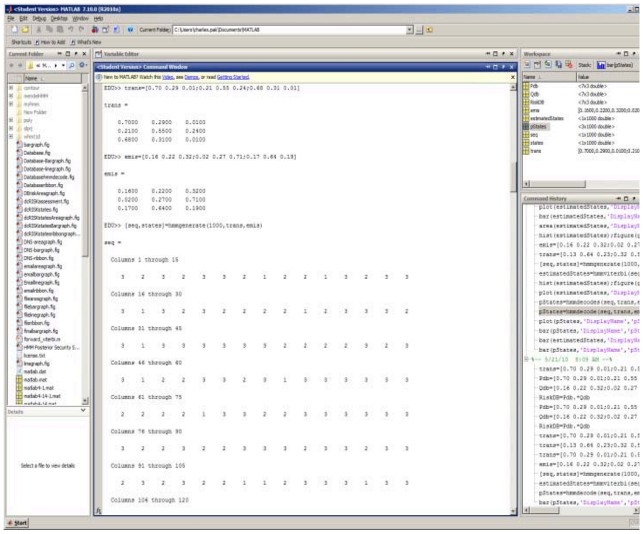
Figure 18 shows the MatLab™ programming environment in which the manual risk assessment validated the near real-time automated risk assessment outputs. While Figure 18 shows the database risk assessment output in a matrix format, Figure 15 and 16 show a total picture of mission critical asset risk assessments for a specific risk assessment period. Select mission critical assets are assessed by plotting their risk factors in response to threat levels in a near real-time manner. The underlying risk assessment was based on a HMM. The output of the near real-time model was plotted for each mission critical asset for each time interval. The plotted graphs reveal some important facts for management to consider in their risk mitigation process as the risk assessment continues over time.

Figure 15 and 16 present near real-time monitoring graphs where each asset's risk levels can be compared with other mission critical assets. These graphs can serve in a decision making process as management determines which assets need to be protected first based on risk levels and asset priorities. For example, at 30 seconds the database risk level is at the highest risk level and this would alert management to mitigate the risk. This type of near real-time risk assessment can provide a clear and timely risk assessment to aid management in planning for mitigation.

4.5 HMM Posterior Security State Probabilities

Given the HMM used in this simulation, the posterior security state probability distribution can be calculated. Because the calculation expected a square matrix format, the time interval has been split into three time intervals to match the three security states. In addition, the Viterbi Algorithm [2] in the following HMM scripts calculates the most probable path in a given HMM with certain specific transition and emission probability distribution matrices. The following HMM posterior probability distribution was calculated using over 1,000 sequential sample populations over each measuring time interval (Figure 19).

```
EDU>> trans=[0.70 0.29 0.01;0.21 0.55 0.24;0.68 0.31 0.01]
trans =
   0.7000    0.2900    0.0100
   0.2100    0.5500    0.2400
   0.6800    0.3100    0.0100
EDU>> emis=[0.16 0.22 0.32;0.02 0.27 0.71;0.17 0.64 0.19]
emis =
   0.1600    0.2200    0.3200
   0.0200    0.2700    0.7100
   0.1700    0.6400    0.1900
EDU>> [seq,states]=hmmgenerate(1000,trans,emis);
EDU>> estimatedStates=hmmviterbi(seq,trans,emis);
EDU>> pStates=hmmdecode(seq,trans,emis);
EDU>> bar(pStates,'DisplayName','pStates');figure(gcf)
EDU>> trans=[0.13 0.64 0.23;0.32 0.53 0.15;0.00 0.01 0.99]
trans =
   0.1300    0.6400    0.2300
   0.3200    0.5300    0.1500
        0    0.0100    0.9900
EDU>> emis=[0.01 0.10 0.89;0.21 0.37 0.42;0.00 0.01 0.99]
emis =
   0.0100    0.1000    0.8900
   0.2100    0.3700    0.4200
        0    0.0100    0.9900
EDU>> [seq,states]=hmmgenerate(1000,trans,emis);
EDU>> estimatedStates=hmmviterbi(seq,trans,emis);
EDU>> pStates=hmmdecode(seq,trans,emis);
EDU>> bar(pStates,'DisplayName','pStates');figure(gcf)
```


<Student Version> MATLAB 7.10.0 (R2010a)
Current Folder: C:\Users\charles.pak\Documents\MATLAB
Variable Editor
<Student Version> Command Window
New to MATLAB? Watch this Video, see Demos, or read Getting Started.
Workspace
Command History

```
EDU>> trans=[0.70 0.29 0.01;0.21 0.55 0.24;0.68 0.31 0.01]

trans =

    0.7000    0.2900    0.0100
    0.2100    0.5500    0.2400
    0.6800    0.3100    0.0100

EDU>> emis=[0.16 0.22 0.32;0.02 0.27 0.71;0.17 0.64 0.19]

emis =

    0.1600    0.2200    0.3200
    0.0200    0.2700    0.7100
    0.1700    0.6400    0.1900

EDU>> [seq,states]=hmmgenerate(1000,trans,emis)

seq =

  Columns 1 through 15

     3     2     3     2     3     3     2     1     2     2     1     3     2     3     3

  Columns 16 through 30

     3     1     3     2     3     2     2     2     2     1     2     3     3     3     2

  Columns 31 through 45

     3     1     3     3     3     3     3     3     2     2     2     2     3     2     3

  Columns 46 through 60

     3     1     2     2     3     3     2     3     1     3     3     3     3     3     3

  Columns 61 through 75

     2     2     2     2     1     3     3     2     2     3     3     3     3     3     3

  Columns 76 through 90

     3     2     2     3     2     2     3     3     3     2     3     3     2     3     3

  Columns 91 through 105

     2     3     2     3     2     2     1     1     2     3     3     3     1     3     3

  Columns 106 through 120
```

Figure 19. HMM Posterior Security State Probabilities

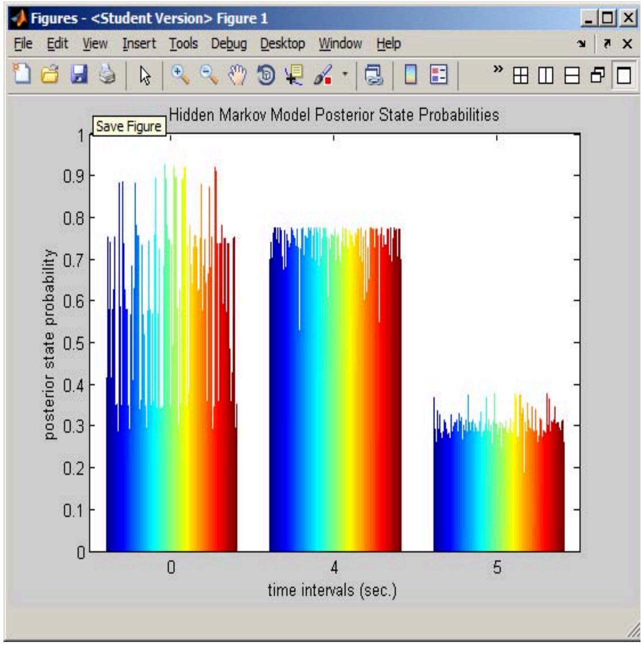


Figure 20. HMM Posterior State Probabilities over Time Interval 0, 4, and 5 seconds

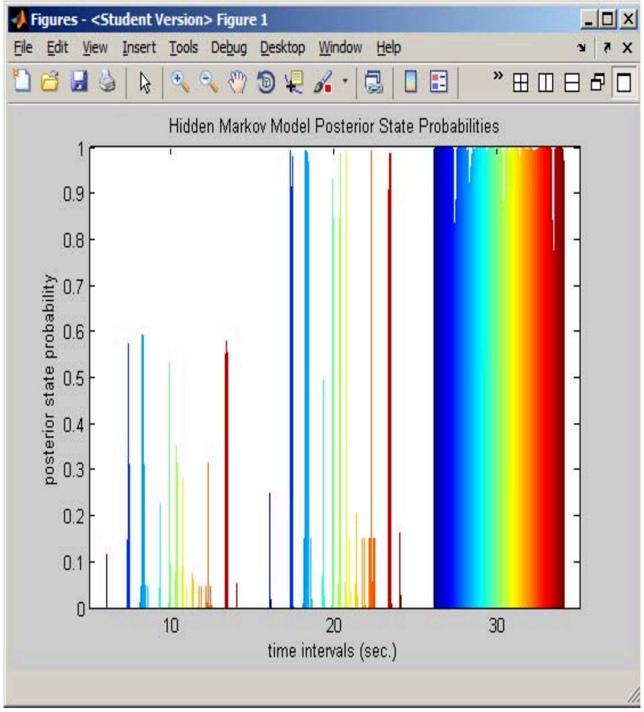


Figure 21. HMM Posterior State Probabilities over Time Interval 10, 20, and 30 seconds

As displayed in Figure 20 and 21, a high rate of risk was observed at 30-second interval. Given the model presented in the HMM calculation, the posterior security state probabilities were predicted (Figure 21). The results indicate that the HMM posterior probability has a similar pattern of risks calculated over a sequence of 1,000 matrix element sets as the previously assessed risk (Figure 16).

5. CONCLUSION

The near real-time risk assessment can provide timely risk levels of dynamically changing information assets by applying HMMs. This research has extensively explored the HMM and its applicability to risk assessment, and has shown that the HMM is a very effective modeling theory to assess risk in near real-time. In this simulation, the model, given the proper HMM inputs, calculated risk levels in real-time. Further, the HMM was able to predict posterior risk levels. The HMM posterior probabilities validated the near real-time risk assessment methodology by producing a similar pattern of outputs over a large sequence (1,000) of matrix element sets. The research has shown that given a HMM, future risk can be predicted using the near real-time risk assessment. As demonstrated in the HMM posterior probability, the HMM is effective in assessing dynamically changing information security states and predicting posterior [2] risk levels on mission critical assets.

The research results of this study offer a valuable contribution to the field of information security. When risk can be forecasted before the actual risk materializes, stakeholders can effectively manage risk before it becomes a reality.

As system administrators strive to remove vulnerabilities and mitigate risk on each network resource, management can review the total organizational risk picture (Figures 15 and 16). Figure 15 shows each asset risk assessment over time and provides stakeholders with a picture of security states in near real-time. The produced risk assessment can serve as a "thermostat" [5, p. 850] for the changing risk environment. Management would be able to monitor the dynamics of security state changes on mission critical assets and determine remedial actions prior to a compromised security state. This type of proactive risk assessment can add value to risk management and ultimately reduce risk management cost.

Optionally, the near real-time risk assessment can present its output in a graphical format and show the dynamics of risk level changes over time as different threats are introduced at different time intervals. The near real-time risk assessment could update its findings as threats or assets change over time. Observing threat changes and asset updates, and assessing risk levels of mission-critical assets in a near real-time basis would provide a valuable risk assessment to organizations, help determine appropriate safeguards to implement, and be able to justify cost-benefit decisions by stakeholders. Risk assessments should be as reliable, accurate, and timely as weather forecasts.

6. REFERENCES

[1] Chapra, S. (2008). *Applied numerical methods with MATLAB for engineers and scientists* (2nd ed.). NY: McGraw Hill.

[2] Eddy, S. (2004). What is a hidden Markov model? *Computational Biology, 10*(2), 1215–1216.

[3] Pak, C. (2008). The near real time statistical asset priority driven (nrtsapd) risk assessment. *Proceedings of the 9th ACM SIGITE, Cincinnati, Ohio,* 105–112.

[4] Pak, C., & Cannady, J. (2009). Asset priory risk assessment using hidden Markov models. *Proceedings of the 10th ACM Conference on Sig-information Technology Education, Fairfax, Virginia,* 65–73.

[5] Power, M. (2009). The risk management of nothing. *Accounting, Organizations and Society*, *34*(6–7), 849–855.