

Implementing Information Integrity Technology – A Feedback Control System Approach

Vijay V. Mandke
Research Leader
Centre for Information Integrity Research
Unitech Systems (I) Pvt. Ltd.,
B-64 (First Floor), Gulmohar Park
New Delhi-110049, INDIA
E-mail : vijaymandke@satyam.net.in

Madhavan K. Nayar
President
Unitech Systems, Inc
1240 E. Diehl Road, Suite 300
Naperville, Illinois 60563, USA.
E-mail : mnayar@unitechsys.com

Abstract

Intrinsic information integrity attributes of accuracy, consistency and reliability are central to any information system in that in their absence the information systems (IS) will have massive amounts of polluted (error – filled) data and useless, even dangerous information. These errors are essentially caused by on-line factors of change, complexity, communication, conversion and corruption which have their presence mainly through system environment which is external to computing (and hence the application) system and overlaps with the user environment. Need, therefore, is for on-line error detection and integrity improvement mechanisms in the IS models in the form of automatic feedback control systems. The paper addresses this research issue of implementing integrity technology. Specifically, the paper begins with the choice of information model for integrity improvement, followed by presentation of some alternatives for quantification of intrinsic integrity attributes and development of integrity profile and cumulative information integrity index as a means for demonstrating integrity improvement. This is followed by presentation of information integrity technology implementation steps. Finally, the paper gives a description of information integrity technology, thus emerging, as a software product and details it.

1 INTRODUCTION

Errors in computerized 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, electronic commerce and distributed computing have changed all this, resulting in loss of integrity in information systems. These errors are essentially caused by on-line factors of **change, complexity, communication, conversion and corruption (The 5 C's)**. Change may be in the content or in configuration of the system environment. Complexity is due to introduction of new component, be it a program, database or network, thereby adding new interfaces. Communication is the movement of data/information within or across enterprises. Conversion means consolidation, decomposition or transformation of data and corruption refers to human behavior, inherited errors and unpredictability. Human behavior includes poor motivation, desire for personal gain, carelessness, and intended or unintended actions of people. Inherited error occurs when an error is propagated beyond the system in which it originated. Inherited errors pollute the information system. Unpredictability is noise of any kind, e.g., communication

channel noise, equipment failure, etc. These five factors, namely, the 5 C's, have their presence in IS mainly through system environment which is external to computing (and hence the application) system and overlaps the user environment. In spite of application controls, it is these external factors that then introduce in information systems, errors that are made but not corrected [Mandke and Nayar, 1997; Nayar, 1996].

It is within above framework of errors in computerized information systems and their integrity implications, that research investigations presented at IFIP TC 11 WG 11.5 Second Working Conference (Mandke and Nayar, 1998) identify intrinsic integrity attributes of accuracy (includes completeness and timeliness implying accuracy in spite of time related changes in data/information), consistency (satisfying domains and constraints) and reliability (accuracy with which information item represents data item in whichever way information system processed it) which, irrespective of nature of use, any IS must satisfy. Research investigations further observe that depending on the context and nature of use there can be other optional integrity attributes of security and privacy which can be seen as extrinsic or subjective integrity attributes specific to area of use [Mandke, 1996]. Other such subjective attributes of integrity could be: interpretability, ease of understanding, tractability, cost effectiveness, flexibility, etc. [Wang and Strong, 1993].

It is to ensure above integrity attributes that research investigations presented at IFIP TC 11 WG 11.5 Second Working Conference [Mandke and Nayar, 1998] have proposed need to incorporate on-line learning and error correcting mechanisms in the IS models. Specifically, to account for errors in IS that are made but not corrected, they propose incorporation of automatic feedback control systems with error detection and correcting technologies for improved information accuracy, consistency and reliability; technologies that maximize integrity of information systems – Information Integrity Technologies. They further argue that, when incorporated, it is such Information Integrity Technology that would also facilitate demonstrating improved integrity of information obtained, rather than merely trusting the computerized information systems.

There are obvious difficulties in implementing such automatic feedback control systems, the most important being to study error patterns. Specifically, it is not possible to track and analyse every bit of data/information for all times as it flows through the information system stages. Way out here is to consider Information Integrity Technology that takes a sample of input data at the output or at an intermediate point of an appropriately identified stage or sub-system of the IS and then follows or keeps track of the sampled records at output or intermediate points of subsequent stages (sub-systems), at a given point of time or at different points of time over a required time interval [Mandke and Nayar, 1997; Mandke and Nayar, 1998].

2 INFORMATION INTEGRITY ATTRIBUTE QUANTIFIERS

This brings forth the central question as to what will be the nature of such an Information Integrity Technology. To answer this and particularly to suggest ability of such an integrity technology to demonstrate integrity improvement in information obtained, it is first necessary to consider the question of quantification of intrinsic integrity attributes of accuracy, consistency and reliability and of overall system integrity. Towards this the investigation at hand draws on the IS model accounting for errors that are made but not corrected, as developed by Mandke and Nayar in their paper presented at Second Annual IFIP TC-11 WG 11.5 Working Conference [Mandke and Nayar, 1998].

2.1 Choice of Model for Data / Information : A Basis for Integrity Quantifiers

Specifically, networked computerized information systems of today see “Data” as raw material, “Data Product or Information” as processed data used to trigger certain management action, “Processing” as the system function, and are characterized by (a) computing processes that include micro-computer and

telecommunication and (b) pre and post-processing stage communication channels at various data/information processing nodes, that are people based and include data communication and transaction processing networks with world-wide reach. Such decentralized structure of IS has certainly facilitated organizations and individuals to work with shared data environments and with capture, use and control of growing, complex and diversified volumes of data and information; in turn making it possible for business to access bigger markets.

In such information systems, data can be modeled by a triple $\langle e_i, a_i, v_i \rangle$ representing input to the information system and information by a triple $\langle e_o, a_o, v_o \rangle$ representing output from the information system; $\langle e, a, v \rangle$ representing datum a triple $\langle \text{entity, attribute, value} \rangle$ as developed by the database research community. This representation which permits treating data/information as formal organized collection allows to segment integrity issue into issues concerning entities, attributes and values thereby making it feasible to study IS integrity analytically.

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 $\langle e, a, v \rangle$ by triple $\langle e, a, v + \eta \rangle$ wherein η represents error or noise component responsible for inaccurate, inconsistent and unreliable information and, thereby, for loss of integrity in IS. It may be mentioned that this more realistic representation of data/information model is most simplistic in that it is only accounting for information item on value (v) for the error that is made but not corrected. Certainly such errors can also be present even at the stage of “view” definition where “view” consists of entity types.

Considering that these error implications are present at each stage of an information system, namely; data origin stage, communication channel prior to processing stage, processing stage, communication channel at post processing stage and output stage, there are integrity implications at each stage of the IS and at the overall system level, as shown in Figure 1 [Rajaraman, 1996].

What is important for the investigation at hand is that integrity of the overall Information system is ensured if the integrity requirements of all parts of the

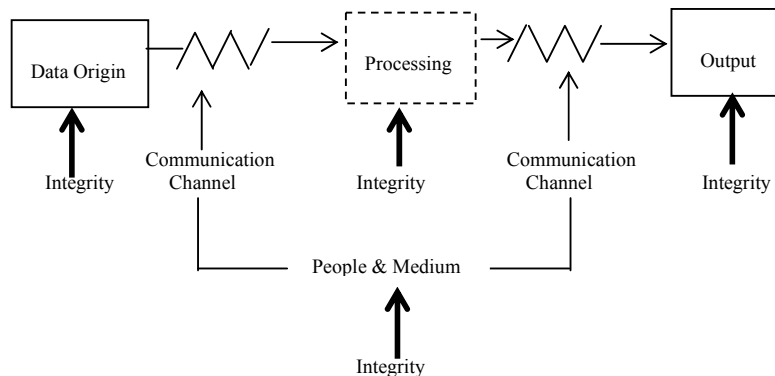


Figure 1 Conceptual Presentation of Integrity of an Information System

system as in Figure 1 are ensured [Rajaraman, 1996]; integrity being defined in terms of attributes of accuracy, consistency and reliability whose quantification being the query to be pursued. In what follows this section addresses this query.

2.2 Accuracy

Accuracy refers to correctness, i.e., preventing unauthorized modification, i.e., degree of conformance between a particular value of data/information and an identified source. The identified source provides

the correct value [AT&T Publication, 1992]. 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.

Given that value of data/information is expressed in a numerical, accuracy of the data/information can be quantified in a number of ways [Redman, 1992; AT&T Publication, 1992; Svanks, 1984; Ameen, 1989]:

i) Difference between the actual value (i.e., value of the identified source) and the value processed by the information system.

$$\text{ii) Error Ratio} = \frac{\text{Actual Error}}{\text{Acceptable Error}}$$

$$\text{iii) Accuracy Index} = \frac{\text{Number of correct values}}{\text{Number of total values}}$$

iv) Number of records examined : R

Number of records with atleast one defect of loss of Accuracy : D1

$$\text{Percent Defective} = \left[\frac{D1}{R} \times 100 \right]$$

$$\text{Accuracy Index (A)} = \left[1 - \left(\frac{D1}{R} \right) \right]$$

Note : Percent Defective is a quantifier used extensively in statistical quality control.

v) Number of defects (cases of loss of accuracy) detected : D
Number of records examined : R

$$\text{Defects/Losses of accuracy per record} = \frac{D}{R}$$

$$\text{Accuracy Index (A)} = \left[1 - \left(\frac{D}{R} \right) \right]$$

It may be mentioned that defect denotes accuracy violation, i.e., presence of error, and hence the absence of accuracy. Ratios based on defects/errors can be converted into accuracy ratio by the transformation:

$$\text{Accuracy Ratio} = 1 - \text{Defect (i.e., Error) Ratio.}$$

Understandably, notion of accuracy quantified as above has many issues not considered here. What if correct value of the identified source is undefined, or simply unknown. And of course what if data/information is say a name or has an alphanumeric value or is a video image; how is error or defect defined then ?

2.3 Consistency

Consistency is with respect to a set of constraints. As pointed out earlier, data/information is said to be consistent with respect to a set of constraints if it satisfies all constraints of the data/information model [AT&T Publication, 1992]. 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 salary level and salary attributes in the entity for a particular employee).

Given the number of constraints specified (CS) and the number of constraints for which error/defect detected in the sense constraints are not satisfied (CE), consistency can be quantified as follows [Svanks, 1984]:

$$\text{Consistency (C)} = \left[1 - \left(\frac{\text{CE}}{\text{CS}} \right) \right]$$

2.4 Reliability

Finally, Reliability (R) may be considered as an accuracy with which the information obtained represents the data item in whatever respect the information system processed it. For this purpose, a model may be considered where any processing of data has a large error component, random in nature. As a result, volume of error in the processed data will be different each time the data processing is repeated, leading to significantly different information in each case; thus reflecting a low reliability of the information. Thus 'Reliability' refers to the extent of existence of random errors in an information, or in other words, the degree of consistency with which an information can be repeated, without any intervening or additional instruction.

Coming to the quantification of reliability (R), in any data/information model, for an entity (i), the value (v_i) for an attribute or processed value for ith data item for the entity may be expressed as $v_i = t_i + e_i$, where ' t_i ' is the true component of the value and ' e_i ' is the error component. It is assumed that :

- (a) v_i takes values on a real line,
- (b) e_i 's are distributed independently and randomly over the whole population of data items (i's) and that $\bar{e}_i = 0$, and
- (c) e_i 's are uncorrelated with t_i 's.

Then reliability 'R' is given by :

$$V_e$$

$$R = 1 - \frac{V_e}{V_v}$$

where

$$V_v = \frac{1}{N} \sum_{i=1}^N (v_i - \bar{v}_i)^2$$

is the variance of the processed value and

$$V_e = \frac{1}{N} \sum_{i=1}^N (e_i - \bar{e}_i)^2$$

is the variance of the error component.

It follows from the above that reliability “R”, also termed as “Coefficient of Reliability” or “Reliability Index”, will have a value between 0–1.

It is appreciated that it may not be possible to repeat every data processing. In such case internal consistency of a data/information set comprising (a) information from processed data, (b) information from relevant identified source, (c) information from another related database, (d) results from relevant computational algorithm, etc. could be studied to obtain reliability.

Various methods exist for calculating the Reliability Index (R); Analysis of Variance (AOV) technique being one such. Choice of a method would depend on advantages, disadvantages and convenience of application in a given situation, while accounting for factors like nature of available data, form of data and computation aids available for processing.

2.5 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. Let W_a , W_c and W_r represent significant weightages for the integrity attributes accuracy, consistency and reliability, respectively, for the application area under consideration. These weightages may take values between [0-10].

Consider a user using the above information system for the application at hand. Let the Information Integrity attribute indices as observed at the user end in this specific example be : Accuracy (A) = 0.78, Consistency (C) = 0.55 and Reliability (R) = 0.85. Then Information Integrity Profile from the user end can be represented as follows :

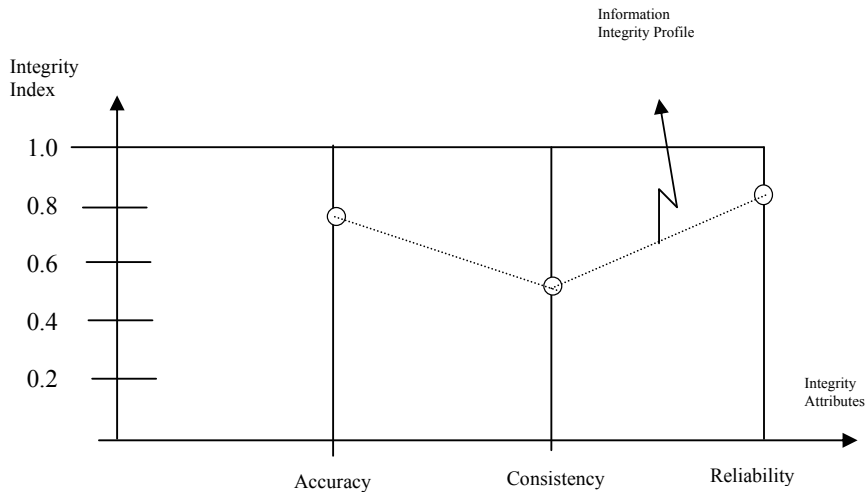


Figure 2 Information Integrity Profile

2.6 Cumulative Information Integrity Index (CIII)

Let Information Integrity attribute, depending on the range in which the attribute index value falls, be assigned a 5-point scale as shown in Table 1:

Attribute Index Value Range	Scale	Points
[1-0.8]	H	5
(0.8-0.6]	G	4
(0.6-0.4]	F	3
(0.4-0.2]	E	2
(0.2-0]	D	1

Table 1 : 5 - Point Scale for Information Integrity Attributes

In the example under consideration, the Information Integrity attributes then have the scales and points as given below :

Attribute Index Value Range	Scale	Points
Accuracy Index (A)	G	4
Consistency Index (C)	F	3
Reliability Index (R)	H	5

Then with a view to quantify the overall Information Integrity Index for the given application by the user, a Cumulative Information Integrity Index (CIII) may be given by :

$$CIII = \frac{4W_a + 3W_c + 5W_r}{W_a + W_c + W_r}$$

For example, if $W_a = 6$, $W_c = 5$ and $W_r = 8$, then

$$CIII = \frac{(4 \times 6) + (3 \times 5) + (5 \times 8)}{6 + 5 + 8} = \frac{79}{19} = 4.158$$

CIII will thus take a value between [1–5]. It could be the situation that this value of CIII may be low from the user point of view and the user may be requiring minimum CIII value of 4.25. Further, user may want to improve CIII with additional requirement of Consistency Index having minimum “G” scale. It is to achieve this integrity improvement that the user would then need to incorporate Information Integrity Technology.

Before one proceeds with further development of Information Integrity Technology Product structure, a word of caution is warranted here. The quantification of integrity attributes is not a trivial task even when it is possible [Redman, 1992]. Quantifiers suggested above do not bring out the complexity involved. In respect of accuracy quantification, it is already mentioned that there could be a problem of correct value of the identified source (also called standard) being undefined, or being simply unknown. In a situation, an assumed standard itself may be incorrect as is often the case with data gathered some time in the past and with no corroborating evidence. In yet another situation there may be more than one correct value. Then there is a problem of how to quantify accuracy if the value does not lie on a real line, i.e.; it is not a numerical.

As regards to consistency quantifiers, it is a relatively simpler concept than accuracy. Even then, it can assume complexities when all real database inconsistencies are to be measured (and which will be the need). Coming to the reliability attribute, as already mentioned, reliability quantifier gives an index of an accuracy with which the information obtained presents data item in whatever way the information system processed it. There can be no one way of calculating the reliability index and there will always be a need to develop one based on the nature of available data, form of data and computation aids available for processing.

Finally, exercise undertaken herein considers problem of quantifying integrity attributes when errors made but not corrected are at the level of information item on value (v). But in data/information modeling exercise, errors can be present even at view defining level itself. Specifically, data/information modeling exercise begins with modeling facts observed from the real world in the form of “view” which consists of one or more structures called entity types, from where one builds attributes, their domains and, in the end, values; thereby forming the triple <e, a, v>. What if errors are present at the stage of “view” defining stage itself and which invariably is the case. How does one define and quantify integrity attributes in such case? Further, so far what all one has discussed is only in terms of intrinsic integrity attributes. Depending on the context and the nature of use of information, one will also have to similarly develop methods for defining and quantifying extrinsic integrity attributes.

All these areas then constitute further research needs in the context of integrity attribute quantifiers for integrity improvement.

3 INFORMATION INTEGRITY TECHNOLOGY IMPLEMENTATION STEPS

With a suggestion for Cumulative Information Integrity Index (CIII) as above, within the framework of Information Integrity attributes of Accuracy (A), Consistency (C) and Reliability (R) argued, one can then identify Information Integrity Technology Implementation steps as follows:

- i) Understand the user application of the computerized information system under consideration.

ii) Establish organizational standard pertaining to data/information visavis requirements of : accuracy, consistency, reliability and cumulative integrity, based on application area and study of organizational practices.

iii) Study data/information flow through the Information System and define database(s).

Note : Apart from knowing how the Information System processes the data and apart from understanding more about the “noise” in the system, the study would also necessitate knowing wherefrom , how and data/information of what integrity flows into the system.

iv) Develop the Information System Model as in Figure3, based on understanding of data/information flow in the system for the identified database.

v) Specify and document data rules, also known as edits, to be implemented to study accuracy and consistency of the data/information

vi) Choose a method for calculating Reliability Index , keeping in view advantages, disadvantages and convenience of application while accounting for factors such as nature of available data , form of data and available computation aids.

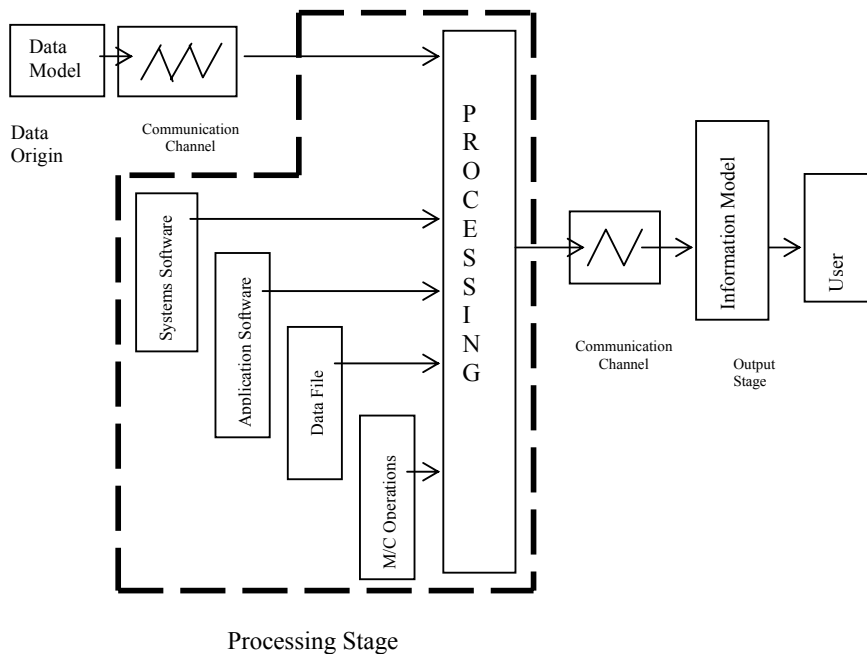


Figure 3 Data/Information Flow Model for an Information System for Implementing Information Integrity Technology

vii) Develop Integrity Analysis Software for analyzing intrinsic Information Integrity attributes of accuracy, consistency and reliability.

In addition, Integrity Analysis Software may also undertake statistical analysis (time series analysis and other techniques) of error patterns signifying irregular changes which contribute to loss of Accuracy and Consistency and of causes which contribute to loss of Reliability. This in turn leads to the development of :

- a) a filter to detect error or cause that occurred sometime in the past at time $(t - \tau)$,
 - b) an estimator to estimate error or cause that occurred in the immediate past at time (t) ,
and
 - c) a predictor to predict error or cause that may occur sometime in future at time $(t + \tau)$.
- viii) For Data/Information Flow Model in Figure 3, select a data sampling point at the output of a subsystem (or at an intermediate point within the subsystem), as close to the beginning of the Information System as possible.
 - ix) Depending on how data arrives at the sampling point (continuously or in batches), develop a continuous or batch processing sampler (a sampling program) to randomly select a sample of records arriving at the sampling point. Along with sampling records, the sampler program should also select some identifier of the sampling point and record of the data and time of sampling.
 - x) Following the selection of a sampling point and development of a sampler, select points for maintaining audit trail for sampled records.
 - xi) These points for maintaining audit trail may be selected at points at the output of subsystems (or at intermediate points within the subsystems) following the sampling point.
 - xii) Once the points for maintaining audit trail for records sampled are identified, develop a Sampled Records' Audit Trail (SRAT) program to separate or pull out (at the points selected) the audit records.
 - xiii) Ensure that sampler program and SRAT program are developed in such a way that they can download sampled records and records for audit trail as in (ix) and (xii) into a database to be set up (see (xiv)).
 - xiv) Accordingly, download the sampled and audit trailed records on mainframe or minicomputer or on personal computer/workstation so as to set up an Error Detection Database, based on hardware and software considerations and number of sampled and audit trailed records..
 - xv) Using the Integrity Analysis Software developed in (vii), analyze the Error Detection Database to :
 - a) identify data rule violations in respect of accuracy and consistency attributes,

- b) establish degree of integrity of data/information in respect of Information Integrity attributes of accuracy and consistency, based on data rule violation statistics
 - c) obtain reliability index for the database along with analysis of factors which contribute to the level of reliability,
 - d) develop Integrity Profile and Cumulative Information Integrity Index, based on indices of accuracy, consistency, and reliability attributes, and
 - e) study changes in database not expected, i.e., irregular changes.
- xvi) Compare the Integrity profile and indices obtained as in [(xv(b))–(xv(d))] with : standards in (ii) – local, regional, national, international as the case may be – and the user specifications on Integrity, so as to know what is expected of Information Integrity Technology. This would also facilitate ordering or ranking of the Integrity attributes from the point of which attribute needs maximum improvement effort.
 - xvii) For each of the Integrity attributes of accuracy and consistency, then, further analyze irregular changes either by subsystem or by field (in that order of priority of choice) and locate separate Integrity improvement opportunities at each of appropriately identified pairs of a given field at a given subsystem.
 - xviii) Similarly locate reliability improvement opportunities at each of the subsystems based on reliability factor analysis in (vii).
 - xix) Having located pairs of a given field at a given subsystem, each for improvements of accuracy and consistency and having located given subsystems for reliability improvement opportunities, further analyze the Error Detection Database and study irregular changes at each of pairs corresponding to accuracy and consistency attributes and study reliability factors at each of the subsystems, so as to understand over the time error patterns and causes contributing to loss of accuracy, consistency and reliability.

This would then facilitate detection of error or cause that occurred sometime in the past ($t - \tau$), or estimating error or cause at time (t), or predict error or cause that may occur at a future time ($t + \tau$).
 - xx) Now develop Information Integrity Improvement Action Plan for locations identified in respect of integrity improvement opportunities based on assessment as in (xvi) of integrity improvement target and based on the understanding of error patterns and factors for loss of intrinsic Information Integrity attributes as in (xix). This Integrity Improvement Action Plan may comprise restructuring subsystem(s) previous to the point of occurrence of error, improving integrity of data origin stage, improving communication channels, etc.
 - xxi) Finally, study performance of the Information System on incorporation of the Information Integrity Technology as outlined above. Accordingly obtain 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 vide [xv(b)], [xv(c)] and [xv(d)], so as to quantify integrity improvement achieved and to check if it is as per customer expectation.

4 THE INFORMATION INTEGRITY TECHNOLOGY PRODUCT

The Information Integrity Technology product thus implemented would then be a SOFTWARE PRODUCT consisting of :

- the user data rules list for error detection
Note : Data rule is that which must hold true in an Information System
- the Integrity Analysis Software for :
 - Accuracy
 - Consistency
 - Reliability
 - Integrity profile for the Information System
 - Integrity Indices
- the sampling program
- the Sampled Records' Audit Trail (SRAT) program
- the program for
 - statistical analysis of errors/causes for loss of integrity
 - factor analysis for reliability
 - time series analysis
- the program for :
 - detecting errors/causes (filter program)
 - estimating errors/causes (estimation program)
 - predicting errors/causes (predictor program)
- the generation of Error Detection Data Base
- the reporting based on analysis of Error Detection Data Base in terms of:
 - errors and causes detected; their locations in the Information System and their significance
 - error and cause patterns and trends obtained through statistical techniques such as time-series analysis
 - detection, estimation and prediction of errors and causes
 - identification of Integrity improvement opportunities
 - deciding and implementing Information Integrity Improvement Action Plan for Integrity Improvement opportunities identified (probabilistic action plan as also manual action plan included)
- obtaining improved Integrity Profile and Index
- documentation :
 - data rules list encoding the specifications for the Integrity Analysis
 - software and the reporting facility. This calls for user interaction.
 - the individual program in accordance with the systems and program documentation within the user organization
 - operating instructions for each program
 - program maintenance and test procedures
 - training material for users

5 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.

The users of computerized information systems have to undertake this computer house-keeping to incorporate Information Integrity Technology in their information systems, so as to avoid serious potential losses occasioned by errors that were made (due to factors external to application control) but not corrected. In concrete terms, Information Integrity Technology will be an application and user - specific software which, for an Information Flow Model as in Figure (3), samples on-line, periodically and systematically, records arriving at an appropriately chosen point, follows or keeps track of sampled records at subsequently identified points through the information system and stores the records so sampled and obtained through follow up (audit trail), to set up error detection database. Information Integrity Technology then analyzes this error database to identify errors, i.e. changes not expected and to quantify resulting loss of integrity therefore, so as to develop Integrity Improvement Action wherein Information Integrity opportunity is identified and implemented.

Understandably, this Information Integrity Technology will have to be developed in a computer language compatible with the information processing environment of the user organization. This calls for organizational IS planning, devising policies, standards, and guidelines pertaining to data. If this is not ensured, net result is non-compatible, and unshareable data/information. An important step in the implementation of Information Integrity Technology, therefore, is data rule specification, defining data rule standard, which is also needed for undertaking Information Integrity Analysis.

Yet another area that calls for standards pertains to degree of integrity. As mentioned, the application area would influence the degree of accuracy, consistency and reliability. The application area would also influence values of W_a , W_c , W_r . Further, quantification of Integrity attribute such as accuracy calls for identification of data/information sources and their standards, i.e. correct values. In implementation of Information Integrity Technologies, it would therefore be necessary to establish these application area specific standards representing requirements of degrees of integrity as also of values of Integrity attribute significance factors.

Finally, it is important to appreciate that the development of standards as above would facilitate implementation of Information Integrity Technology products for different subsystems of the information system as also for the total system. This would call for support of reputable software developers and vendors for the purpose. Further, these Information Integrity Technology products would cover data/information in various forms – numerical, alphabetic, alphanumeric, video-images or any other – and that too for different application areas. This would open a new vista in terms of design, development, commissioning, operation and maintenance of data technologies, hitherto not attended, for ensuring on-line integrity of computerized information system.

6 REFERENCES

1. Ameen, D.A.(March 1989) *Systems Performance Evaluation*, Journal of Systems Management, pp. 33–36.
2. AT&T Publication (1992) *Data Quality Foundations*, Published by AT&T Quality Steering Committee, USA.
3. Mandke Vijay V. (1996) *Research in Information Integrity : A Survey and Analysis*, Proceedings of the JNCASR and SERC Discussion Meeting at IISc Campus, Bangalore on

- Information Integrity – Issues and Approaches, Edited by Rajaraman V. and Mandke Vijay V., published by Information Integrity Foundation, New Delhi, India.
4. 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.
 5. Mandke Vijay V., and Nayar M.K. (1998) *Design Basis for Achieving Information Integrity – A Feedback Control System Approach*, IFIP TC 11 Working Group 11.5 Second Working Conference on Integrity and Internal Control in Information Systems, Edited by Sushil Jajodia, William List, Graeme W. McGregor and Leon A.M. Strous, Published by Kluwer Academic Publishers
 6. Nayar M.K. (1996) *A Framework for Achieving Information Integrity*, Proceedings of the JNCASR and SERC Discussion Meeting at IISc Campus, Bangalore on Information Integrity – Issues and Approaches, Edited by Rajaraman V. and Mandke Vijay V., published by Information Integrity Foundation, New Delhi, India.
 7. Rajaraman V. (1996) *Information Integrity – An Overview*, Proceedings of the JNCASR and SERC Discussion Meeting at IISc Campus, Bangalore on Information Integrity – Issues and Approaches, Edited by Rajaraman V., and Mandke Vijay V., published by Information Integrity Foundation, New Delhi, India.
 8. Redman T.C. (1992) *Data Quality : Management and Technology*, Bantam Books, NY.
 9. Svanks Maija I., (1984) *Integrity Analysis : A Methodology for EDP Audit and Data Quality Assurance*, EDP Auditors Foundation, Inc.
 10. Wang R. and Strong D. (1993) *An empirical Investigation of Data Quality Dimensions : A Data Consumer's Perspective*, TDQM-93-12, The TDQM Research Program, MIT, Sloan School of Management, Cambridge, USA.
