

SCS5607 ETHICAL HACKING AND DIGITAL FORENSICS

UNIT-V

FRAUD SELECTION & DETECTION

Key Fraud Indicator selection process

Customized Taxonomies

The following taxonomies were created within this research to assist in gaining a clearer understanding of the many facets of this complex and elusive topic:

1. The Universal ICF Taxonomy (Original Diagram)
2. Macro ICF Taxonomy (Original Diagram)
3. Taxonomy of Computer Fraud—Perpetration Platform (Lucian VasIU, Deakin University, Australia, and Ioana VasIU, Babeş-Bolyai University, Romania)
4. Taxonomy of Computer Fraud—Perpetration Method (Lucian VasIU, Deakin University, Australia, and Ioana VasIU, Babeş-Bolyai University, Romania)
5. Micro Insider Computer Loan Fraud Taxonomy
6. Insider Loan Taxonomy (Key Fraud Indicators [KFIs] and Key Fraud Metrics [KFMs])
7. Forensic Foto Frame Taxonomy (Original Diagram)
8. Metadata Taxonomy (Original Diagram)
9. Application Defect Taxonomy (Lucian VasIU, Deakin University, Australia, and Ioana VasIU, Babeş-Bolyai University, Romania)

Listed below are the primary areas in which the use of the taxonomies developed for this research were applied:

- ICF Journaling Workflow Diagram
- Development and Use of KFIs
- Development and Use of KFMs
- Development and Use of Key Fraud Signatures (KFSs)

Customized Taxonomies for Detecting ICF—The Universal ICF Taxonomy:

This taxonomy provides a comprehensive listing of potential ICF activities. However, the primary focus of this research is on the manipulation of data input, which is one the most prevalent forms of ICF.

Macro Computer Fraud Taxonomy:

This taxonomy (Figure 8.2) provides a comprehensive listing of potential ICF threats. The contents of these criteria represent a roll-up of the categories of ICF activities based upon the results of the *ICF Summary Report* and the *ICF Taxonomy* documents listed below. All of the criteria contained within the ICF Summary Report and the ICF Taxonomy were based on a collection of actual cited cases of ICF activities and listed in the public domain.

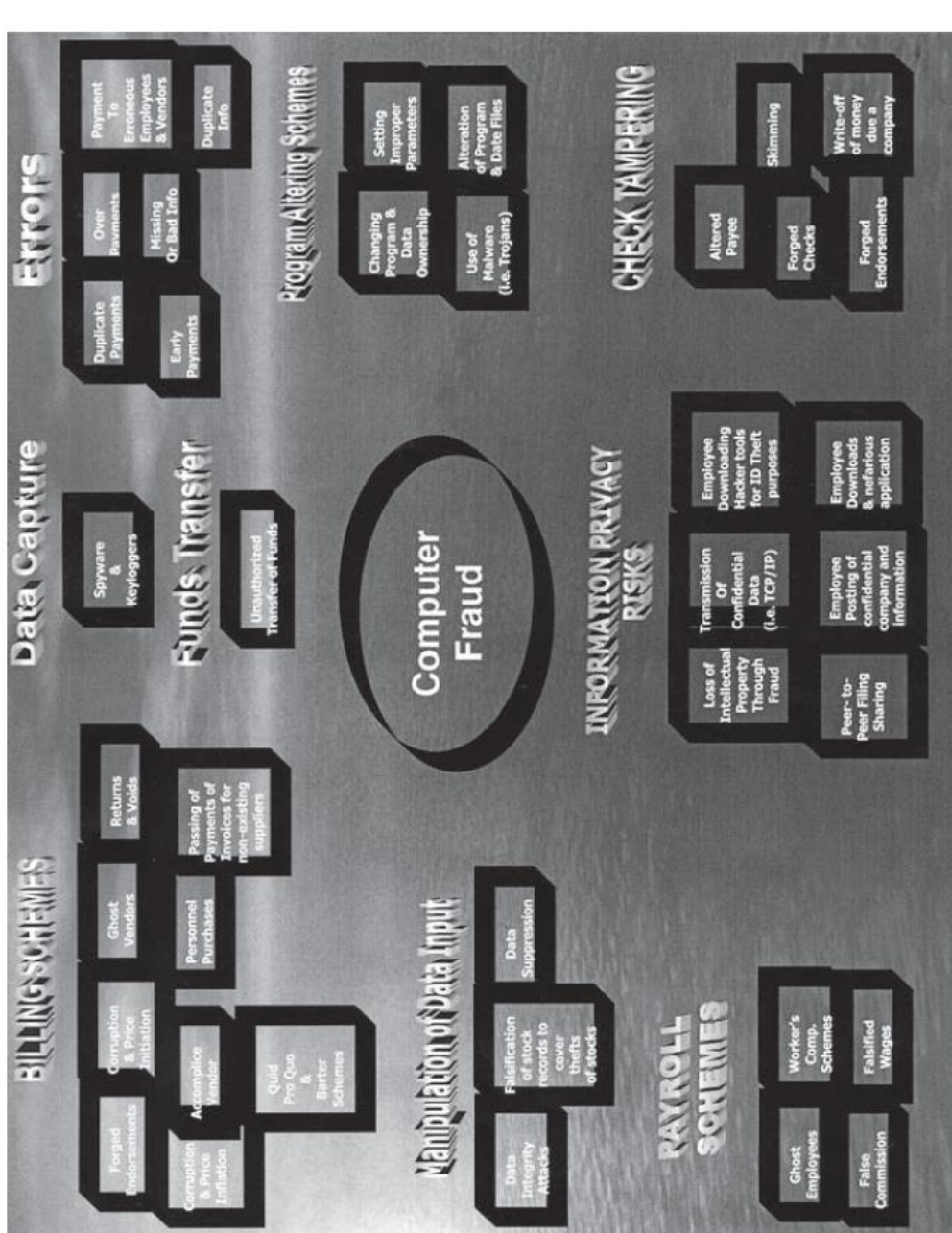


Figure 8.2 The universal ICF taxonomy.

Listed below are the names and sequences of reports prepared in support of developing the macro ICF taxonomy (Table 8.1):

- Macro ICF Taxonomy (Final Step)
- ICF Summary Report (Summary Report)
- ICF Taxonomy Heatmap (Interim Report)
- ICF Decomposition—ICF Case Analysis Report

Table 8.1 Macro ICF Taxonomy

	<i>Ontological Category (Parent)</i>	<i>Ontological Category (Child)</i>	
GENESIS	Data	Data input manipulation	HYBRID ICF
		Data destruction	
	System	Misuse of system capabilities (authorized users)	
		Hardware destruction or damage	
		Access control misuse	
		Hacking (unauthorized user)	
	Software	Unauthorized system access through the fraudulent use of ex-employees	
		Code manipulation	
		Logic bomb	
		Malcode injection	
	Trojan horse		

The perpetration methods in a taxonomy of computer fraud are generally described by the authors as input, program, and output. The authors state that the greatest concerns are the frauds that involve manipulation of data records or computer programs to disguise the true nature of transactions, cracking into an organization’s computer system to manipulate business information, and unauthorized transfers of funds electronically (Table 8.3).

Table 8.3 Taxonomy of Computer Fraud (Perpetration Method)

Data	Insert	Improper data	
		Data improperly	
	Improper obtaining or use		
	Integrity attacks		
	Availability attacks		
Program	Run attacks	Without authorization	
		In excess of authorization	
		Improper parameters	
		Transit attacks	Interruption
			Interception
			Modification
			Fabrication
	Integrity attacks		
	Availability attacks		

Source: Vasiu, Lucian and Vasiu, Ioana. Dissecting computer fraud from definitional issues to a taxonomy, *Proceedings of the 37th Annual Hawaii International Conference on Systems Sciences*. © 2004 IEEE. (Reprinted by permission.)

Micro Insider Computer Loan Fraud Taxonomy

The bank insider loan fraud taxonomy was developed based upon a review and analysis of a white paper produced by the Federal Financial Institution Examination Council (FFIEC), entitled “Insider detection, investigation and prevention of insider loan fraud,” for the FFIEC fraud investigation symposium, held October 20–November 1, 2002. (See Table 8.4.)

Table 8.4 Micro Taxonomy of Insider Computer Fraud—Bank Insider Loan Fraud

Data Manipulation	Insert	Falsified data
		Nominee loan name
		Improper use of loan proceeds
		Integrity issues
		Preferential rate and term for loan

Insider Loan Taxonomy (KFI and KFM)

This taxonomy (Figure 8.3) was developed based upon my analysis of the aforementioned FFIEC document that was used as the basis for determining KFIs and KFMs, which assisted in the illustration of how the framework could be implemented, using insider loan fraud within banks.

Metadata Taxonomy

The metadata taxonomy provided an integral component in establishing the criteria for the attribute selection for each KFI, KFM, KFS, and training and testing dataset for the novelty neural network. A good analogy between identifying the role of metadata and data is closely aligned with relational database design, where the primary key in the database schema would equate to the data element, and the attributes of the table would equate to the metadata.

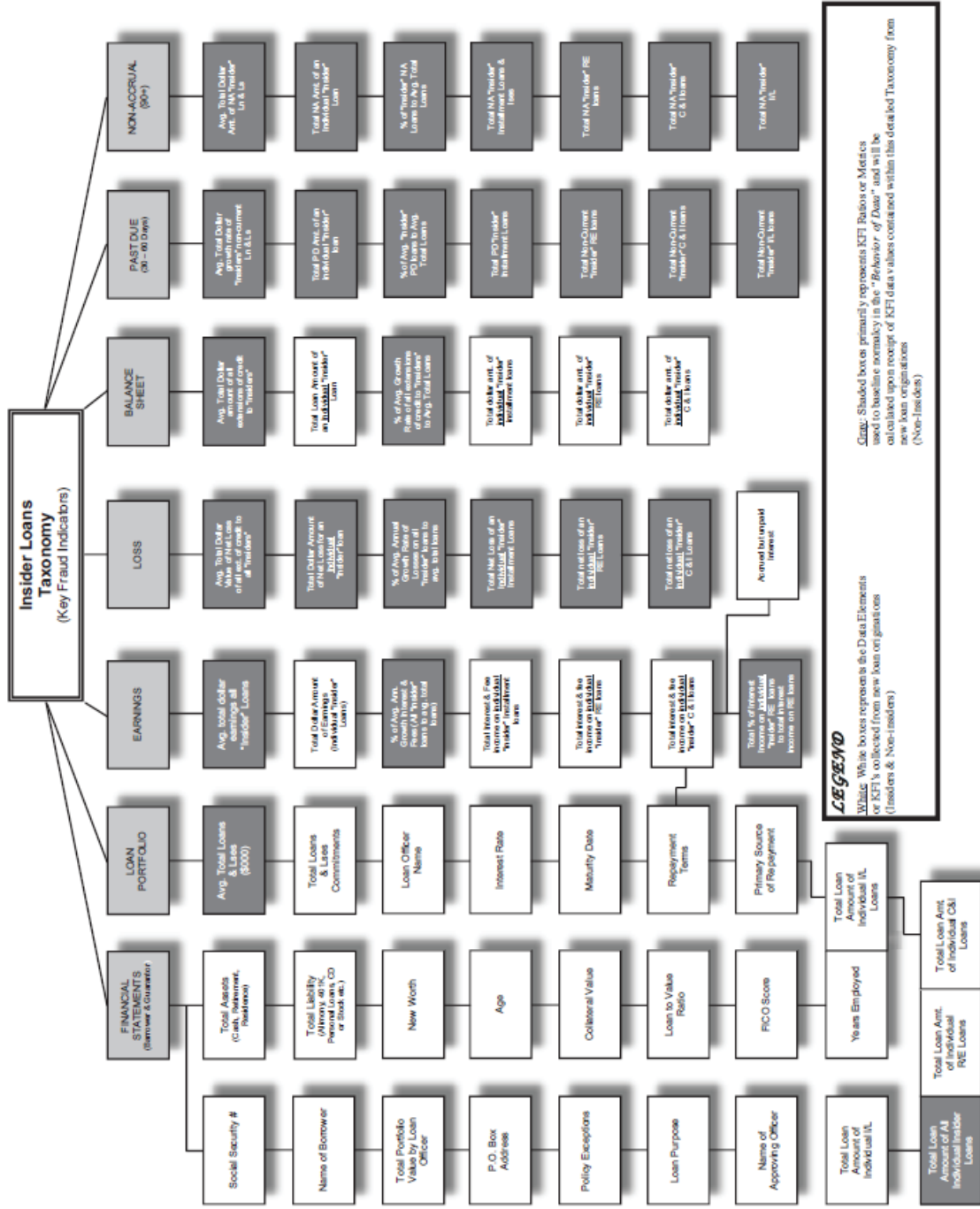


Figure 8.3 Insider loan taxonomy.

Key fraud signature selection process

The KFS selection and implementation processes are significant components to Layer 2 of this Defense in Depth insider computer fraud (ICF) Framework.

Just as a recap, the levels of my Defense in Depth Model include the following three components; however, it is important to note that each of these three layers do not have to be performed sequentially, but rather should be performed in concert given their close interrelationships:

Layer 1: Application and information technology (IT) control risk assessment

Layer 2: Application journaling

Layer 3: Training and testing the novelty neural network

The KFS selection process is initially more of an art than a science and will need the benefit of time and experience for users to more fully gain from the benefits of its use. Over time, when the ICF architectural framework has seen refinements based on a clearer understanding of the risks of a particular application or system, the identification, deletion, and refinement of an existing KFS will become more mature and repeatable and this process will eventually evolve into more of a science.

It would be beneficial at this point to introduce the KFS triangle (Figure 5.4) that graphically depicts the interrelationships between a KFS, key fraud metrics (KFM), and key fraud indicators (KFI). One approach for introducing any new method or process is to illustrate through example.

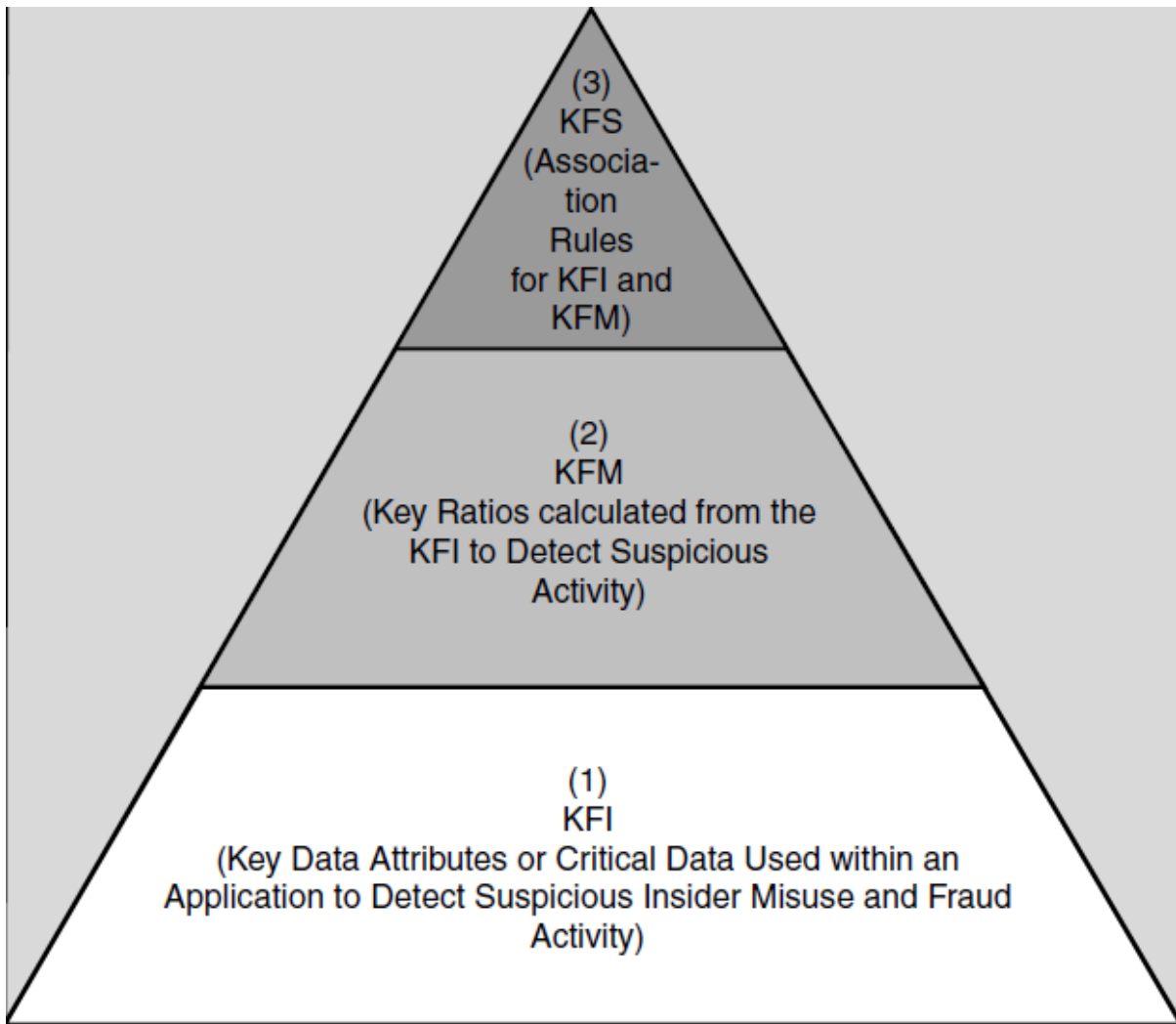


Figure 5.4 The key fraud signature (KFS) pyramid.

Accounting Forensics

The important aspect to note at this point is that under the aforementioned insider loan fraudulent scenario, the analysis reflects speculation regarding which specific accounts would be potentially impacted given the suspected insider misuse of the application and data.

In brief, an impact analysis would have to be performed of the suspected fraudulent transaction,

with a financial statement and forensic accounting analysis of that enterprise.

For illustrative purposes, the example below represents a simple format that could be used in identifying a KFS candidate, for use in KFS preparation, journaling, and neural network dataset preparation.

Example of KFSAR (Macro and Micro ICF Taxonomy)—Insider Loan Fraud Scenario

Although no KFSAR rules exist, one baseline rule that could be considered would be a basic “If Then Else” statement to determine associations between KFI and KFM attributes. To keep the illustrations simple, the following example will use only KFIs, not KFMs.

To illustrate the “If Then Else” statement, one rule format may take the form of the following:

Example of KFS Format:

If (Day = Saturday *and* Time = P.M. *and* Item = Beer), *then* cost <\$10.00, which just says that on Saturday evenings people normally buy small quantities of beer.

Example KFSAR:

If (sale values decrease *and* production costs increase *and* marketing costs decrease), *then* the financial risk is low.

Macro Taxonomy:

Unauthorized use or misuse of system access privileges or capability to change access controls to perpetrate fraud (high risk).

Software modification to funnel purchases into a “dummy” account and then erase any trace of fraud (moderate risk).

Misuse of system capabilities (low risk).

Based on the macro taxonomy, one potential KFSAR might include the following:

KFSAR 1 (Macro Taxonomy):

If (user access level increases *and* this is a new employee), *then* the financial risk is high.

To illustrate this example in another manner, which more closely aligns with our ICF scenarios involving insider computer loan fraud, perhaps the following format and content would provide another perspective. Look now at an actual KFS that was used within the insider computer fraud operational which reflects the following KFI value changes to give the appearance of a potential insider fraudster, who decided to engage in data input manipulation.

KFSAR 2 (Micro Taxonomy):

If (FinData _LoanBal declines by 15 percent *and* FinData_LoanBal_IL_Mod_Time_Meta changes *and* ABUI increases by 5 percent *and* Earn_ABUI_Mod_Time_Meta changes), *then* Alert KFS 1.

(Additional KFIs and KFMs can be added as appropriate to more accurately reflect new malicious ICF patterns.)

Assuming a transaction meets the criteria of KFSAR 2 and the data behavior is now considered suspicious and meeting the criteria of KFS 1, a KFS 1 alert should be transmitted to notify the appropriate InfoSec personnel to potentially trigger their computer incident response team (CIRT) processes to mitigate the risks associated with this activity.

There are, however, several mitigating factors that may influence the severity of these alerts. They include the following factors:

1. *A KFS Designation Does Not Apply to Each Forensic Foto Frame*: Although the KFS designation might be appropriate and technically meets the criteria outlined within the KFSAR, each KFS does not apply universally to every Forensic Foto Frame. Based on the ICF Service Oriented Architecture diagram listed below, there may be numerous control gates where Forensic Foto Frames will be taken. Each Forensic Foto Frame will be taken at various stages within the journey of the transaction and its data. What might be considered as a KFI or KFS anomaly, indicating a suspicious transaction, for Forensic Foto 1, may paradoxically be considered as normal behavior for Forensic Foto 2, based on different processing activities (i.e., calculations) that will change the behavior of the data.

2. *Direct and Indirect Correlation Conditions Have Not Been for a High-Risk KFS Designation*:

a. Direct Correlation: A general rule for a direct correlation to exist may require that certain conditions be satisfied. For example, a KFS 1 designation may also require the incorporation of a KFM to be considered a high risk, which may be included within another signature, say KFS 2.

b. Indirect Correlation: From an indirect correlation perspective, there may be ontological rules that establish pre- and postconditions to occur which taken together would warrant a high risk designation, which has similarities to the concepts of direct correlation.

3. *Novelties Detected from the Neural Network Will Change the Significance of KFS*: One of the major advantages of incorporating the use of a novelty neural network within my ICF Framework is to validate the accuracy, relevance, and significance of existing KFSs, KFIs, and KFMs.

Computer Forensics:

- The term “computer forensics” involves the discovery of computer-related evidence and data. Computer forensics is commonly used by law enforcement, the intelligence community, and the military.
- There are many technical implications involving the identification and collection of data, along with an equal number of legal implications in the identification, collection, preservation, and analysis of computer forensic data.
- The concepts of computer forensics, journaling, and computer incident response team (CIRT) processes are all inextricably linked together. Specifically, computer forensics is the science behind the collection and analysis of computer journaling and other evidence, and journaling is the practice of capturing key data for security monitoring and during a computer forensics examination, if ICF activity arises.
- When a suspected problem does arise, then the CIRT processes are activated to ensure the survivability of the organization and to determine a root cause.
- Journaling is the heart and soul of computer forensics and represents the evidentiary data that will aid those involved in the investigatory process in conducting a root cause analysis and investigation.
- Given the high level of importance of journaling and its direct relationship to computer forensics, there are obviously legal implications in the collection, handling, and analysis of this information that will only be briefly introduced in this section.
- It is important to note that an organization develops comprehensive CIRT policies and procedures that map the connections between CIRT processes, computer forensics, and journaling. Specifically, the aforementioned policy and standards should address the journaling requirements and recommend

journaling as part of the evidentiary data collection requirement for assessing the existence of hacking and other computer crime (for example, fraud, money laundering, embezzlement, or other misuse of the system).

Audit logs need to be stored in a secure place where attackers will not have access to the files. Following are a few ways to ensure protection of the logs:

- Setting the logical protection on the audit log so that only privileged users have write access.
- Storing the audit log to another computer dedicated to storing audit logs where no one has access to the machine.

Types of Evidence

Direct: This category of evidence is basically oral testimony given by an individual to either validate or dispute a given fact. The source of direct evidence is any of an individual's five senses (e.g., observing the physical location of computer equipment at the alleged crime).

Real: This category of evidence is made up of tangible objects (e.g., the computer and storage media used during an alleged crime).

Documentary: This category of evidence is tangible (e.g., computer printouts). It is important to note that the actual printout of data is considered hearsay evidence, because it is only evidence of the original evidence, which is the original data element stored within the computer. For additional details on documentation evidence, refer to the best evidence and hearsay rules noted below.

Demonstrative: This category of evidence is created to illustrate or further support criminal activity (e.g., a flowchart that graphically illustrates how a computer fraud occurred).

Best Evidence Rule: As previously described, documentary evidence, although admissible in a court of law, does not comply with the best evidence rule, which prefers the original evidence and not a copy.

Journaling and its requirements

The term “journaling” describes the creation of activity log records and the capture of key information about all security-relevant information technology (IT) systems.

Journaling is not considered a real-time activity, but rather an after-the-fact analysis of a transaction and data. Typically, such activities include the capture of the following information:

- Date and time of activity, actions taken, and users involved.
- Successful and unsuccessful log-on and log-off activity.
- Successful and unsuccessful accesses to security-related files and directories.
- Denial of access to excessive failed log-ons.

The National Industrial Security Program Operating Manual (NISPOM).

NISPOM sets the standards for protection of classified information. Covered under NISPOM are all commercial contractors who have access to classified information.

Security auditing involves recognizing, recording, storing, and analyzing information related to security-relevant activities. The audit records can be used to determine which activities occurred and which user or process was responsible for them.

Audit 1 Requirements

1. Automated Audit Trail Creation: The system shall automatically create and maintain an audit trail or log. (On a PL-1 system only: In the event that the operating system cannot provide an automated audit capability, an alternative method of accountability for user activities on the system shall be developed and documented.)

Audit records shall be created to record the following:

- a. Enough information to determine the date and time of action (e.g., common network time), the system locale of the action, the system entity that initiated or completed the action, the resources involved, and the action involved.
- b. Successful and unsuccessful log-ons and log-offs.
- c. Successful and unsuccessful accesses to security-relevant objects and directories, including creation, open, close, modification, and deletion.
- d. Changes in user authenticators.
- e. The blocking or blacklisting of a user ID, terminal, or access port and the reason for the action.

f. Denial of access resulting from an excessive number of unsuccessful log-on attempts.

2. *Audit Trail Protection*: The contents of audit trails shall be protected against unauthorized access, modification, or deletion.

3. *Audit Trail Analysis*: Audit analysis and reporting shall be scheduled and performed. Security-relevant events shall be documented and reported. The frequency of the review shall be at least weekly and shall be documented in the SSP.

4. *Audit Record Retention*: Audit records shall be retained for at least one review cycle or as required by the CSA (Cognizant Security Agency).

Audit 2 Requirements

In addition to Audit 1, Individual accountability (i.e., unique identification of each user and association of that identity with all auditable actions taken by that individual). Periodic testing by the ISSO or ISSM of the security posture of the IS.

Audit 3 Requirements

In addition to Audit 2, Automated Audit Analysis: Audit analysis and reporting using automated tools shall be scheduled and performed.

Audit 4 Requirements

In addition to Audit 3, An audit trail, created and maintained by the IS, that is capable of recording changes to mechanism's list of user formal access permissions.

Journaling Risk/Controls Matrix:

The following matrix details all the KFI and KFM attributes, which incorporates all the metadata previously discussed. The maturity date KFI (MATDT) was selected because of its pervasiveness in use by insider loan fraudsters at financial institutions, according to the FFIEC Insider Detection, Investigation and Prevention of Insider Loan Fraud: A White Paper Produced for the FFIEC Fraud Investigation Symposium, October 20–November 1, 2002.

It is noteworthy to mention that selecting the appropriate KFI and KFM will take time and careful planning. Establishing a well-conceived journaling risk/controls matrix is an important first step determining how many KFIs and KFMs will be selected and what attributes will actually be logged.

There are numerous logging possibilities for the MATDT KFI; however, only a few were selected in comparison to the entire population. For a complete listing of all the KFIs and KFMs.

Controls	Risks									
	Direct ^b			Indirect ^b						
<i>Attributes^a (Data and Metadata)</i>	<i>Fraud Scenario^c</i>			<i>Fraud Scenario^c</i>						
<i>Key Fraud Indicators (KFIs)</i>	2	3	4	6	7	75	25	56	57	70
KFM_NA_Tot_IL										
KFM_NA_Tot_Re										
KFM_NA_Tot_CI										

Standardized Logging Criteria for Forensic Foto Frames:

1. Administration/summary data:

- a. The number of metadata elements in each Forensic Foto Frame
- b. Author
- c. Author e-mail
- d. Data owner/maintainer
- e. Description
- f. Name of approving officer
- g. Attribute name, author, date, time, and frequency of TOTAL data object CREATIONS
- h. Attribute name, author, date, time, and frequency of TOTAL data object ACCESSES
- i. Attribute name, author, date, time, and frequency of TOTAL data object DELETIONS
- j. Attribute name, author, date, time, and frequency of TOTAL data object ADDITIONS
- k. Attribute name, author, date, time, and frequency of TOTAL data object MODIFICATIONS
- l. Attribute name, author, date, time, and frequency and TOTAL VIOLATIONS OF DATA ACCESS RULES
- m. Attribute name, author, date, time, and frequency and TOTAL number of embedded graphics (objects) creation, additions, and deletions
- n. TOTAL of algorithmic transformations (i.e., calculations)

2. Frame statistics:

- a. Creation date
- b. Creation time
- c. Last save time
- d. Revision number
- e. Total edit time (minutes)

3. Data access rules violations (access Level 1: read only—loan officer; access Level 2: write only—data entry personnel only; access Level 3: read/write—supervisory loan officer):

- a. Social security number
- b. Name of borrower

- c. Total portfolio value by loan officer
- d. P.O. box address
- e. Policy exceptions
- f. Loan purpose
- g. Name of approving officer
- h. Total assets
- i. Total liabilities
- j. Borrower net worth
- k. Collateral value
- l. Loan to value (LTV)
- m. FICO score
- n. Years employed
- o. Loan officer name
- p. Interest rate

- q. Maturity date
- r. Repayment terms
- s. Primary source of repayment
- t. Total dollar amount of earnings (individual “insider loan”)
- u. Total interest and fee income on individual “insider” installment loan
- v. Total interest and fee income on individual “insider” RE loan
- w. Total interest and fee income on individual “insider” C&I loan
- x. Total loan amount of all individual “insider” loan(s)
- y. Total loan amount of all individual “insider” I/L
- z. Total loan amount of all individual “insider” RE loan(s)
- aa. Total loan amount of all individual “insider” C&I loan(s)

4. Graphics/objects:

- a. Number of embedded objects:
- b. Date of embedded object creation, deletion, addition, modification
- c. Time of embedded object creation, deletion, addition, modification
- d. Frequency of embedded object creation, deletion, addition, modification
- e. Source of embedded object

5. Algorithmic transformations (i.e., calculations)

- a. Number of algorithmic transformations:
- b. Date of algorithmic transformations creation, deletion, addition, modification
- c. Time of algorithmic transformations creation, deletion, addition, modification
- d. Frequency of algorithmic transformations creation, deletion, addition, modification
- e. Source of algorithmic transformation

Neural networks – Misuse detection and Novelty detection:

One of the primary objectives of this research will be to understand the basic concept of neural networks and how they impact the detection of insider computer fraud (ICF) activities.

Computer Forensic Benefits of Neural Networks

The forensic journaling that will be built into the software engineering process for new application development will also assist in the development of NNs to detect unknown or anomalous insider user behavior. The use of NNs for ICF detection has many advantages and is well suited to the elusive nature of ICF activities:

- Has the ability to handle nonlinear problems.
- Needs no processing algorithm.
- Has the ability to model chaotic time series.

The Neural Network Development Process

Prior to understanding the nexus between the use of digital forensics data and the development of neural nets for capturing key journaling criteria, there is a need to establish a fundamental understanding of the NN development process.

Specifically, anomaly detection in NNs is created by having systems learn to predict the next user command based on a sequence of previous commands by a specific user.

Basically, the building of a NN for use within intrusion detection systems consists of three phases:

1. Collect training data by obtaining the audit logs for each user for a certain period. A vector is formed by each day and each user, which shows how often the user executed each command.
2. Train the NN to identify the user based on the command distribution vectors.
3. Command the NN to identify the user based on the command distribution vector. If the network's suggestion is different from the actual user an anomaly is signaled.

To address this increasing information security threat, there has been a growth in the industry in the use of ADSs and IDSs.

There are many cited issues involving the use of anomaly/intrusion detection, which include the following:

- Problems with scalability
- False positives
- An inability to determine what is really important information
- A lack of a complete, comprehensive database of attack signatures

Novelty Detection (Saffron Technologies)

By definition, novelty detection identifies abnormal or nonrandom behavior that demonstrates a process is under some influence of special causes of variation, without impeding the normal learning process that is so vital to creating associative memory.

The detection of novelty is an important concept, particularly when dealing with ICF activities, because it provides a feedback mechanism to the user and validates the effectiveness and accuracy of the training and testing dataset or perhaps flaws within the initial underlying logic of the software. Substantive user acceptance and quality assurance testing would have to be conducted prior to any conclusions in either scenario.

Using Saffron's didactic tool, LabAgent and companion documentation, the properties of this software and the underlying concept behind novelty NNs in general involve the following fundamental characteristics or properties:

- *Incremental*: Start from zero, learn case-by-case.
- *Nonparametric*: No knob-tweaking to build.
- *Malleable*: Adapt on the fly to new features.
- *Unified Representation*: Various inferences can be computed at query time.
- *No Overtraining*: Do not get worse as more data is seen.

Anomaly Detection Using Neural Networks (Fuzzy Clustering)

- Technologies that include fuzzy clustering are beginning to be used. Fuzzy clustering is being chosen instead of relying on the use of classifiers, which may not deal as effectively with detecting events that do not neatly fall into any predefined cluster.
- Basically, the term *fuzzy clustering* works by ostensibly training itself, through the creation of a baseline profile of the network in various states, to determine what happens under normal conditions. It then determines what different

users do and the resources they normally request, and what types of files they transfer and other activity.

Misuse Detection Using Neural Networks

- As previously noted, the use of attack signatures alone is not as effective as if they were combined with other forms of prevention and detection when it comes to network or ICF attacks involving Web-based or traditional applications accessed only in-house.
- The signature-based attack detection process can be effective if tuned and continually baselined against known networks or can be compared against application attacks.

REFERENCES:

1. Kenneth C.Brancik, "Insider Computer Fraud", Auerbach Publications Taylor & Francis, Group 2008.
2. Caudill, Maureen and Butler, Charles, *Naturally Intelligent Systems*, MIT Press, Cambridge, MA, 1992.
3. Hawkins, Jeff, *On Intelligence*, Times Books, Henry Holt, New York, 2004..
4. Saffron Technologies, Technical White Paper, Morrisville, NC, 2004 (www.saffrontech.com).
5. Nigrini, Mark, Fraud Detection—I've Got Your Number. *Journal of Accountancy*, May, 79–83, 1999.