Alternative Form of Dempster's Rule for Binary Variables

Rajendra P. Srivastava
Ernst & Young Professor and Director
Ernst & Young Professor of Accounting and Information Systems
The University of Kansas, Lawrence, KS 66045
Phone: 785-864-7590, Fax: 785-864-5328
Email: rsrivastava@ku.edu

December 2004

Alternative Form of Dempster's Rule for Binary Variables

ABSTRACT

This article develops an alternative form of Dempster's rule of combination for binary variables. This alternative form does not only provide a closed form formulae for efficient computation but also enables researchers to develop closed form analytical formulae for assessing risks such as information security risk, fraud risk, audit risk, independence risk, etc., involved in assurance services. We demonstrate the usefulness of the alternative form in calculating the overall information security risk and also in developing an analytical model for assessing fraud risk.

Alternative Form of Dempster's Rule for Binary Variables

Dempster's rule is the fundamental rule in Dempster-Shafer theory of belief functions (Shafer 1976) for combining items of evidence pertaining to a variable. The general form of Dempster's rule as presented by Shafer (1976) is conceptually easy to understand but computationally very complex to operationalize. In particular, if one is trying to develop analytical models for assessing such risks as audit risk (Srivastava et al 2004a), fraud risk (Srivastava et al 2004b), and auditor's independence risk (Turner et al. 2004), the general form is not of much help.

In this research note, we develop an alternative form of Dempster's rule for combining items of evidence that pertain to a binary variable. The alternative form of Dempster's rule presented in this article provides not only a closed form formula for efficient computation but also enables researchers to develop closed form analytical formulae for assessing risks as mentioned earlier. Such analytical formulae are needed, especially when empirical evidence shows that auditors do think of uncertainties in terms of belief functions as demonstrated by Harrison et al. (2002).

Dempster's Rule of Combination for Binary Variables

Dempster's rule of combination of beliefs from two independent items of evidence is given by (Shafer 1976):

$$m(A \neq \varnothing) = \sum_{A=A_1 \cap A_2} m_1(A_1)m_2(A_2)/K, \qquad (1)$$

$$K = 1 - \sum_{A_1 \cap A_2 = \varnothing} m_1(A_1)m_2(A_2). \qquad (2)$$

where m(A) represents the combined m-value on A, m_1 and m_2 represent the two sets of m-

values[1], on the frame Θ, and K represents the renormalization constant. The second term in K

represents the conflict between the two items of evidence. If the conflict term is 1, i.e., if the two

items of evidence exactly contradict each other, then K = 0 and, in such a situation, the two items

of evidence are not combinable. In other words, Dempster's rule cannot be used when K = 0.

In order to derive the general form of Dempster's rule for a binary variable, we first

consider the above formula for two items of evidence and then generalize it for n-items of

evidence. Let us consider a binary variable X with two values: 'x', that the variable X is true and

'$\sim x$' that the variable is not true. The frame of discernment, Θ, is given by $\Theta = \{x, \sim x\}$. Let us

assume the following m-values to represent the two set of beliefs obtained from two independent

items of evidence pertaining to variable X:

Evidence 1: $m_1(x)$, $m_1(\sim x)$, and $m_1(\Theta)$.

Evidence 2: $m_2(x)$, $m_2(\sim x)$, and $m_2(\Theta)$.

The combined m-values using Dempster's rule given in Equations (1) and (2) can be

written as:

$$m(x) = [m_1(x)\, m_2(x) + m_1(x)\, m_2(\Theta) + m_1(\Theta)\, m_2(x)]/K, \qquad (3)$$

$$m(\sim x) = [m_1(\sim x)\, m_2(\sim x) + m_1(\sim x)\, m_2(\Theta) + m_1(\Theta)\, m_2(\sim x)]/K, \qquad (4)$$

$$m(\Theta) = m_1(\Theta)\, m_2(\Theta)/K, \qquad (5)$$

where K is given by

---

[1]Shafer (1976) calls these m-values as the *basic probability mass assignment function* whereas Smets (1990a, 1990b, 1998) calls them the *basic belief mass assignment function*.

$$K = 1 - \{m_1(x)\, m_2(\sim x) + m_1(\sim x)\, m_2(x)\}. \tag{6}$$

Equations (3) and (4) can be rewritten[2] as:

$$m(x) = 1 - (1 - m_1(x))(1 - m_2(x)\,)/K, \tag{7}$$

$$m(\sim x) = 1 - (1 - m_1(\sim x))(1 - m_2(\sim x))/K, \tag{8}$$

and the renormalization constant K in Equation (5) can be rewritten as:

$$K = (1 - m_1(x))\,(1 - m_2(x)) + (1 - m_1(\sim x))\,(1 - m_2(\sim x)) - m_1(\Theta)\, m_2(\Theta). \tag{9}$$

Equations (5), (7) and (8) give the combined m-values for binary variable, X, when two

independence items of evidence pertaining to the variable are combined using Dempster's rule.

By extending the above results from two independent items of evidence to n

independence items of evidence pertaining to variable X, one obtains m-values as given in the

following proposition:

Proposition: Dempster's rule yields the following m-values when n independent items of
evidence pertaining to a binary variable, X are combined:

---

[2] Since $m(x) = [m_1(x)\, m_2(x) + m_1(x)\, m_2(\Theta) + m_1(\Theta)\, m_2(x)]/K$, and
$$K = (1 - m_1(x))\,(1 - m_2(x)) + (1 - m_1(\sim x))\,(1 - m_2(\sim x)) - m_1(\Theta)\, m_2(\Theta),$$
we can write m(x) as:
$$m(x) = [\{m_1(x) + m_1(\Theta)\}\{m_2(x) + m_2(\Theta)\} - m_1(\Theta)m_2(\Theta)]/K,$$
$$= [\{1 - m_1(\sim x)\}\{1 - m_2(\sim x)\} - m_1(\Theta)m_2(\Theta)]/K,$$
$$= 1 - \{1 - m_1(x)\}\{1 - m_2(x)\}/K.$$
Similarly, we can write m(~x) as:
$$m(\sim x) = [\{m_1(\sim x) + m_1(\Theta)\}\{m_2(\sim x) + m_2(\Theta)\} - m_1(\Theta)m_2(\Theta)]/K,$$
$$= [\{1 - m_1(x)\}\{1 - m_2(x)\} - m_1(\Theta)m_2(\Theta)]/K,$$
$$= 1 - \{1 - m_1(\sim x)\}\{1 - m_2(\sim x)\}/K.$$

$$m(x) = 1 - \prod_{i=1}^{n}(1 - m_i(x))/K, \tag{10}$$

$$m(\sim x) = 1 - \prod_{i=1}^{n}(1 - m_i(\sim x))/K, \tag{11}$$

$$m(\Theta) = \prod_{i=1}^{n} m_i(\Theta)/K, \tag{12}$$

where K is given by

$$K = \prod_{i=1}^{n}(1 - m_i(x)) + \prod_{i=1}^{n}(1 - m_i(\sim x)) - \prod_{i=1}^{n} m_i(\Theta). \tag{13}$$

The plausibility functions can be written[3] as:

$$Pl(x) = \prod_{i=1}^{n}(1 - m_i(\sim x))/K = \prod_{i=1}^{n} Pl_i(x)/K, \tag{14}$$

$$Pl(\sim x) = \prod_{i=1}^{n}(1 - m_i(x))/K = \prod_{i=1}^{n} Pl_i(\sim x)/K. \tag{15}$$

The proof of the above proposition is straight forward extension of Equations (5), (7)-(9) through induction. As one can see, Equations (10)-(13) provide a way to not only compute the resultant beliefs efficiently but also help one derive analytical formulae for the overall beliefs in a complex situation where the decision maker has several independent items of evidence pertaining to a binary variable. Such situations are quite common in the real world. For example, the auditor while conducting a financial audit encounters multiple items of evidence for a given account. The above results have been used by Srivastava et al (2004), Turner et al (2004a,

---

[3] By definition: $Pl(x) = 1 - Bel(\sim x)$ which yields:

$$Pl(x) = 1 - m(\sim x) = \prod_{i=1}^{n}(1 - m_i(\sim x))/K = \prod_{i=1}^{n} Pl_i(x)/K,$$

where K is defined in (13).

2004b) in deriving analytical formulae for assessing the overall audit risk[4], fraud risk and

independence risk in a financial audit.

Applications to Business Decisions

In this section we show the usefulness of the alternative form of Dempster's rule in

computing the overall beliefs on a binary variable and also show how this form makes it easy to

develop analytical formulae for assessing fraud risk.

Numerical Example: Information Systems Security Risk

Sun et al. (2004) have recently developed an evidential reasoning approach to assessing

information systems security risk. They use the Dempster-Shafer theory of belief functions to

model the uncertainties involved in the evidence. We use part of their model, as depicted in

Figure 1, to illustrate the value of the alternative form of Dempster's rule. As seen in Figure 1,

we have seven items of evidence pertaining to the assertion "Customer information is protected

from unauthorized internal access and is used in ways associated with the entity's business." Let

us assume that we have the following set of m-values from the seven items of evidence that the

assertion is true (t), not true (~t), or we don't know whether it is true or not true, {t, ~t}.

Evidence 1: $m_1(t) = 0.3$, $m_1(\sim t) = 0.1$, $m_1(\{t, \sim t\}) = 0.6$.

Evidence 2: $m_2(t) = 0.2$, $m_2(\sim t) = 0.1$, $m_2(\{t, \sim t\}) = 0.7$.

Evidence 3: $m_3(t) = 0.3$, $m_3(\sim t) = 0$, $m_3(\{t, \sim t\}) = 0.7$.

---

[4] Audit risk is defined as the risk that the auditor has given a clean opinion on the financial
statements but there is a possibility that financial statements may contain material
misstatements due to errors and fraud. Fraud risk is the risk that financial statements contain
material misstatements due to management fraud. Independence risk is the risk that the auditor
is not independent of the audit client while conducting the financial audit.

Evidence 4: $m_4(t) = 0.5, m_4(\sim t) = 0.1, m_4(\{t, \sim t\}) = 0.4.$

Evidence 5: $m_5(t) = 0.6, m_5(\sim t) = 0.1, m_5(\{t, \sim t\}) = 0.3.$

Evidence 6: $m_6(t) = 0.3, m_6(\sim t) = 0.2, m_6(\{t, \sim t\}) = 0.5.$

Evidence 7: $m_7(t) = 0.6, m_7(\sim t) = 0.1, m_7(\{t, \sim t\}) = 0.3.$

Using the alternative form of Dempster's rule given in Proposition 1, one can easily

obtain the combined m-values. The renormalization constant K for the above case is given by

$$K = \prod_{i=1}^{7}(1 - m_i(t)) + \prod_{i=1}^{7}(1 - m_i(\sim t)) - \prod_{i=1}^{7}m_i(\{t, \sim t\}) = 0.7\text{x}0.8\text{x}0.7\text{x}0.5\text{x}0.4\text{x}0.7\text{x}0.4$$

$$+ \; 0.9\text{x}0.9\text{x}1.0\text{x}0.9\text{x}0.9\text{x}0.8\text{x}0.9 - 0.6\text{x}0.7\text{x}0.7\text{x}0.4\text{x}0.3\text{x}0.5\text{x}0.3 = 0.489052,$$

and the m-values as:

$$m(t) = 1 - \prod_{i=1}^{7}(1 - m_i(t))/K = 1 - 0.7\text{x}0.8\text{x}0.7\text{x}0.5\text{x}0.4\text{x}0.7\text{x}0.4/0.489052 = 0.955113,$$

$$m(\sim t) = 1 - \prod_{i=1}^{7}(1 - m_i(\sim t))/K = 1 - 0.9\text{x}0.9\text{x}1.0\text{x}0.9\text{x}0.9\text{x}0.8\text{x}0.9/0.489052 = 0.034066,$$

$$m(\{t, \sim t\}) = \prod_{i=1}^{7}m_i(\{t, \sim t\})/K = 0.6\text{x}0.7\text{x}0.7\text{x}0.4\text{x}0.3\text{x}0.5\text{x}0.3/0.489052 = 0.01082.$$

Thus, we can easily determine the overall beliefs and plausibilities that the assertion is true or not

true as given below:

$$Bel(t) = 0.955, \; Bel(\sim t) = 0.034,$$

$$Pl(t) = 1 - Bel(\sim t) = 1 - m(\sim t) = \prod_{i=1}^{7}(1 - m_i(\sim t))/K = 0.966, \text{ and}$$

$$Pl(\sim t) = 1 - Brl(t) = 1 - m(t) = \; = \prod_{i=1}^{7}(1 - m_i(t))/K = 0.045.$$

Plausibility that the assertion "Customer information is protected from unauthorized internal access and is used in ways associated with the entity's business" is not true represents the risk related to this assertion. Sun et al. (2004) discuss information security risk in great detail.

As we can see from the above example, the alternative form of Dempster's rule makes it very convenient to compute the combined beliefs in one step. In fact, one can easily program the logic in MS Excel Spreadsheet to compute the combined m-values for a large number of independent items of evidence.

Fraud Risk

Srivastava et al. (2004) have developed a comprehensive analytical model for assessing the risk of fraud in financial statements using Dempster-Shafer theory of belief functions. Also, Turner et al (2004) have used Dempster-Shafer theory of belief functions to develop an analytical model for assessing the overall audit risk that financial statements will contain material error due to random errors and/or due to management fraud while the auditor has given a clean opinion. In all such cases, one needs to combine various independent items of evidence pertaining to a single binary variable. The alternative form of Dempster's rule becomes very useful in such situations. Here, we consider a much simpler example to illustrate the value of the alternative form of Dempster's rule for binary variables.

The auditing literature (see, e.g., Ramos 2003) discusses that management will commit fraud when the following three conditions exist: management lacks integrity, management has incentives, and there are opportunities to commit fraud. The auditor is required through professional standards (AICPA 2002) that he/she should assess the risk of management fraud by assessing whether management lacks integrity, has incentives and opportunities to commit fraud.

For simplicity, we represent this combined assessment of the risk of fraud by one item of

evidence (**RF**) in Figure 2. Once the auditor has identified that there is a potential for

management fraud then he/she evaluates the existing mitigating factors (**MF**) that may be present

in the situation and would reduce the risk of fraud. Security and Exchange Commission's

regulations and penalties to CEOs (Chief Executive Officers) for committing fraud are few

examples of such mitigating factors. Also, while performing the routine audit procedures such as,

analytical procedures (**AP**: ratio analyses, comparison of the current year account balance with

the last year account balance and with the industry average, etc), test of controls (**TC**), and test of

details of the account balance (**TD**), the auditor might be able to detect fraud. While these routine

audit procedures are not very effective in detecting fraud, they do detect sometimes. Thus, one

should not attach too much weight on these items of evidence when they pertain to "No Fraud"

assertion. Ultimately, if the assessed risk of fraud is high then the auditor should perform

forensic procedures (**FP**) to detect fraud. All these items of evidence are depicted in Figure 2

through rectangular boxes. They all pertain to one assertion or variable that "There is no fraud in

objective O of account A". We assume that this variable has two values, "yes, fraud is present

(yf)" and "no, fraud is not present (nf)."

The opportunity to commit fraud and also the type of routine audit procedures depend on

the specific assertion[5] of the account in interest. For example, accounts receivable balance would

---

[5] According to the auditing literature (AICPA 1980), there are five management assertions:
'Existence or Occurrence' 'Completeness' 'Valuations or Allocations' 'Rights and
Obligations' and 'Presentations and Disclosures'. When management publishes the financial
statements of the company, they implicitly imply that the above assertions are true for all the
transactions. In other words, management implies that all recorded tansactions exist or have
occurred (Existence or Occurrence), all transactions are recorded (Completeness), all
transactions are valued properly, the management has the right of assets to use and obligation
of liabilities to pay (Rights and Obligations), and all relevant disclosures and presentations are

be overstated if management created fictitious sales on credit. In order to detect this fraud, the auditor needs to perform audit procedures that are specific to the objective "Existence" of accounts receivable, which implies that the account receivable balance is not overstated due to fictitious transactions. Also, the type of forensic procedures performed by the auditor will depend on the nature of account and its objective.

Using Equation (15), the plausibility that fraud is committed in objective O of account A, given the evidence in Figure 2, can be written as:

$$Pl(yf) = Pl_{RF}(yf)Pl_{MF}(yf)Pl_{AP}(yf)Pl_{TC}(yf)Pl_{TD}(yf)Pl_{FP}(yf)/K, \qquad (16)$$

where K is determined by Equation (13), and $Pl_{..}(..)$'s on the right side represent various plausibilities that there is fraud based on the corresponding evidence labeled as the subscript.

Srivastava and Shafer (1992) argue that plausibility that error exists in the financial statements is a better measure of risk. Using their definition of risk, we write the following formula for fraud risk (FR) by using (16):

$$Fraud\ Risk = FR = FR_{RF}\ FR_{MF}\ FR_{AP}\ FR_{TC}\ FR_{TD}\ FR_{FP}/K \qquad (17)$$

The above model suggests that the overall fraud risk is the product of six fraud risks assessed by the following evidence: 1) factors such as lack of management integrity, presence of incentives, and presence of opportunities, 2) mitigating factors, 3) analytical procedures, 4) test of controls, 5) test of details, and 6) forensic procedures. The above formula can be used to determine the

---

made (Presentations and Disclosures). We will use the term 'Objective' in place of 'Assertion' in this article.

overall fraud risk by assessing the individual components. The simple form of the fraud risk formula as given in (17) is possible only because of the alternative form of Dempster's rule.

Conclusion

We have derived an alternative form of Dempster's rule of combination of evidence for binary variables. This form allows one to compute the combined m-values more efficiently and also allows one to develop analytical formulae for real world problems such as audit risk, fraud risk, independence risk, and etc. Such closed form analytical formulae are needed for assessing the above risks during an audit/assurance engagement.

References

American Institute of Certified Public Accountants (AICPA). 2002. *Consideration of Fraud in a Financial Statement Audit*. SAS No. 99. New York, NY: AICPA.

American Institute of Certified Public Accountants. 1980. *Statement on Auditing Standards, No, 31: Evidential Matter,* New York: AICPA.

Harrison, K., R. P. Srivastava, and R. D. Plumlee. 2002. Auditors' Evaluations of Uncertain Audit Evidence: Belief Functions versus Probabilities. In *Belief Functions in Business Decisions,* edited by R. P. Srivastava and T. Mock, Physica-Verlag, Heidelberg, Springer-Verlag Company: 161-183.

Ramos, M. 2003. Auditors' Responsibility for Fraud Detection. *Journal of Accountancy*, May: 28-36.

Shafer, G. 1976. *A Mathematical Theory of Evidence*. Princeton University Press.

Smets, P. 1990a. The Combination of Evidence in the Transferable Belief Model. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* 12, 5 (May).

Smets, P. 1990b. Constructing the Pignistic Probability Function in a Context of Uncertainty. *Uncertainty in Artificial Intelligence 5.* ed. by Henrion, M., Shachter, R.D., Kanal, L.N., and Lemmer, J.F. North-Holland: Elsevier Science Publishers B.V.

Smets, P. 1998. The Transferable Belief Model For Quantified Belief Representation. *Quantified Representation for Uncertainty and Imprecision,* Vol. 1. Edited by P. Smets. Kluwer Academic Publishers

Srivastava, R. P., and G. Shafer. 1992. Belief-Function Formulas for Audit Risk. The *Accounting Review*, April: 249-283

Srivastava, R. P., T. Mock, and J. Turner. 2004. The Effects of Integrity, Opportunity, Incentive and Mitigating Factors on Fraud Risk. Working paper, University of Kansas.

Sun, L., R. P. Srivastava, and T. Mock. 2004. An Information Systems Security Risk Assessment Model under Dempster-Shafer Theory of Belief Functions. Working Paper, School of Business, The University of Kansas.

Turner, J., R. P. Srivastava, and, T. Mock. 2004a. Integrating Fraud Risk into the Audit Risk Model: The Evolution of Audit Risk. Working paper, University of Kansas.

Turner, J., T. Mock, and R. P. Srivastava. 2004b. An Analysis of the Auditors' Independence: Effects of Incentive, Opportunity, and Rationalization. Working paper, University of Kansas.

Figure 1: Seven Items of Evidence pertaining to one Assertion

E1: Policies are set for ways to use customer information.

E2: Only a small number of operations staff has access rights to the database.

E3: All requests for data are channeled through a DBA who then requests from operation staff.

E4: Staff can not directly access the underlying SQL DB that stores customer information.

E5: Policy restricts the staff from disclosing private customer information to any third party.

E6: Clearly states on the web that ABC can give access to customers' private data to other divisions.

E7: Policies and monitoring procedures to ensure only certain employees can access to customer information.

Customer information is protected from unauthorized internal access and is used in ways associated with the entity's business.

Figure 2: Fraud Risk Assessment*

RF: Management Lacks Integrity, has Incentives and Opportunities to Commit fraud.

No Fraud Related to Objective O of Account A {nf, yf }

MF: Mitigating Factors Pertaining to Opportunities to Commit Fraud related to objective O of account A.

AP: Analytical Procedures related to Objective O of Account A.

TC: Test of Controls pertaining to objective O of account A.

TD: Test of Details pertaining to Objective O of Account A.

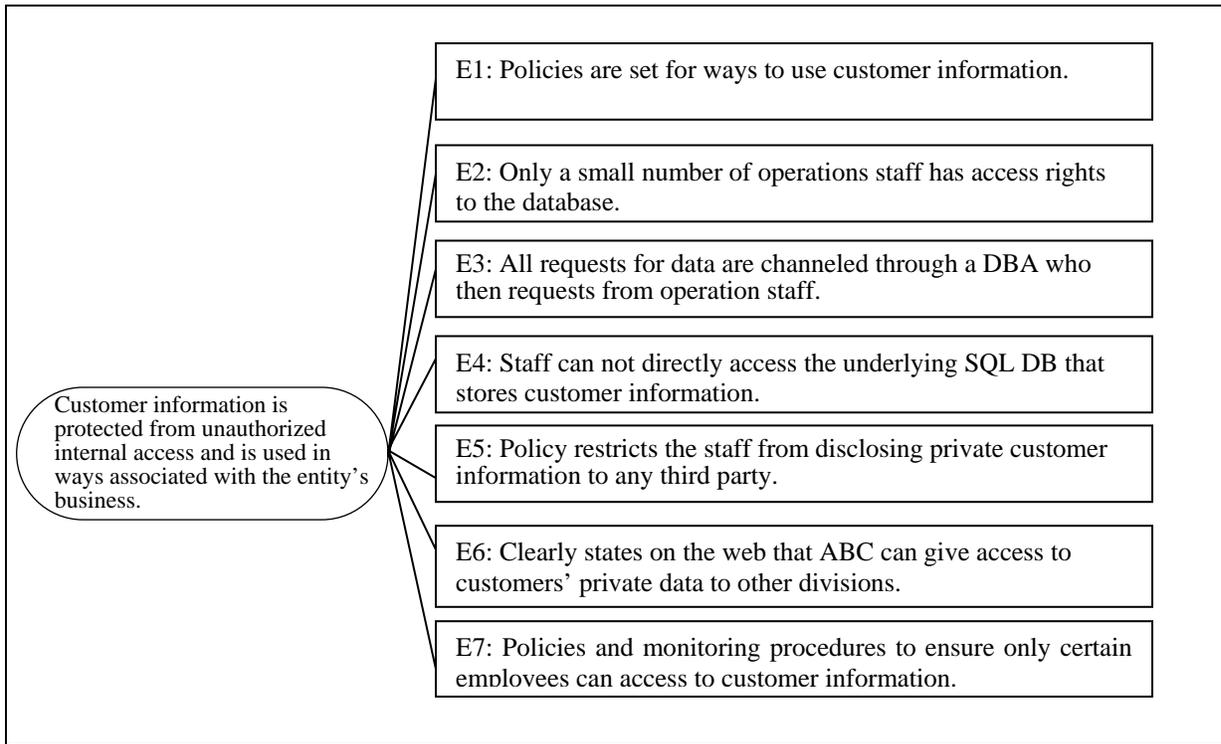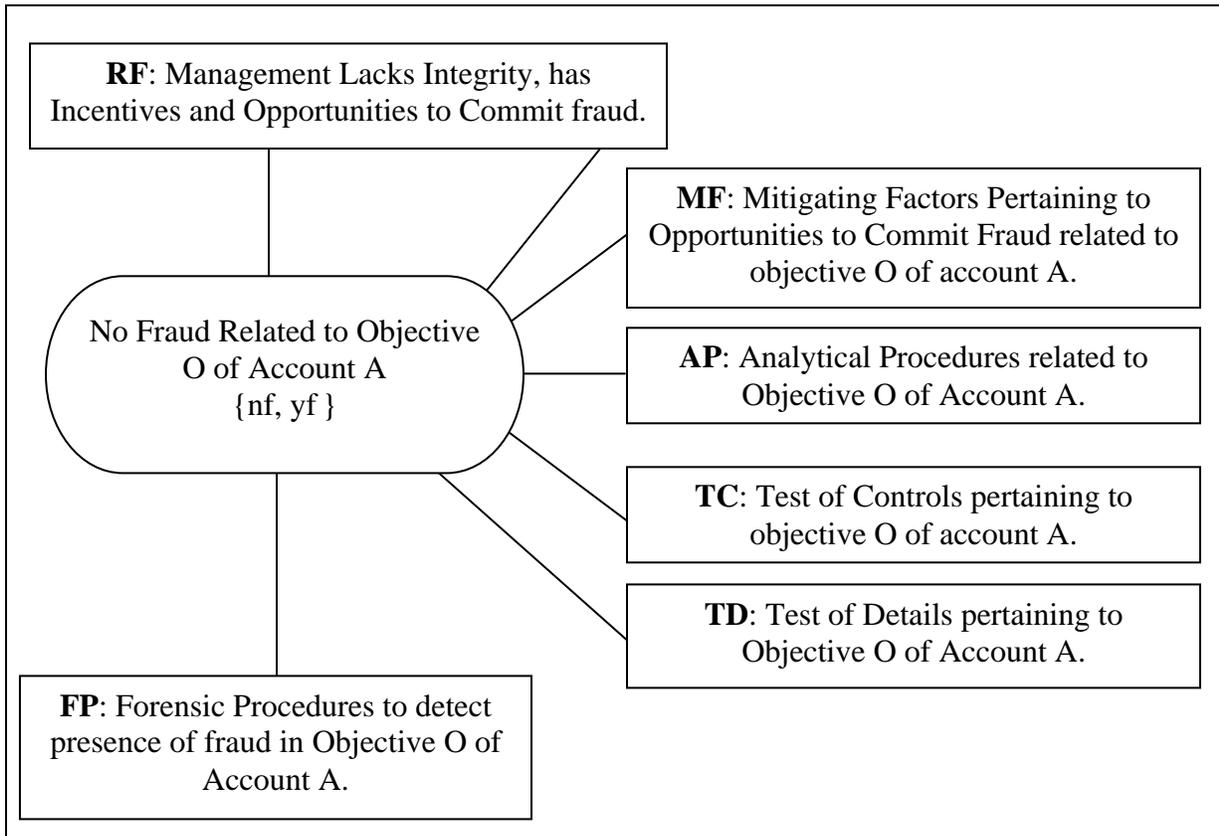**FP**: Forensic Procedures to detect presence of fraud in Objective O of Account A.

*The oval shape box represents the assertion or the variable and the rectangular boxes represent items of evidence.