ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

A Novel Approach Based Information Integrity Modeling

QamarRayees Khan, Muheet Ahmed Butt, Majid Zaman, Mohammad Asger
Baba Ghulam Shah Badshah University, Rajouri (J&K), India.
University of Kashmir, Srinagar, (J&K), India
University of Kashmir, Srinagar, (J&K), India
Baba Ghulam Shah Badshah University, Rajouri (J&K), India

Abstract- Errors in computer based information systems were relatively manageable as long as there was homogenous system environment and centralized control over information. Emerging trends of globalization, changing organizational patterns, strategic partnering, and electronic commerce and distributed computing have changed all this, resulting in loss of integrity in information systems. These errors are essentially caused by online factors of change, complexity, communication, conversion and corruption (the 5 C's). These factors have their presence in IS mainly through system environment that is external to computing (and hence the application) system and overlaps the user environment. The proposed research paper tries to suggest the methodsto improve the optimum Integrity in Information System by quantifying intrinsic integrity attributes of accuracy, consistency and reliability on the data to be mined.

I. INTRODUCTION

The field of information security has grown and evolved significantly in recent years. The end of the 20th century and early years of the 21st century saw rapid advancements in telecommunications, computing hardware and software. The availability of smaller, more powerful and less expensive computing equipment made electronic data processing within the reach of small business and the home user. These computers quickly became interconnected through a network generically called the Internet or World Wide Web. The rapid growth and widespread use of electronic data processing and electronic business conducted through the Internet, along with numerous occurrences of international terrorism, fueled the need for better methods of protecting the computers and the information they store, process and transmit as defined by CNSS [1]. The academic disciplines of information security and information assurance emerged along with numerous professional organizations - all sharing the common goals of ensuring the security and reliability of information systems. Security of information resources must include controls and safeguards to offset possible threats as well as controls to ensure timeliness, availability, integrity, confidentiality, etc. [2] [3].

Being one of the most fundamental security properties, integrity is subject to a vast area of research. The study of integrity [4] [5] as a formal security property has received little attention within the research community. Integrity is the core concepts of information security. Integrity, in Information Technology terms, means that data remains unchanged while stored or transmitted. Unauthorized changes to stored data violate integrity. We refer to security textbooks [6, 7] that discuss assorted flavors of integrity, and integrity surveys [8] [9] [10] and tutorials [11] that develop integrity classifications. To the best of our knowledge, our framework is the first to unify information integrity for programs. As mentioned previously, our departure point is the classification by Li et al. [12]. Our contribution compared to this classification is a more general model of invariants (Li et al. only discuss predicate invariants), a more general model of information flow (Li et al. do not consider endorsement), and a unified view, where we show that program correctness subsumes invariance policies. In addition, we also offer a unified enforcement mechanism that guarantees all aspects of integrity at once. Information-flow integrity dates back to Biba's integrity model [13], which dualizes Bell and LaPadula's model [14] [15] for mandatory access control. The Clark-Wilson integrity model [16] is a classical model that focuses on separation of duties and transactions.

II. INFORMATION INTEGRITY ATTRIBUTE QUANTIFIERS

Having developed the foundation for Integrity objective, next question is to look for approaches to quantification of intrinsic integrity attributes of accuracy, consistency and reliability, and to propose a method for demonstrating integrity improvement in information system.

A. Choice of Data/Information Model

This describes for first deciding on a practicable data/information model. Exclusively for the problem objective at hand, data taken form the information model can be modeled by a triple $<e_i, a_i, v_i>$ that represents input to the core information system, and information by a triple $< e_0, a_0, v_0>$that represents output from the information

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

system; < e, a, v > represents datum of a triple <entity, attribute, value> as developed by the database research community. This form of representation, which authorize treating data/information as formal organized collection, allows dividing integrity issue into issues relating to entities, attributes and values thereby making it feasible to study Information system integrity analytically [17] [18] [19].As networked computerized information systems contain errors that are made but not corrected, it is the above data/information model that needs to be further improved by replacing triple <e, a, v> by triple < e, a, v + η >where represents error or noise component that is responsible for inaccurate, inconsistent and unreliable information and, thereby, for loss of integrity in Information systems.

B. Accuracy

Accuracy is another key-attribute of information. It means that information is free from mistakes and errors, is clear and accurately reflects the meaning of data on which it is based. It conveys an accurate picture to the recipient, who may require a presentation in graphical form rather than tabular form. Accuracy is the degree of correctness and precision with which information in an automated system represents states of the real world. It is a very important quality dimension that on which many early information quality studies have focused [20][21][22]. Within information production processes inside organizations, accuracy can be improved by implementing institutional procedures, like having information double checked by two independent people, or by installing technical means, like calibrating sensors or verifying shipping address information received through a website against an address database. The concept of accuracy implies the assumption that information can be captured in an objective fashion. Thus, accuracy is not applicable to subjective information, like destructive impact, public perception or political views. Inaccurate information may be worse than no information at all. Accuracy is also the degree of agreement between a particular value and an identified source. It can be assessed by identifying the relevant established source (standard) and by determining an acceptable tolerance. Specifically, the identified source provides the correct value – preferably the value corresponding to the optimum integrity. It can be an object or relationship in the real world; it can also be the same value in another database, or the result of a computational algorithm.

C. Consistency

Consistency summarizes the validity, accuracy, usability and integrity of related data between applications and across an IT enterprise. This ensures that each user observes a consistent view of the data, including visible changes made by the user's own transactions and transactions of other users or processes. Consistency is the degree to which multiple instances of a value satisfy a set of constraints. The multiple instances may exist across space (such as databases or systems) or over time. Thus, consistency is with respect to a set of constraints and data/information is said to be consistent with respect to a set of constraints if it satisfies all constraints of the data/information model [23].Constraints can apply to the same attributes in different entities (such as the salary attribute in the entities of several employees); they can also apply to different attributes in the same entity (such as the marks percentage and marks attributes in the entity for a particular student).

Various kinds of consistency have been identified. These include Application Consistency, Transaction Consistency and Point-in-Time (PiT) Consistency. Data is point-in-time consistent if all of the interrelated data components (either a group of data sets or a set of logical volumes) are as they were at any single instant in time. Point-in-time consistency is an important property of backup files and a critical objective of software that creates backups. Transaction consistency is also frequently referred to as atomicity. A transaction is a logical unit of work that may include any number of file or database updates. A good example of the importance of transaction consistency is a database that handles the transfer of money. Application Consistency is similar to Transaction consistency, but instead of data consistency within the scope of a single transaction, data must be consistent within the confines of many different transaction streams from one or more applications.

D. Reliability

Reliability, which traditionally is a large concern in the system development lifecycle model, is a little complex attribute to define as it has a dual meaning in modern technical usage. In the broad sense, it refers to a wide range of issues relating to the design of large systems (complex computerized information system [CIS] included), which are required to work well for specified periods of time. In such a case, the term “reliability” includes descriptors such as “quality” (commonly understood from the traditional “standard’ product angle)and “dependability”, and is interpreted as a qualitative measure of how a system matches the specifications and expectations of a user. From this point of view for an IS the definition of reliability given as “accuracy with which information obtained represents data item in whatever respect the information system processed it" [17] can be seen to define the reliability requirement for the IS as a whole; reliability index being amenable to quantification through techniques such as Analysis of Variance (AOV) [24].

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

E Integrity Profile

Consider an information system designed and developed for an application area. It is appreciated that each application area, consistent with information usage requirements, will have application area specific order of significance for integrity attributes.

III. INFORMATION INTEGRITY TECHNOLOGY IMPLEMENTATION STEPS

Within the framework of Information Integrity attributes of Accuracy (A), Consistency (C) and Reliability (R) argued, we can then identify Information Integrity Technology Implementation steps as follows:

- Understand the user application of the computer based information system under consideration and Establish organizational standard pertaining to data/information vis-à-vis requirements of: accuracy, consistency, reliability and cumulative integrity, based on application area.
- Study data/information that flows through the Information System and define the source and destination data sources.
- Develop the Model of the Information System under consideration as in Figure (1), based on understanding of data/ information flow in the system for the identified data sources and also specify and document data rules, also known as edits, that is to be implemented to study accuracy and consistency of the data/information in proposed system.
- Choose a method for calculating Reliability Index, keeping in view advantages, disadvantages and convenience of application while accounting for factors such as nature and form of available data, and available computation aids.
- Propose Integrity Analysis method for analyzing intrinsic Integrity attributes of accuracy, consistency and reliability. In addition to this, the proposed method may also undertake statistical analysis (time series analysis and other techniques) of error patterns signifying irregular changes, which contributes to loss of accuracy and consistency and of causes, which contribute to loss of Reliability.
- The proposed method can be adapted for analyzing the error Detection Database constructs so as to identify the data rule violations in terms of accuracy and consistency attributes and set up the degree of integrity of data/ information in respect of Information Integrity.
- Make Comparative study of the Integrity profile and indices acquire with: standards in local, regional, national, and international as the case may be – and the user specification on Integrity, so as to know what is to be expected of Information Integrity Technology. This would also assist in ordering or ranking of the Integrity attributes from the points of which attributes needs utmost improvement effort. This can also provide efficiencies and deficiencies for anticipated system and will also pave way to propose certain adjustment on the anticipated system so that these deficiencies are curtailed.
- After locating the pairs of a given field at a given subsystem, each for enhancement of accuracy and consistency and having located given subsystems for reliability improvement opportunities, further analyze the Error Detection Database and [25] study uneven changes at each of pairs consequent to accuracy and consistency attributes and also study reliability factors at each of the subsystems, so as to understand over the time the error patterns [25] and causes that contributes to the loss of accuracy, consistency and reliability. This would then assist in detecting the error or cause that occurred in due course of time in the past or estimating error or cause at time (t), or predict error or cause that may probably occur in future.
- Also propose a Comparative Study performance of the Information System on inclusion of the Information Integrity Technology as discussed above. Accordingly obtained the intrinsic Information Integrity attribute indices, Integrity profile and Cumulative Information Integrity Index and compare them with appropriate reports before implementation of information Integrity Technology available.
- Finally after applying various attributes on various Information Systems/ Data Sources the proposed system is evaluated for whether these intrinsic attributes have modified the data from its original version or not. These changes could be detrimental for the data and need to be analyzed properly. To overcome these situations effective data mining techniques will be

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

implemented before and after these operations are carried on the proposed data like Classification, Clustering, Association, Predication and Sequential Patterns.

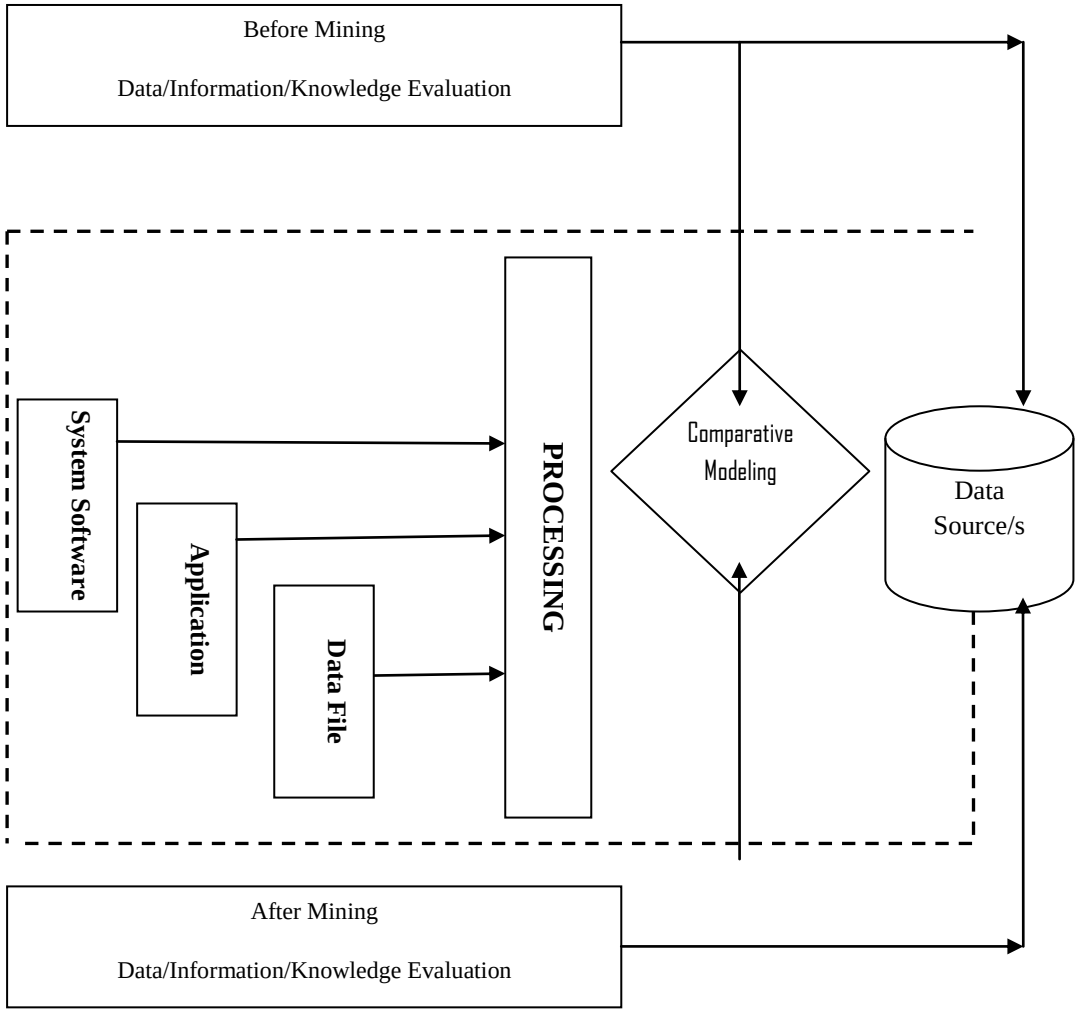

Fig 1. Data/Information Flow Model for an Information System for Implementing Information Integrity Technology

IV. CONCLUSION

Computerized information systems contain errors that are made but not corrected by controls built-in at system analysis and design stage of the Information System. Therefore, the confirmation of potential or suspected anomalies in a live database and subsequent integrity improvement becomes an essential facility (beyond application controls) within an Information System. This facility is the Information Integrity Technology. Specifically what it means is that benefits of increase in information use by achieving integrity are compared with costs of acquiring information as also of evaluating and applying integrity. There is a competitive advantage in the form of an overall benefit, as long as increase in information use value is more than costs. This then advocates an optimum integrity value that is desirable for maximum competitive advantage. It is this observation that then offers the basis for developing Information Integrity attribute quantifiers leading to the statement of Information Integrity Technology based on a feedback control system approach.

REFERENCES

[1]. Delone M. H., and McLean E. R. (1992) InformationSystems Success: The Quest for the Dependent Variable, Information Systems Research 3 (1).

[2]. Mandke Vijay V., and Nayar M.K. (1999) ModelingInformation Flow for Integrity Analysis, Proceedings of the 1999 Conference on Information Quality, Edited byY. V. Lee and G. K. Tayi, MIT, Cambridge, Massachusetts, USA.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

[3]. Made Vijay V. (1996) Research in InformationIntegrity: A Survey and Analysis, Proceedings of theJNCASR and SERC Discussion Meeting at IISc Campus, Bangalore on Information Integrity Issues and Approaches, Edited by Rajaraman V. and Mandke VijayV., published by Information Integrity Foundation, New Delhi, India.

[4]. K.J. Biba. Integrity considerations for secure computer systems. Technical Report MTR-3153 Rev 1 (ESD-TR-76-372), MITRE Corp Bedford MA, 1976.

[5]. Boritz, J.E. 2005. ARE practitioners' views on core concepts of information integrity? International Journal of Accounting Information Systems 6:260-279.

[6]. C. P. Pfleeger and S. L. Pfleeger. Security in Computing (4th Edition). Prentice Hall, 2006.

[7]. D. Gollmann. Computer Security (2nd Edition). Wiley, 2006.

[8]. T. Mayfield, J. E. Roskos, S. R. Welke, J. M. Boone, and C. W. McDonald. Integrity in automated information systems. Technical Report P-2316, Institute for Defense Analyses, 1991.

[9]. R. S. Sandhu. On five definitions of data integrity. In Proceedings of the IFIP WG11.3 Working Conference on J. Guttman. Invited tutorial: Integrity. Presentation at the Dagstuhl Seminar on Mobility, Ubiquity and Security, Feb. 2007. http://www.dagstuhl.de/07091/. Slides at http://web.cs.wpi.edu/˜guttman/.

[10]. Mandke Vijay V., and Nayar M.K. (1997) Information Integrity □ □ A Structure for its Definition, Proceedings of the 1997 Conference on Information Quality, Edited by Diane M. Strong and Beverly K. Kahn,MIT, Cambridge, Massachusetts, USA.

[11]. C. Dima, C. Enea, and R. Gramatovici. Nondeterministic no interference and deducible information flow. Technical Report 2006-01, University of Paris 12, LACL, 2006.

[12]. K. J. Biba. Integrity considerations for secure computer systems. Technical Report ESDTR- 76-372, USAF Electronic Systems Division, Bedford, MA, Apr. 1977. (Also available through National Technical Information Service, Springfield Va., and NTIS AD-A039324.).

[13]. D. E. Bell and L. J. LaPadula. Secure computer systems: Mathematical foundations. Technical Report MTR-2547, Vol. 1, MITRE Corp., Bedford, MA, 1973.

[14]. L. J. LaPadula and D. E. Bell. Secure computer systems: A mathematical model. Technical Report MTR-2547, Vol. 2, MITRE Corp., Bedford, MA, 1973. Reprinted in J. of Computer Security, vol. 4, no. 2–3, pp. 239–263, 1996.

[15]. D. D. Clark and D. R.Wilson. A comparison of commercial and military computer security policies. In Proc. IEEE Symp. on Security and Privacy, pages 184–193, May 1987. [38] P. Ørbæk. Can you trust your data? In Proc. TAPSOFT/FASE'95, volume 915 of LNCS, pages 575–590. Springer-Verlag, May 1995.

[16]. P. Ørbæk and J. Palsberg. Trust in the λ-calculus. J. Functional Programming, 7(6):557– 591, 1997.

[17]. Mandke Vijay V. and Nayar M.K. (1998) Design Basis for Achieving Information Integrity. A Feedback Control System Approach, IFIP TC 11 Working Group11.5 Second Working Conference on Integrity and Internal Control in Information Systems, Edited by SushilJajodia, William List, Graeme W. McGregor and LeonA.M. Strous, Published by Kluwer Academic Publishers.

[18]. Mandke Vijay V., and Nayar M.K. (2001)Information Envelope and its Information IntegrityImplications: For a complex, changing environment,modeling a generic business process as an integral to aclosed loop information and control system characterized by uncertainty, Proceedings of the 2001 Conference on Information Quality, Edited by Elizabeth M. Pierce andRaissaKatzHass,MIT, Cambridge, Massachusetts, USA.

[19]. Nayar M.K. (1996) A Framework for AchievingInformation Integrity, Proceedings of the JNCASR and SERC Discussion Meeting at IISc Campus, Bangalore on Information Integrity □ □ Issues and Approaches, Edited byRajaraman V. and Mandke Vijay V., published byInformation Integrity Foundation, New Delhi, India.

[20]. Alexander, J. E. & Tate, M. A. (1999). Web wisdom: how to evaluate and create information quality on the Web. Mahwah, NJ: Lawrence Erlbaum Associates.

[21]. Katerattanakul, P. and Siau, K., 1999. Measuring Information Quality of Web Sites: Development of an Instrument. Proceeding of the 20th International Conference on Information System.p: 279-285.

[22]. Wang R. and Strong D. (1993) an empirical Investigation of Data Quality Dimensions: A Data Consumer's Perspective, TDQM9312, The TDQMResearch Program, MIT, Sloan School of Management, and Cambridge, USA.

[23]. AT&T Publication (1992) Data Quality Foundations, Published by AT&T Quality Steering Committee, USA.

[24]. Vijay V. Mandke, Madhavan K. Nayar. Implementing Information Integrity Technology- A Feedback Control System Approach.

ISSN: 2319-5967
ISO 9001:2008 Certified
International Journal of Engineering Science and Innovative Technology (IJESIT)
Volume 2, Issue 1, January 2013

[25]. Khan Q. Rayees, Butt Muheet A, Baba M. Zaman, Asger M. Data Cleansing Approaches for Automatic Error Detection in Data Sets. IJSC January, 2013.