7/25/68

A SYSTEMS APPROACH
TO INTELLIGENCE DATA STORAGE AND RETRIEVAL

A THESIS

Presented to

The Faculty of the Division of Graduate

Studies and Research

by

Thomas Austin Wilson II

In Partial Fulfillment

of the Requirements for the Degree

Master of Science

in the School of Industrial and Systems Engineering

Georgia Institute of Technology

August, 1972

A SYSTEMS APPROACH
TO INTELLIGENCE DATA STORAGE AND RETRIEVAL

Approved:

Paul Gray, Chairman

Donovan Young

Terrence Connolly

Date Approved by Chairman: 6-1-72

ACKNOWLEDGMENTS

I would like to thank all those who helped me in the preparation of this thesis.

In particular, I am grateful to Lt. Robert E. Hightower, Miss Ann Searles, and Mrs. Judith Williamson of the Georgia Bureau of Investigation. Lt. Hightower provided invaluable assistance in developing the experimental model shown in this thesis, and Miss Searles and Mrs. Williamson devoted much of their own time and effort in data collection and model evaluation.

My sincere appreciation goes to Dr. Paul Gray, chairman of my thesis committee, whose guidance, advice, and encouragement have added immeasurably to the quality of this research. Thanks are due also to Dr. Terry Connolly and Dr. Donovan Young, members of the reading committee, for their assistance in the research and writing.

A very special note of thanks is due to Sue, Kacy, and Lori--my wife and daughters. Without their patience, love, and understanding, this thesis could not have been written.

TABLE OF CONTENTS

TABLE OF CONTENTS (continued)

LIST OF TABLES

LIST OF ILLUSTRATIONS

Figure Page

SUMMARY

Lack of development of efficient file management procedures has been identified by many investigators as a major cause of problems in the storage and retrieval of intelligence data. This research develops proposed procedures for (1) screening out irrelevant intelligence data to prevent its entry into the files, (2) updating information in the data base, (3) purging obsolete entries, and (4) integrating new data into the files.

Using a systems engineering approach, the problems are examined in the context of police intelligence systems, and methods are proposed for their solution. Application of the proposed methods involves the development and use of a weighted, additive scoring model to determine the relative utility of file entries and candidates for entry. An experimental model was developed and tested using actual data and personnel of the Georgia Bureau of Investigation. The tests show the model to be an appropriate solution and easily implementable.

The proposed methods are applicable to both manual and automated systems and are attainable with available equipment and routine software. Modifications which may be required in a computer-based system are presented and discussed.

CHAPTER I

INTRODUCTION

Background

In spite of the advances in information technology, managers continue to seek more complete and timely information for use in planning and decision making. This continuing demand has resulted in the philosophy that information and intelligence systems must be designed to provide an infinite pool of data from which managers can extract the specific data they desire. Accordingly, the emphasis in the design of these systems is placed on the generation and collection of data.

There is, however, an increasing feeling that a central problem in management's search for information is not the lack of pertinent information for planning and decision making; it is the receipt of too much irrelevant information (1). This feeling has aroused an interest in the problems associated with file management. Specifically, the problems of interest are (1) screening out irrelevant data to prevent its entry into the data pool, (2) updating information in the data base, (3) purging obsolete entries, and (4) integrating new data into the system.

These problems are applicable to information and intelligence systems of all types. Their solution requires a series of decisions regarding the value of information.

Purpose

The purpose of this research is to examine intelligence systems by using a generalized systems engineering approach similar to that presented by Hall (17), Goode and Machol (15), and Engstrom (13). Having examined an intelligence system, modifications will be proposed to alleviate some of the problems related to file management.

The specific objective of this research is to develop methods to:

1. Alleviate the file management problems pointed out in the previous section.

2. Routinize the file management procedures to minimize the need for managerial decisions.

Scope of the Study

This research is concerned only with the flow of material which has been sent to the intelligence system. Although the collection of data and the use of the resulting intelligence are vital to the system, these problems will not be addressed. The boundaries for the research are, at one extreme, receipt of data by the intelligence system and, at the other, the purging of material from the system because of loss or initial lack of intelligence value.

Interactions between the intelligence system and other systems are discussed only as they directly affect the management of the intelligence files. Generally, these interactions either take form of constraints or of factors that affect the value of material in the system.

One further limitation on scope is that indexing (that is, the classification of information by key words and categories) will not be addressed. This research will concentrate on methods that are independent of the type or extent of indexing used in the storage of intelligence data.

Procedure

Chapter II outlines the pertinent literature regarding information/ intelligence systems and decision-making techniques. Chapter III begins by defining an intelligence system and develops a cycle which illustrates the major functions performed in this system. The police intelligence system is chosen as an exemplar and is examined using a systems engineering approach. Its weaknesses are discussed, and methods are proposed to eliminate them. These methods are synthesized, and a framework is given for integrating them into the existing system.

In Chapter IV, various decision-making techniques are analyzed with respect to their use in the methods discussed in Chapter III. A suggested design procedure is given for developing a decision model for use in an intelligence system, and an experimental model is shown in the Appendix. Details of its design and test results are presented in

Chapter VI.

Chapter V points out some modifications that may be required if the system is to be automated, and the type of equipment that could be used in a computer-based intelligence system is discussed.

Conclusions drawn from this research are presented in Chapter VII.

CHAPTER II

LITERATURE SURVEY

Introduction

This literature survey will outline only the areas of particular importance to this research, viz., intelligence systems and decision-making techniques. In examining intelligence systems, it is necessary to include some basic literature on information systems because the fundamentals are the same in both areas.

Information/Intelligence Systems

Information systems have been defined in a variety of ways. For example, Murrish (30) defines an information system as one that provides for the collection of internal and external information in a form accessible to managers for use in planning and control decisions. Prince (31) states that a management information system (MIS) must provide relevant data for decision making and be capable of implementing changes made by management. The United States Army defines military intelligence as knowledge acquired by collecting, evaluating, and interpreting all available information concerning an area of interest and states that this intelligence provides the basis for the plans and estimates of the commander (2). Whisenand and Tamaru (37) characterize an information system as being oriented to the collection, analysis, and utilization of data to promote interaction among personnel, machines, and procedures for planning and decision making. Schultz and Norton (35) classify intelligence as the product resulting from the collection, evaluation, and interpretation of information which concerns criminal activity and which is significant to planning.

The theme common to all of these definitions is that information is provided to managers to assist in their planning and decision making. This theme prevails whether the objective is a business information system, a military intelligence system, or a police intelligence system.

A second theme common to all of these definitions is the differentiation among data, information, and intelligence. Eilon (9) declares that data is the raw material from which information is produced. Symonds (41) says, "Data becomes information only when it has been collected, analyzed, and presented in such a form that it results in the communication or receipt of intelligence." Heaney (19) and Williams (39) also identify this difference between data and information whereas Whisenand and Tamaru call data the raw material for information processing (37). The Army classifies information as the raw material from which intelligence is produced (2), and the Law Enforcement Assistance Administration (LEAA) says, "The intelligence process is a continuous series of interrelated activities directed toward converting raw information into informed judgments" (3).

The third common theme is the existence of a cycle or process that describes the functions around which the information system is organized. Symonds (41) identifies a process which includes the functions of collection, analysis, and dissemination. The Army uses a four-stage cycle which includes (1) collection planning, (2) collection, (3) processing, and (4) dissemination and use (2). The steps involved in processing, as defined by the Army, are shown in Figure 1. Schultz and Norton (35) describe an intelligence process which includes collection, processing, use of intelligence, and guidance of the collection effort. The steps involved in processing according to this cycle are shown in Figure 2. The LEAA describes a five-stage process: (1) collection and evaluation, (2) collation, (3) analysis, (4) reporting, and (5) dissemination. Collation includes screening the information, arranging it in an orderly manner, and storing it.

These cycles are compared in Figure 3. The figure shows a high degree of commonality between the cycles, but it is apparent that certain terms have conflicting meanings in different cycles and that some cycles are more comprehensive than others. For example, the cycles described by Symonds and the LEAA do not account for frequent reassessment of the planning of the collection effort; the LEAA cycle is the only one which directly accounts for systematic storage of information. Although the term analysis is used in three of the cycles, the meanings are not equivalent.

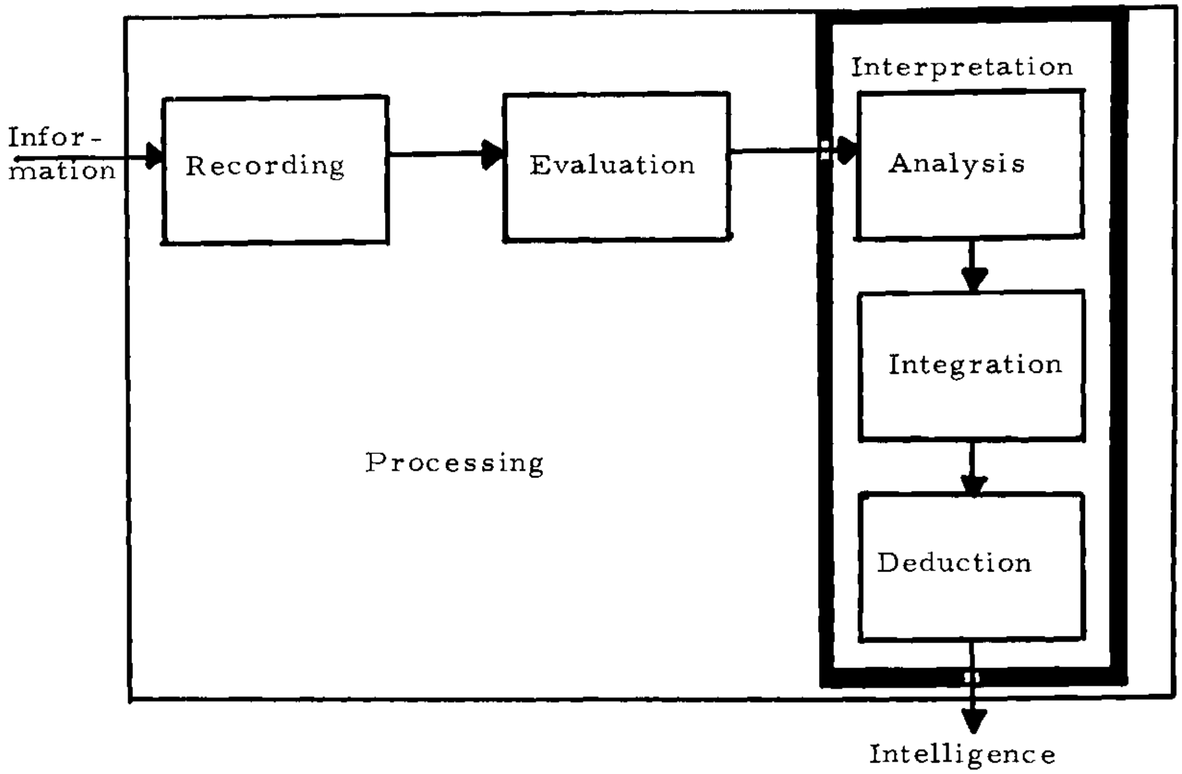


Figure 1. Processing Information into Intelligence--the Army Intelligence Cycle.

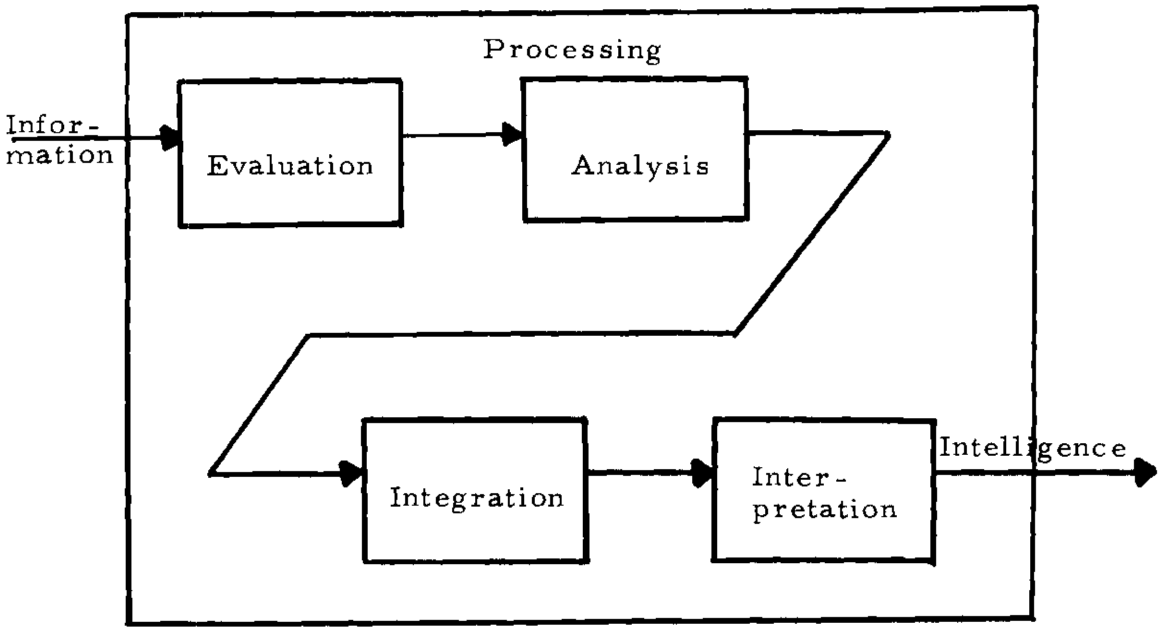


Figure 2. Processing Information into Intelligence--Schultz and Norton.

Process Function	Symonds	U.S. Army	LEAA	Schultz & Norton
Planning the collection effort	-----	Collection planning	-----	Guidance of the collection effort
Collection of data	Collection	Collection	Collection and evaluation	Collection
Processing	Analysis	Processing a. recording b. evaluation c. interpretation 1. analysis 2. integration 3. deduction	Analysis	Processing a. evaluation b. analysis c. integration d. interpretation
Use of the intelligence	Dissemination	Dissemination and use	Dissemination and reporting	Use
Storage	-----	-----	Collation a. screening b. arranging c. storage	-----

Figure 3. Comparison of Information/Intelligence Cycles.

In Chapter III, an intelligence cycle that adopts the best characteristics of those described above will be developed, and definitions of the terms will be given so that further confusion may be avoided.

In addition to conflicting terminology and incomplete governing processes, other problem areas related to information systems require attention.

In a summary of studies dealing with command information systems, Ringel (33) points out that their input data are received from different sources and vary in content, form, and completeness. Such data require extensive handling and processing and must be screened for relevance to the requirements that the information system is designed to satisfy. Yet the problem of screening has received little attention (33). Williams (39) also pointed out the need for a selection process that will filter out information which does not fit the user's requirements, and Ackoff (1) states that managers using a MIS may suffer from an overabundance of irrelevant information. He further states that one of the most important functions of an information system is filtration, yet "the literature on MIS's seldom refers to (this) function, let alone considers how to carry (it) out" (1).

Other problems are also pointed out by Williams (39). These include purging the system, updating information in the system, and integrating new information into the system.

Information loses value over time and should be discarded or devalued so that its continued presence does not interfere with the retrieval of more valuable information (39). The result of failing to do this was illustrated in a study of the field interview files of the Los Angeles Police Department. Conroy (5) found that there had been no meaningful purging of the files and that satisfactory information could not be obtained because of the volume of the files and the length of time required to search them. This problem exists in most intelligence systems because of a general reluctance to accept the responsibility for purging data. The problem is further complicated by the absence of any set procedures or policy for determining which records should be kept and which should be destroyed (37).

Kelly (24) establishes that the problems associated with file update become significant when the update includes the integration of new records into the file and the addition of information to existing records. He also states generally that there is a need for file organization in information systems.

Decision-Making Techniques

Decision making involves (1) the identification of the possible alternatives, (2) a prediction of the results of each alternative, (3) a set of priorities or preferences, and (4) the selection of the preferred alternatives (i.e., the decision). With the rapid expansion of computer technology and the increased awareness of cost accounting, the problem of decision making has received great attention in the professional literature of many fields. The scope of the literature dealing with decision making runs the gamut from highly complex mathematical decision systems to simple, one-step rules.

Much of the literature deals with the development of models for decision making. These models fall into four categories: (1) scoring models, (2) economic models, (3) constrained optimization models, and (4) risk analysis models.

Application of the latter two types of models requires a large amount of data and computation time. Since the objective of this research is to develop methods which are readily adaptable to both manual and computer-based intelligence systems, these categories were considered unsuitable, and no extensive research was conducted regarding their use. The other two categories are discussed below.

Economic Models

Economic models generally involve Decision Analysis as typified by the work of Howard (21, 22, 23) and Raiffa (32). It is concerned primarily with the economic impact of decisions. In this category, one of the best known methods for dealing with decision problems is the decision tree. Raiffa's (32) development of this method is a logical sequence of steps which requires only a fundamental knowledge of probability.

Howard's (21, 22, 23) approach to decision making is more sophisticated mathematically, but it is similar to that of Raiffa; i.e., by obtaining expected values of outcomes, the value to the decision maker of information that would reduce the uncertainty can be computed (21). Howard extends this concept to show that availability of information affects the probability density function of profit (22) and shows how statistical decision theory can be used in the solution of systems engineering problems (23).

Other widely publicized methods include those of Hertz (20), Dean (7), and Solomon (36) which provide rankings of sets of stimulus objects based on economic characteristics such as rate of return (20), minimum attractive rate of return (7), and average rate of return (36).

Scoring Models

Scoring models are concerned with the decision process itself; i.e., by what process does an individual reach a decision, and what decision model will approximate the results obtained by a consistent decision maker? Much of the work in this category has come from the field of psychology.

Einhorn states that "one of the main problems in decision making has been the attempt to understand how individuals assess the utility of each of a set of stimulus objects when each stimulus can be evaluated in terms of a set of multidimensional attributes" (10). He has studied the use of nonlinear models to approximate the decision process and has achieved excellent results (10, 11). He has also shown that the use of nonlinear models to approximate the decision process is dependent on the type of task and the amount of information presented to the decision maker (11).

Although Einhorn's experimentation shows that his conjunctive and disjunctive models often give a better approximation of the actual decision process than does the linear model, a great amount of evidence has been gathered to show that if decisions are averaged over time or across individuals, the results are approximately linear. Yntema and Torgerson (40) have shown that linear models are an excellent approximation to data even when there are nonlinear relations present. This has also been the conclusion of Goldberg (14), Wiggins and Hoffman (38), and Hammond and Summers (18); "the simple linear model appeared to characterize quite adequately the judgmental process involved" (14).

The use of a model to approximate the decision process makes it possible to use as input information that has been evaluated subjectively by experts. The input information is combined "mechanically" by the model. Yntema and Torgerson recognized the need for this man-machine cooperation early in the use of computers as decision-making tools (40). More recently, Sawyer stated that "the clinician may be able to contribute most not by direct prediction, but rather by providing, in objective form, judgments to be combined mechanically" (34).

The distinction that must be made is that, even with expert opinion, the global judgment is a subjective combination of components which frequently are also subjective judgments. Einhorn's recent studies ". . . argue for the quantification of the components of the judgments as well as the global judgments themselves" (12).

The application of the above principles to practical, modern decision problems has been most notable in the field of research and development (R&D) and, more specifically, in the selection of R&D projects. This class of applied decision models has been termed scoring models and has received attention in the literature on operations research, engineering management, and management science.

The advantages of scoring models are that (1) they are the only models that account for the explicit inclusion of subjective factors, (2) they use simple and inexpensive methods of data acquisition, and (3) they allow the decision maker to determine the weight of each of the factors in arriving at a decision (27).

Scoring models achieve the same purpose as the decision tree of Raiffa and the economic models used by Dean, Hertz, Howard, Solomon, et al. However, application of the scoring models eliminates the need for the detailed analysis used in economic models, substituting expert clinical judgment of the components of the global decision for piece-by-piece evaluation.

Mottley and Newton (29) applied a multiplicative model to evaluate proposed research projects on the basis of five criteria whereas a weighted, additive model was used by Dean and Nishry (6) for the same purpose. This latter model was tested on a sample of six projects, and the resulting ranking was identical to that yielded by a profitability model. The chief contribution by Dean and Nishry was the use of expert personnel to determine the criteria weights used in the model. Several drawbacks to the scoring model are also pointed out in this study, but it is believed that these are overcome by careful attention to the design details formulated in more recent studies by Moore and Baker (27) and in the application by Goodwin (16).

Moore and Baker compared a scoring model, a profitability model, and a linear programming model to determine whether scoring models can be constructed which are consistent with the other decision models and, if they can, to determine what properties of the scoring model provide this consistency. Their results showed that "it is possible to construct a scoring model which exhibits a high degree of rank-order consistency with other models of project selection" (26). Other results were that (1) the additive scoring model produced better results than the multiplicative model, and (2) inter-model consistency is dependent on both the effective range over which the criterion is measured and the ability of the model to distinguish between levels of performance (26).

Based on the results presented in (26), a method for scoring model design and verification was developed by Moore and Baker. The method consists of eight steps beginning with selection of the criteria and culminating with complete specification of the model; the discussion of each step provides clear, workable guidelines for the designer (27). In a recent study of evaluation of alternate designs, Goodwin (16) used essentially the method outlined by Moore and Baker and achieved satisfactory results.

With slight modifications required by the different environment, the method developed by Moore and Baker is used in the development of this thesis.

CHAPTER III

SYSTEM DEVELOPMENT

Introduction

For the purpose of this research, an intelligence system is defined as any system which transforms data and information into knowledge concerning an area of interest useful in planning or decision making. The cycle describing the functioning of an intelligence system may be thought of as a regulatory feedback system as shown in Figure 4. This cycle includes the five major functions defined in the various processes discussed in Chapter II; viz., (1) planning the collection effort, (2) collection of the data, (3) processing the data into intelligence, (4) dissemination and use of the intelligence, and (5) collation for future use.

A set of intelligence needs or requirements are evaluated, and a plan is formulated for the collection of data. Based on this plan, requests (or orders) are issued to collection agencies who then gather the data and submit it to the requester. This data is processed (as shown in Figures 5 and 6) into intelligence and disseminated for use in planning and decision making. This intelligence can then be compared to the initial intelligence needs so that a new set of requirements can

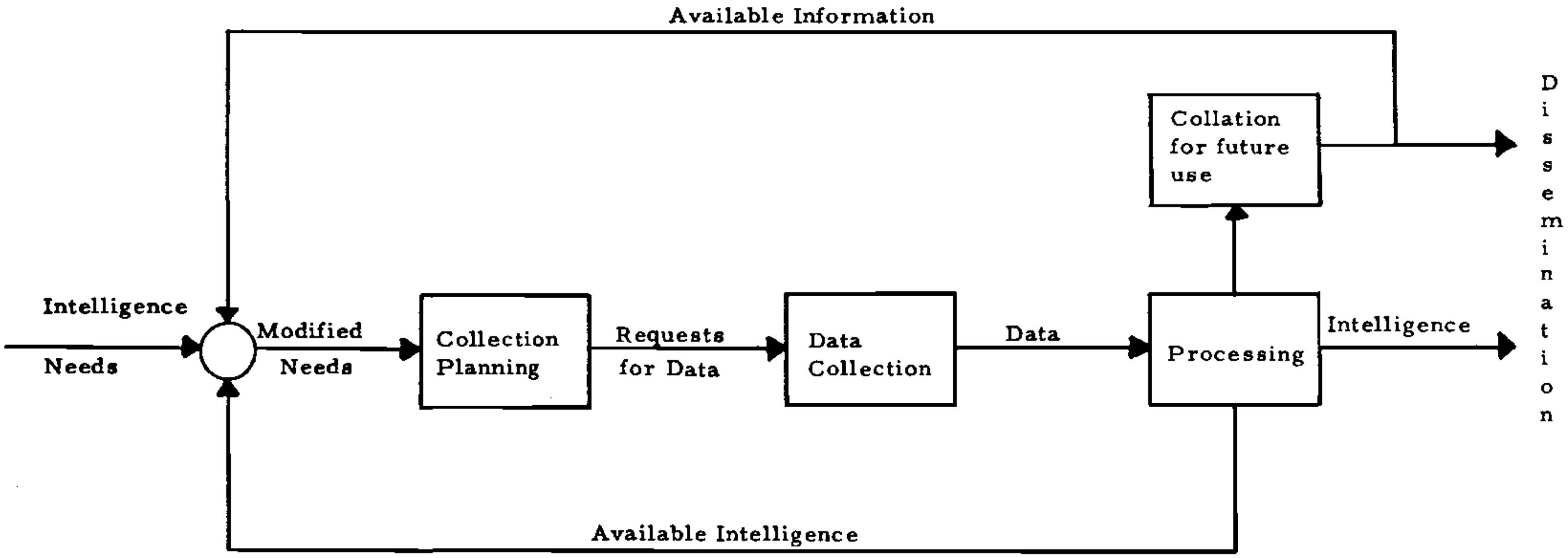


Figure 4. The Basic Intelligence Cycle.

be established. It is also collated for future dissemination and the satisfaction of further intelligence needs.

Processing

Figure 2 illustrates the transformation of data into intelligence, and Figure 3 shows the steps involved in evaluation. The following definitions of the terms involved in these figures are provided for added clarity.

Data. Any facts, documents, materials, or actions (or their symbolic representations) that may be recorded.

Information. Reduced data, i.e., data which has been summarized.

Intelligence. Knowledge concerning events or courses of action that have occurred, are occurring, or may occur.

Recording. The reduction of data to writing or some other form of symbolic representation, if this has not already been done at the time of collection.

Analysis. Determining the meaning of the data when considered by itself.

Integration. Combination of new information with other known information or intelligence.

Interpretation. Determining the significance of new information in relation to other known information or intelligence.

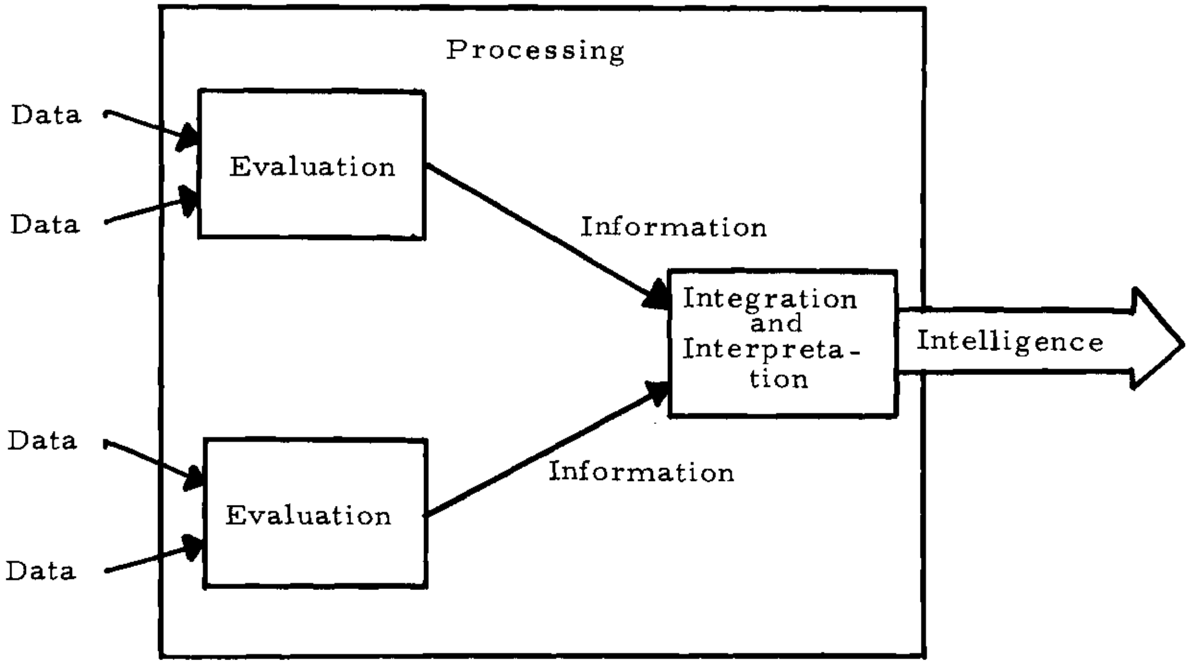


Figure 5. Processing--the Transformation of Data into Intelligence.

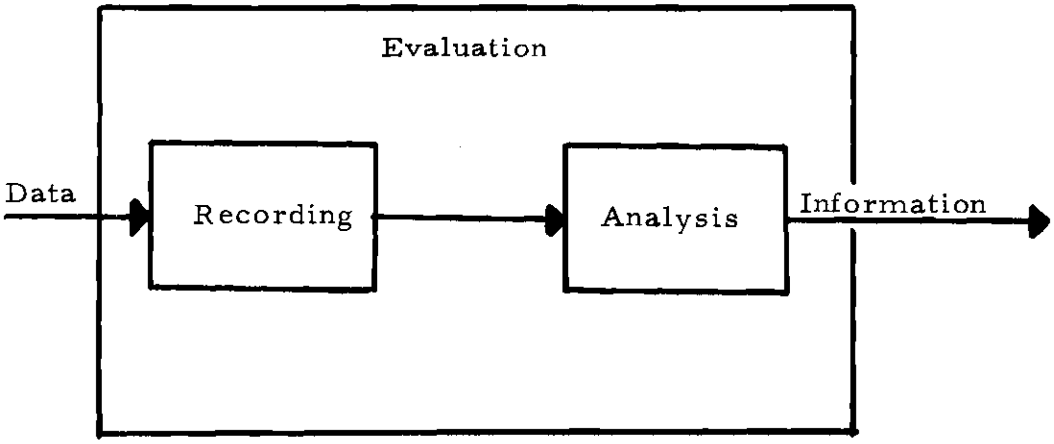


Figure 6. The Function of Evaluation.

Collation

The function of collation consists of three separate steps--screening, arranging, and storage. Screening is the filtering of data, information, and intelligence to remove those items which either contain errors, are duplicates of items already stored in the system, are irrelevant to the system, or fail to meet the standards required for storage in the system. Arranging is the ordering of the material which is to be retained in the system so that it may be easily retrieved for further use.

Storage is the actual placement of material in the intelligence files and the creation of the index entries corresponding to the material. The intelligence files may contain raw data, information, and intelligence; however, for the sake of simplicity, the contents of these files will be referred to variously as items, elements, material, and information. In this sense, information is a generic term for data, reduced data, and that part of intelligence which is not knowledge.

The structure of the files will vary depending on the nature of the intelligence system. The methods presented here are independent of the structure.

System Description

The basic intelligence system is essentially an information flow system with decision points at various locations in the flow. It has, as its primary objective, the provision of information to aid managers in their planning and decision making. Operation of the system is governed by the cycle illustrated in Figure 4.

The number of sub-systems in an intelligence system will generally not exceed five. These five sub-systems correspond closely to the major functions defined in the intelligence cycle; they are (1) planning and control, (2) collection, (3) processing, (4) administration, and (5) storage. The exact number and composition of the sub-systems depend on the size and complexity of the intelligence system. Large, complex systems have five distinct sub-systems, whereas smaller, less complex systems will consolidate some of the functions.

Inputs to the systems are data which come from different sources, vary extensively in content, form, and completeness, and require extensive handling and processing. Outputs may take the form of intelligence estimates, probability statements, predictions of future events, or identification of alternatives. The specific nature of both the inputs and outputs depends on the system environment.

To clarify the system description further and to decrease the degree of abstraction, police intelligence systems will be used as an examplar for the remaining discussion. Placing the intelligence system in this environment permits identification of the inputs to the system. These include (1) reports from sources such as intelligence unit investigators, informers, other units of the same agency, liaison officers, federal agencies, and other law enforcement agencies; (2) data from public records, newspapers, and business records; (3) technical data such as fingerprints, blood analyses, and paint samples; and (4) "hard" data such as weapons, vehicles, and clothing.

Some typical outputs of a police intelligence system are reports of developing criminal activities, reports relating to cases under investigation, activity reports on known or suspected underworld figures, and transactions of pawn shops.

An analysis of the environment of the police intelligence system identifies some of the factors which influence its operation. The intelligence system affects, and is affected by, the scope of police operations, public opinion, the extent of criminal activity in the locale, and the degree of centralization and automation of the police organization and related systems. These factors also affect the constraints on the system--the size of the intelligence unit, the degree of mechanization/automation, the territory in which the unit is effective, the extent of cooperation with the public and related agencies, and the available budget.

The objective of a police intelligence system is to provide intelligence that will assist in police planning and decision making.

Weaknesses of Existing Systems

A police intelligence system may have all the weaknesses pointed out in Chapter II for information/intelligence systems in general. The system presently used for finding outdated or valueless material is to review each file entry and reassess its worth. This process requires a skilled analyst, and in a large file, requires a great amount of time. Since time is usually scarce, valueless items are removed from the file infrequently, and then only on a random basis.

In addition to information which has lost its value, intelligence files are overburdened by the entry of a large amount of irrelevant information. Currently there is no methodical procedure for uniformly screening incoming material to ensure that irrelevant items are excluded from the files.

Integration of new material into the files is a problem from two points of view: indexing and assessment of relative value or utility of data. In a large file, it is impractical to reassess every item in the file so that the relative value of new material may be determined. However, there are no other operational methods for accomplishing this task.

The police intelligence system is dynamic. Data are received constantly which affect the value of material already in the files, and this change in value of the stored material affects the entire intelligence cycle. However, the change in value of the elements in the file is reflected only in the subjective evaluation carried in someone's mind. The environment of the intelligence system is also dynamic; the occurrence of new crimes and changes in priorities affect the value of elements in the file. Again, this change in value or utility is expressed only in the subjective evaluation of someone judging the files.

Another weakness of police intelligence systems is in decision making. Bristow (4) notes that the average American police administrator spends approximately 20 per cent of his time making decisions, and most of this time is spent on routine decisions that could be delegated to subordinates. File management decisions are of this type. If a decision rule or policy can be established by the head of the intelligence system which will permit subordinates to make decisions reflecting his priorities and preferences, then the administrator will have more time available for planning and for analysis of nonroutine problems which require a management decision.

In summary, the weaknesses of police intelligence systems are essentially the file management problems described by Kelly (24), Williams (39), and Ringel (33). They include:

1. Inadequate purging of information no longer of value.
2. Insufficient screening to prevent irrelevant items from entering the files.
3. No practical procedure for integrating new material into the files based on its value to the intelligence system.
4. Time-consuming procedures for regular update of information in the file.

One additional weakness is the excessive requirement for management decisions.

System Objectives

The objective of this research is to propose modifications to the existing intelligence system which will mitigate the weaknesses just described. Specifically, the system must include some method(s) to accomplish the following:

1. Purge elements of little or no intelligence value from the file.

2. Screen all inputs and products of the system to prevent the storage of irrelevant or valueless data, information, and intelligence.

3. Integrate new material into the file in such a way that the relative utility of nearly all entries in the file is known.

4. Facilitate the update of information stored in the intelligence files.

5. Routinize the file management procedures to minimize the requirement for management decisions.

Ideally, the method(s) adopted to accomplish these goals should be adaptable to either manual or automated intelligence systems and should be attainable with minimal training and expenditure of funds.

Discussion

Use of a Utility Measure

To assess the relative value of new material in a large existing file, it would be necessary to evaluate every entry in the file. Concerning this problem, Williams (39) states that "analysis of the relative utility of items in a system and predictions of the potential utility of candidate items to the system, it is believed, can be accomplished by means of an analysis model."

It will be shown in this research that a quantitative measure can be obtained of the utility of each item in the file as well as each item being considered for entry into the file. This measure can be used in three ways: (1) as a discriminator to prevent the entry into the files of items having relatively low utility; (2) to order the entries in the file based on their relative intelligence value; and (3) if the file is full, to determine which item should be removed if new material is to be added.

Thus, this measure forms the basis for routine file management decisions. The process by which each of the above actions may be accomplished is explained in detail below.

During the processing stage of the intelligence cycle, a measure of utility, U, is associated with each input. A minimum level of utility, U_{min}, is established for entry into the file; any item with less than the minimum utility is excluded. Those items meeting the minimum utility criterion are indexed and entered into the file. The utility measure is also associated with the index entries and these are filed at the same time.

If total file size (i.e., maximum number of entries) is an active constraint on the system, the index can be reviewed to determine which file entries have the lowest utility, and these items can be removed to permit the entry of more important material. (In an implementation of this method, rules would have to be established to determine which entry would be removed in the event that several items were tied for the lowest value.) The sorting time required to locate the low-valued entries in the index can be reduced by using a system such as the McBee Keysort* card.

In Chapter IV, a decision model is discussed which will yield a measure of relative utility, and a recommended design procedure is given. An experimental model, developed to test the techniques described in this chapter, is presented in the Appendix.

File Update

Update of material in the files may be necessary for a number of reasons. These include (1) receipt of contradictory information, (2) receipt of confirming information, (3) changes in priorities, (4) routine retrieval for use, and (5) the forcing of information from the files by higher-valued material.

Update can be facilitated by identifying file material that is (1) related to new material (which may be contradictory, confirmatory, or supplementary), (2) affected by changes in priorities, (3) removed

---

*A brief description of this system is given by Williams (39, p. 269).

for routine review, or (4) forced out of the file by higher-valued items. Intelligence analysts may then review this material. At the time of this review, a new utility measure, which reflects its current intelligence value, may be associated with it.

Updating the intelligence files by adding new material has already been discussed in the preceding section.

Other Steps

Other steps necessary to alleviate the file management problems of police intelligence systems are the removal from the file of outdated information and the rapid identification of inputs which contain errors or are related to file material. The latter can be accomplished, in most cases, in a preliminary subjective screening by the intelligence analyst. At this time inputs containing errors can be discarded, and related file information can be identified for possible update.

During the processing stage of the intelligence cycle, the person doing the processing should estimate the expected length of time for which the material will have some value to the intelligence system. This time span is translated into an expiration date, t_{exp}, to be associated with each item which passes through the system and with its associated index entries. Then, whenever elements in the file have passed their expiration date, they are removed from the file. These elements may be identified for possible update to hedge against the possibility that some information may have increased in value or may still be current. If a review so indicates, a new expiration date can be assigned.

Again, the use of an implementation such as Keysort would reduce the time involved in locating expired file entries.

Routine Procedures

The intelligence system is an operation that involves considerable routine effort. As was pointed out previously, if decision rules can be established which represent the priorities and preferences (i. e., the utilities) of the head of the intelligence unit, then routine decisions can be delegated to subordinates. Therefore, proper assignment of utility measures aids in making file management decisions. As a result, the manager will have more time for planning and analysis of non-routine problems. This aspect of model design is considered in Chapter IV.

Summary

It is helpful at this point to list the actions intelligence personnel must take to implement the proposed system. Further clarification of the steps required will be given by presenting separate steps in a system flow representation and then synthesizing these steps into a single system representation.

Use of a Utility Measure

These steps are illustrated in Figure 7.

1. Develop and test a model to determine the relative utility, U, of items which have been processed through the system.

2. Apply the model to determine the relative utility of each item and associate this utility with the index entries corresponding to this item.

3. Compare the utility measure of the item with a minimum utility, U_{min}; if the item utility equals or exceeds the minimum utility, file the item; otherwise, dispose of it.

4. If the file is full, compare the utility of the candidates for entry with the utility of the lowest valued item in the file, U_{low}; purge the lowest valued item to create space in the file.

File Update

These steps are illustrated in Figure 8.

1. Identify pertinent material.

2. Retrieve this material from the files.

3. Review the material and, if necessary, reassess its utility.

4. Recycle the reevaluated material through the system and purge the rest.

Preliminary Screening and Purging of Obsolete Information

These steps are shown in Figure 9.

1. Conduct preliminary screening to identify data which contains errors or is related to information in the files.

2. Assign an expiration date, t_{exp}, and associate it with the item

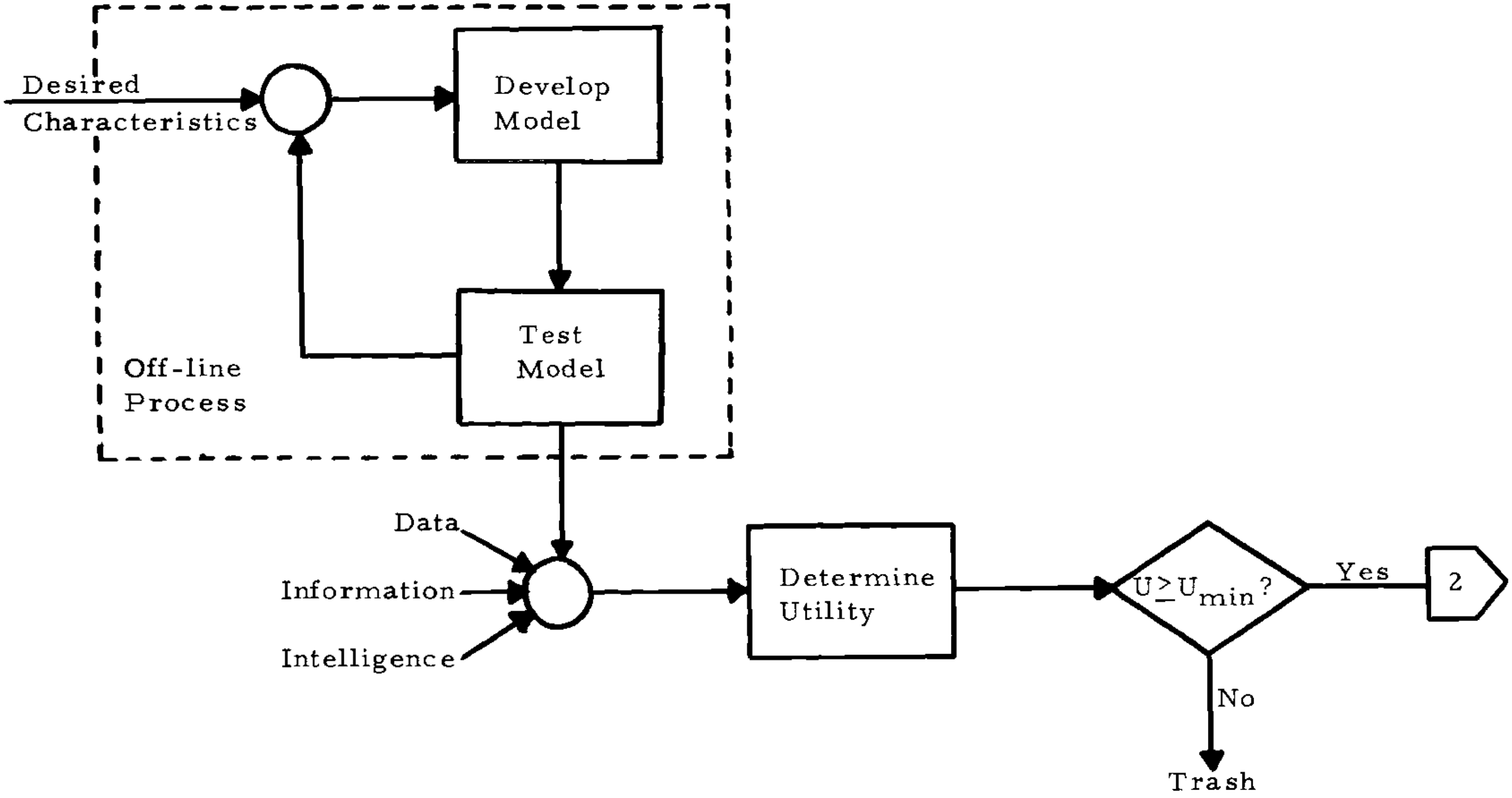


Figure 7. Use of a Utility Measure, Part 1.

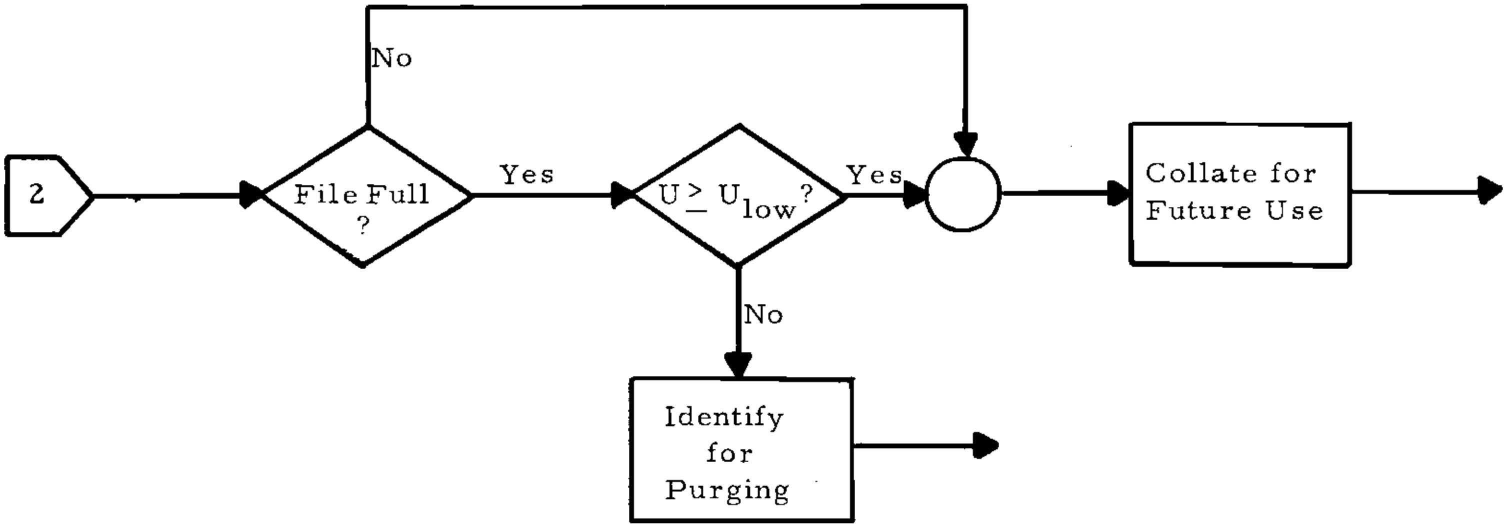


Figure 7 (continued). Use of a Utility Measure, Part 2.

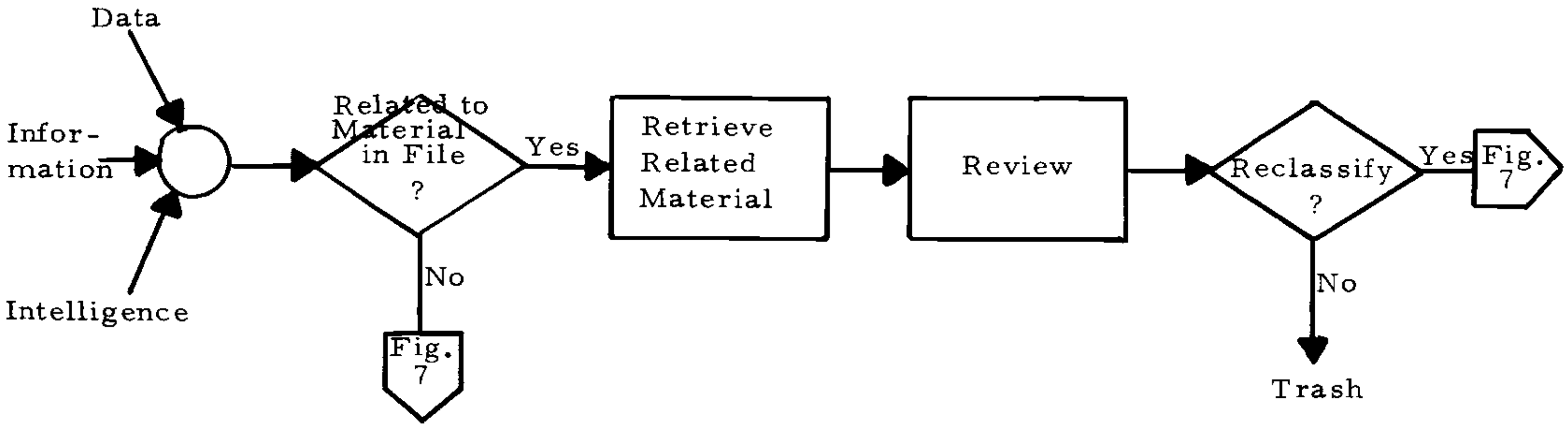


Figure 8. File Update.

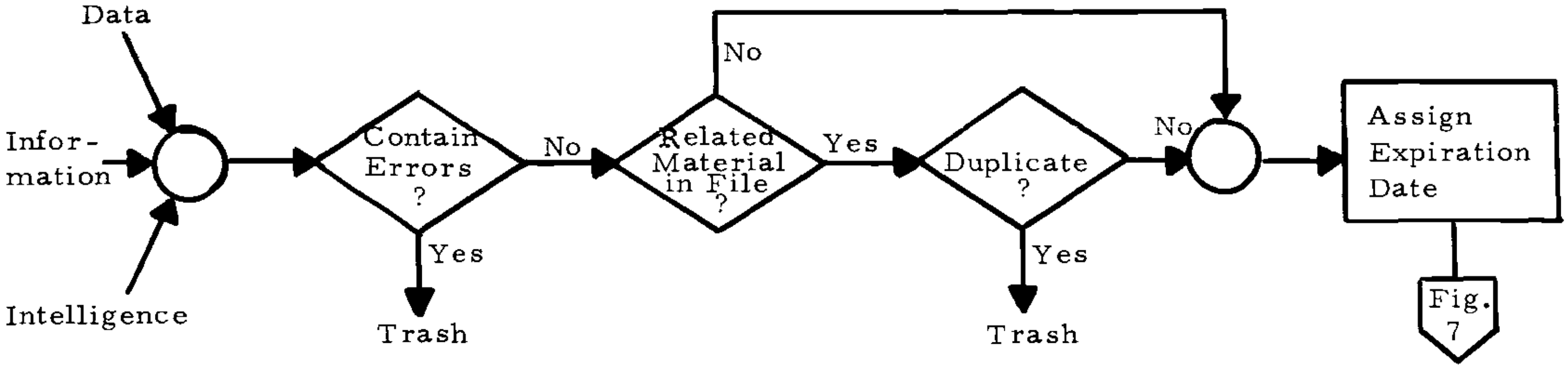


Figure 9a. Preliminary Screening Process.

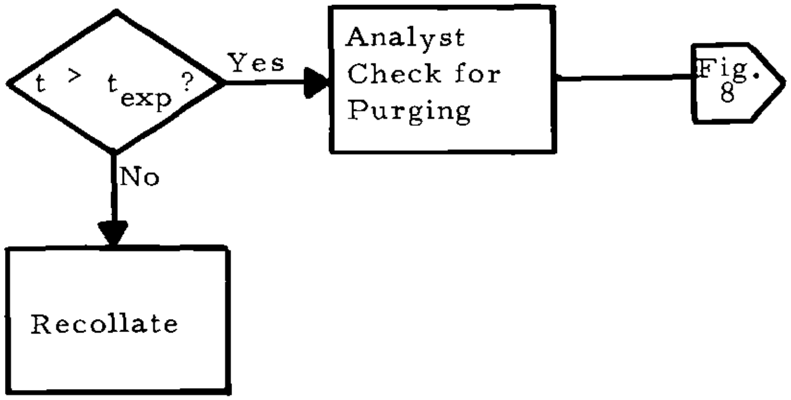


Figure 9b. Identifying Outdated Material.

and its corresponding index entries.

3. Periodically identify all file items which have passed their expiration date, i.e., $t > t_{exp}$.

4. Update expired entries if necessary.

<u>Synthesis</u>

When all of the above procedures are integrated into the existing system, the proposed system will be that shown in Figure 10. This system alleviates all of the problems discussed in this chapter, and it can be implemented in both manual and computer-based intelligence systems.

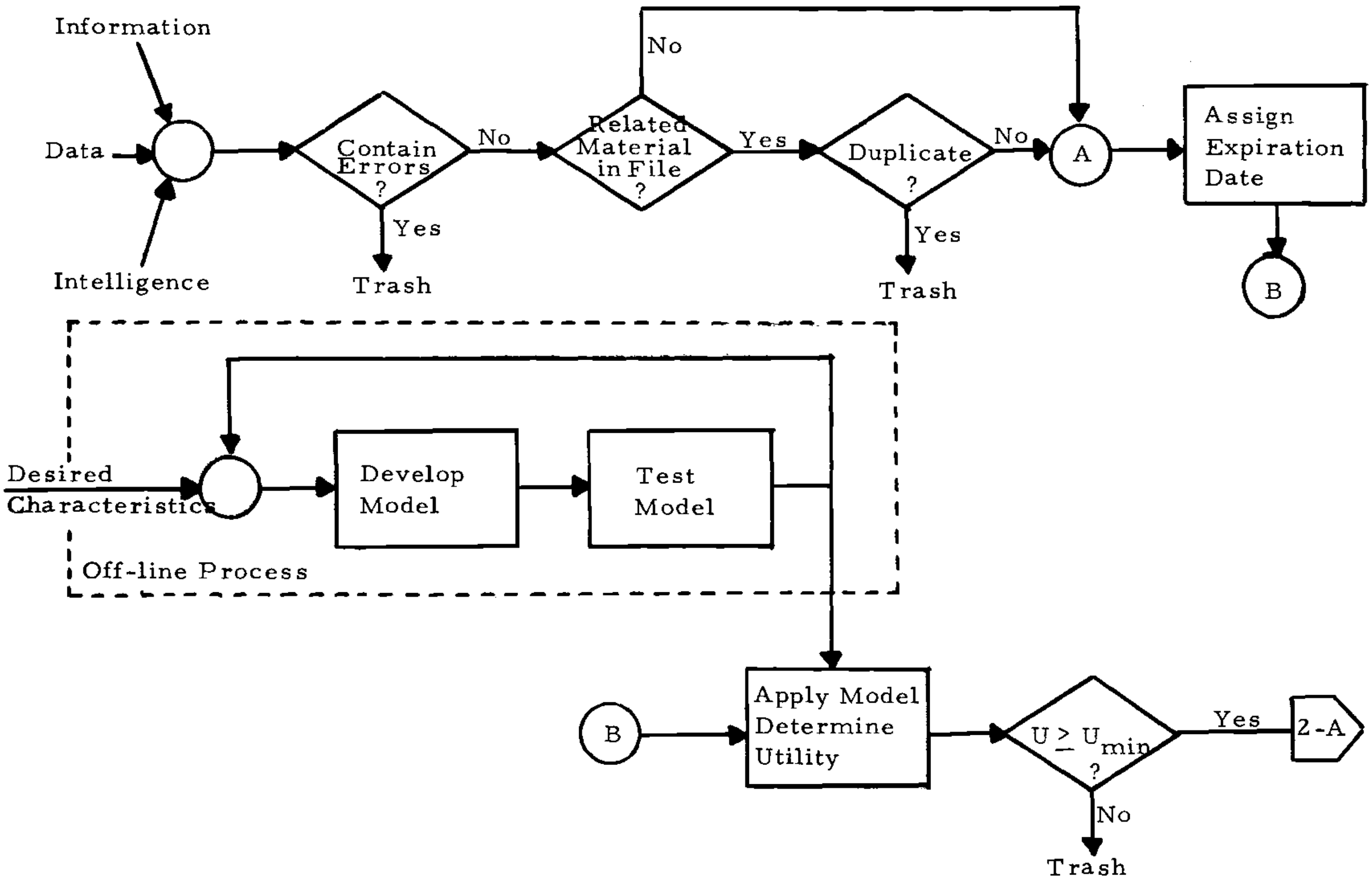


Figure 10. The Modified Intelligence System, Part 1.

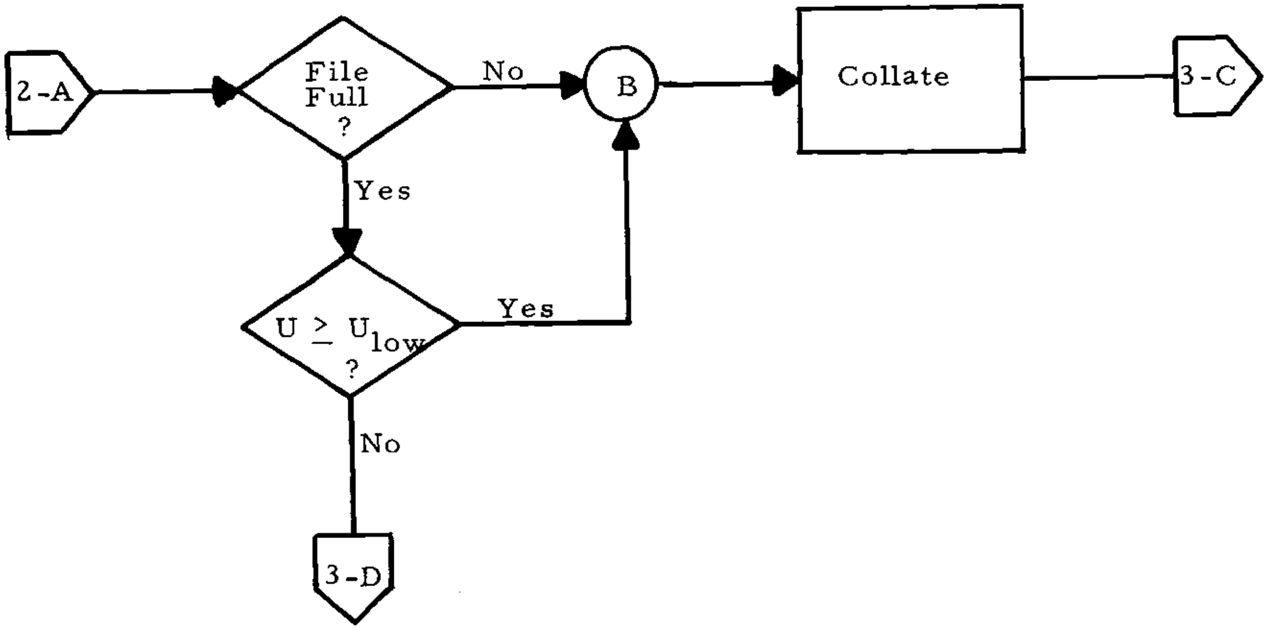


Figure 10 (continued). The Modified Intelligence System, Part 2.

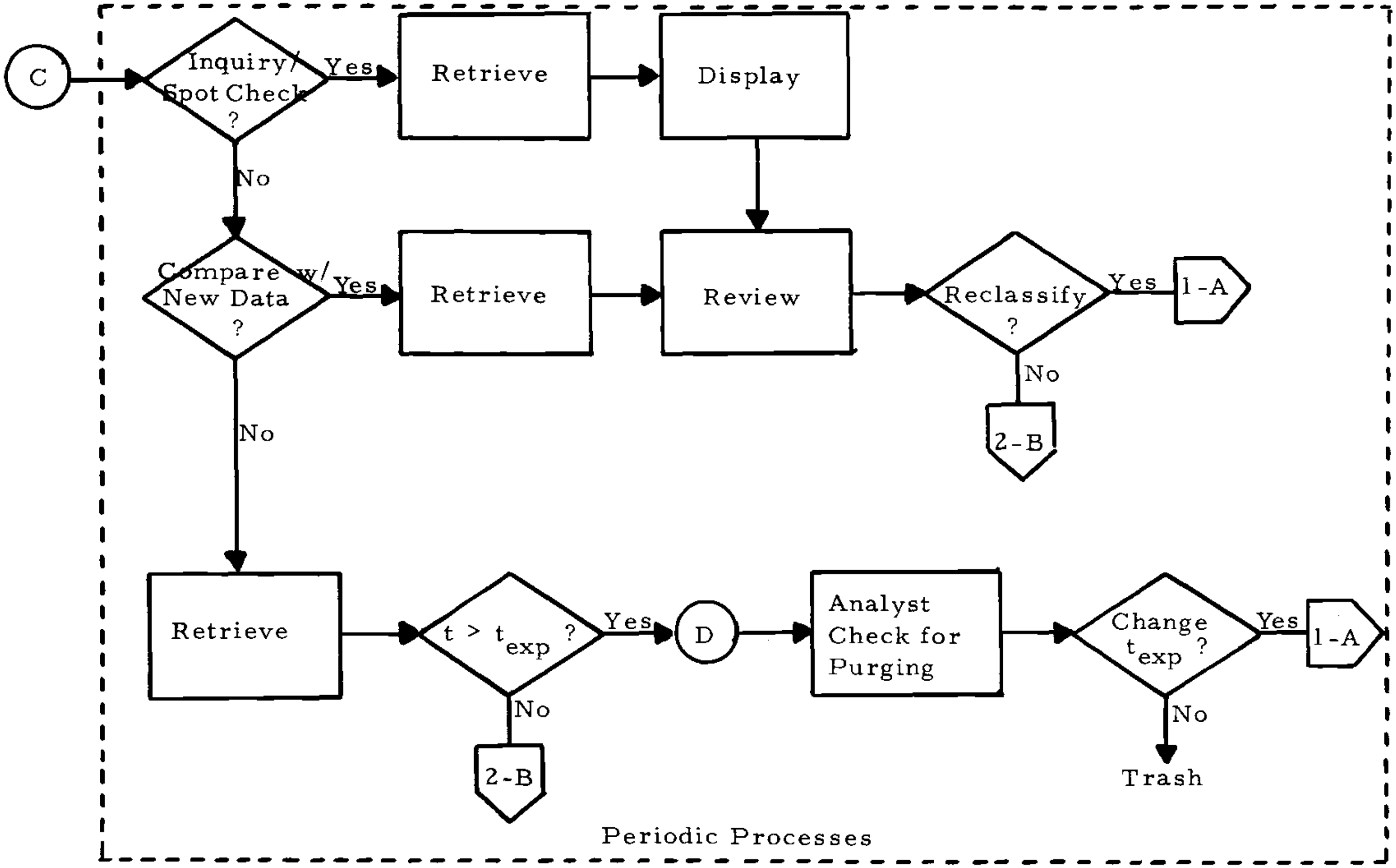


Figure 10 (continued). The Modified Intelligence System, Part 3.

CHAPTER IV

DECISION MODEL DEVELOPMENT

Introduction

There are essentially three approaches to decision making: (1) the use of intuition, (2) a clinical or subjective judgment based on a collection of available facts and data, and (3) the systematic collection of data to be used as input to a formal decision model. It is the last approach that is of interest in this research.

As described in Chapter II, there are four categories of decision models. These are (1) scoring models, (2) economic models, (3) constrained optimization models, and (4) risk analysis models.

Desirable characteristics of a decision model designed for use in an intelligence system are:

1. It must be easy to apply.
2. The user should be able to understand it.
3. The data requirements for its use should be minimal.
4. It should be inexpensive to implement.
5. It should be adaptable to both manual and automated systems.

In a given situation, all four categories of decision models may possess these characteristics; however, it was decided to use scoring models for the purpose of this research. This decision was made because the use of scoring models to arrive at a measure of the utility of information is morphologically adjacent (i.e., a logical next step) to the present practice of purely subjective evaluation of information.

Based on the literature survey concerning scoring models, the linear scoring model was chosen for use in this research.

Scoring Models

Scoring models compute an overall score for a piece of information based on ratings assigned for each relevant decision criterion. They are designed to accept subjective input data. Specifically, criteria which determine the relative utility of information are determined, and a rating scale is developed for each criterion. Each piece of information is then evaluated with respect to each criterion and given a rating. These ratings are then combined for each piece of information, and an overall weighted score is computed.

An experimental scoring model, developed as part of this research, is shown in the Appendix. The procedure for selecting criteria and criterion weights is explained in the following section.

Model Design

Moore and Baker have analyzed scoring models for R&D project selection (26) and have developed an approach to the design of scoring models (27). The design methodology outlined by them for the linear, additive model is essentially the approach that will be used here. Certain modifications are necessary because the environment of intelligence systems in general, and police intelligence systems in particular, is much less sophisticated mathematically than is the environment of R&D; these modifications will be incorporated freely as the need arises.

Since the actual models to be developed are a function of the specific system for which they are designed, this discussion of design will bring out the details which are considered pertinent to the design of scoring models for use in intelligence systems in general. Details of the design of the experimental model developed to test this system are presented in Chapter VI.

Criterion Selection

Prior to selection of the criteria, it is essential that the decision maker have a thorough understanding of the purpose of the decision model. He must then develop a list of criteria that affect his decisions.

It is desirable for the decision maker to consult with the designer during the selection process. Frequently, the designer can lend a measure of objectivity to the assessment of goals, and it may be necessary for him to redirect and refocus the thoughts of the decision maker during the selection process. This consultation also serves to ensure that the final list of criteria does not omit major factors, that each criterion is relevant and measurable, and that overlap between criteria is minimal.

An extensive list, in which no important factors of evaluation are neglected, will ensure that the model gives a close approximation of the decision maker's preferences and will facilitate tuning the model to give a better approximation of the desired outcome. (Tuning will be discussed in a later section.)

The relevance of each criterion should be challenged before it is placed on the final list. Although the accuracy of the model may be increased if the list of criteria is more complete, it should also be recognized that, as the list of criteria grows, the cost and complexity of data acquisition and processing grows. The decision maker's preferences regarding the trade-off between completeness and relevance can be satisfied through consultation between the decision maker and the designer.

Each criterion must be measurable; i.e., it must be possible to construct a scale for rating the extent to which the criterion is satisfied.

Criterion overlap should be avoided where possible; this will minimize the possibility of overweighting the importance of a particular factor in the evaluation process and will facilitate the evaluation of trade-offs between criteria. Generally, overlapping criteria can be combined into a single factor or redefined so that there is no overlap.

Criterion Weights

A weight will be assigned to each criterion to indicate its relative importance in evaluating an item of information and establishing utility of the item. These weights will reflect the utilities of the decision maker and define the trade-off rates between criteria.

There are numerous methods for determining the weights to be assigned to multiple criteria. In one of the most widely quoted studies on weighting, Eckenrode (8) compared six of the most frequently used methods; viz., ranking, rating, two methods of partial paired comparisons, complete paired comparisons, and successive comparisons. These six methods were compared by Eckenrode in three different judgment situations (8).

The results of this study were that "the various methods were equally reliable for collecting such judgment data, but that one (ranking) was much more efficient in terms of the time required to use it than any other method" (8).

In a more recent study, Goodwin used a combination of three methods--ranking, successive comparisons, and rating--to arrive at the weights for his figure of merit (FOM) model. Goodwin altered Eckenrode's ranking procedure slightly to ensure that there are no zero weights (16). This is desirable. Zero weight indicates that the criterion is irrelevant, but a careful selection of criteria will not include any irrelevant criteria.

Since the time of skilled intelligence personnel is so valuable to the intelligence system, the results of Eckenrode's study indicate that either ranking or rating can be used to establish the initial criterion weights.

Ranking. The method presented here for obtaining weights by ranking is that presented by Eckenrode (8) and modified by Goodwin (16). First, the raw ranks are converted by the following formula.

$$R_{cj} = m - (r_{cj} - 1) \tag{1}$$

where R_{cj} = converted rank assigned to criterion c by judge j,

r_{cj} = raw rank assigned to criterion c by judge j,

m = number of criteria.

These converted ranks are then summed over the number of judges, n.

$$R_c = \sum_{j=1}^{n} R_{cj}. \tag{2}$$

The composite weight of criterion c across all judges, w_c, is then given by

$$w_c = R_c / \sum_{c=1}^{m} R_c. \tag{3}$$

Rating. The rating method of obtaining weights is accomplished by first having each judge rate each criterion on a scale of zero to ten.

The following formula is then applied.

$$w_{cj} = b_{cj} / \sum_{c=1}^{m} b_{cj} \tag{4}$$

where w_{cj} = weight computed for criterion c based on the rating given by judge j,

b_{cj} = rating given to criterion c by judge j,

m = number of criteria.

Then,

$$w_c = \sum_{j=1}^{n} w_{cj} / \sum_{j=1}^{n} \sum_{c=1}^{m} w_{cj} \tag{5}$$

where w_c = composite weight for criterion c,

n = number of judges.

When this rating method was used by Eckenrode, it produced weights with the greatest mean deviation and the narrowest range. This is normally not a desirable characteristic for weights (8).

It is recommended that weights be obtained by both ranking and rating when possible--particularly in the case where the number of experts involved in establishing the weights is small. Consistency between the methods can be compared by any of the usual rank-order statistics, such as the Kendall Coefficient of Concordance (25) or the Spearman rank correlation coefficient. If necessary, further testing can be conducted to establish the final weights.

One note of caution is in order. Criterion weights will not remain constant; they will change as environmental forces change. Therefore, the operating intelligence system must include a procedure for periodic revision of the weights.

Initial Specification of the Model

Since nearly all criteria used in the evaluation of intelligence data are subjective, it is necessary that initial specification of the model be done in close cooperation with the decision maker.

Moore and Baker recommend the use of a scale on the closed interval $[1, 9]$ and define their scoring function by assigning an integer value to each scoring interval (27). However, for models such as this one, which incorporate only subjective data, the measurement scales are continuous, and occasionally, experienced decision makers can differentiate ratings of criteria more accurately than to one significant figure. Accordingly, it is recommended that the initial specification use a continuous scale on the interval $[1, 9]$ with certain points on the scale defined, but the user should not be restricted to use of these defined points in his use of the model.

The same interval should be used for all criteria since any other scheme would counteract the weights assigned in the previous step. For example, a criterion evaluated on the scale $[0, 4]$ would automatically receive less weight than one rated on the interval $[1, 9]$.

An example of a scale is shown in Figure 11.

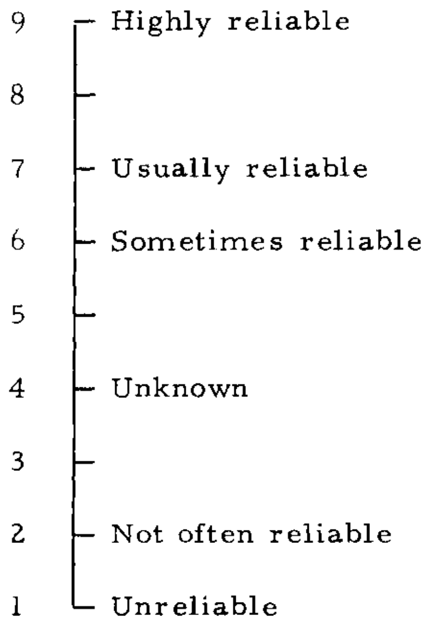


Figure 11. Example Scale for Judging Reliability of the Source.

It is possible that an experienced decision maker could differentiate the reliability more accurately than to one significant figure; therefore, if his ability and experience permit him to assess reliability as being equal to 3.5, for example, he should be allowed to do so. Permitting the user to rate the criteria in this way has the effect of providing finer differentiation between the utility measures of different items of information.

The defined points on the scale should be chosen carefully, as improper placement may give undue emphasis to a particular criterion. An excellent example of this is provided by Goodwin in his study on evaluation of alternate sub-system designs. In this study, he used a scale on the interval [0, 10]; for the criterion of performance values, 5.0 was the value initially assigned for meeting basic specifications. Evaluation of this scale by experts revealed that this caused excessive emphasis to be placed on exceeding specifications. The scale was then redefined so that the interval [9, 10] was used for those alternatives which met or exceeded the basic specifications, and the interval [0, 9) was used for rating other acceptable alternatives (2).

Since this model may be used by personnel other than those who initially defined the criteria, initial specification of the model must include an accurate description of the criteria. This will permit consistent results from the model regardless of who is using it.

Benchmark Determination

Since it is the results of the decision-making process that the model is designed to approximate, the benchmark or basis of comparison for the model rankings should be the subjective rankings of the decision maker. It is against this set of preferences that the validity of the model will be tested.

Initial Verification

This step is really nothing more than testing of the model to ensure a satisfactory level of agreement with the benchmark. The following sequence of operations is recommended for the initial verification.

Step 1. Testing should be done with samples large enough to satisfy the assumptions of the Moore and Baker design process. Typically, 15 samples would be sufficient. The data is ranked subjectively by the decision maker to provide the benchmark and is then evaluated in the scoring model by the formula

$$S_j = \sum_{c=1}^{n} w_c v_{cj} \tag{6}$$

where S_j = score assigned to data piece j by the user of the model (note that S_j is a dimensionless number which measures the relative worth or utility of the information to which it is assigned),

w_c = weight assigned to criterion c by the user of the model,

v_{cj} = value assigned to criterion c for data piece j by the user of the model.

Data is then ranked according to the score received from the scoring model.

Step 2. A correlation analysis is used to determine how closely the results of the scoring model approximate the benchmark. This analysis can take the form of rank-correlation statistics or concordance statistics; Kendall's coefficient of concordance is frequently used for this purpose.

Step 3. The measures of intermodel consistency derived in Step 2 for each set of data should be averaged over several sets to avoid possible effects of chance groupings within a set.

Step 4. If the average level of intermodel consistency is not satisfactory, then the designer must alter the model, and the verification must be repeated. If the consistency is acceptable to the designer, then no further action is necessary unless he wishes to conduct some form of sensitivity analysis.

Tuning the Model

Alteration of the model to achieve closer correlation between model results and the benchmark is called tuning. In the basic model proposed here, there are two areas that can be investigated--the criterion weights and the scales used to evaluate the criteria. The first step in tuning should be to alter the criterion weights until reverification yields a level of consistency satisfactory to the designer. If altering the weights does not achieve this, then the scales used to evaluate the criteria should be adjusted by redefining points on the scales.

The designer must specify the satisfactory level of consistency between the model results and the benchmark. Given the day-to-day inconsistencies of a subjective ranking procedure, it seems that a correlation coefficient of 0.90 indicates a satisfactory level in this case.

If there is still a poor fit between the model rankings and the benchmark after adjusting the weights and the scales, it is likely that the initial model will not work. In this event, the designer has three options: (1) discard the initial model and select new criteria, (2) add more criteria to the initial model, or (3) go to a nonlinear model.

For a more rigorous mathematical discussion of this process, the reader is referred to the five-stage process presented by Moore and Baker (27).

Sensitivity Analysis

Having achieved a satisfactory level of consistency between the benchmark and the model results, the designer may want to perform some type of sensitivity analysis on the model. One form of such an analysis involves varying the criterion weights according to some plan so that the degree of change in the model results may be related to the amount of change in the weights. Knowledge of this would permit the user of the model to determine if and when the model should be retuned.

Another type of sensitivity analysis can be achieved by comparing the value profiles for different items of information. Using these profiles, difference profiles can be constructed which graphically portray the advantages of one item over another. Figure 12 illustrates this procedure.

A third type of sensitivity analysis that can be performed is comparison for dominance of alternatives. This comparison will aid the decision maker in identifying exceedingly important data. An example is shown in Figure 13.

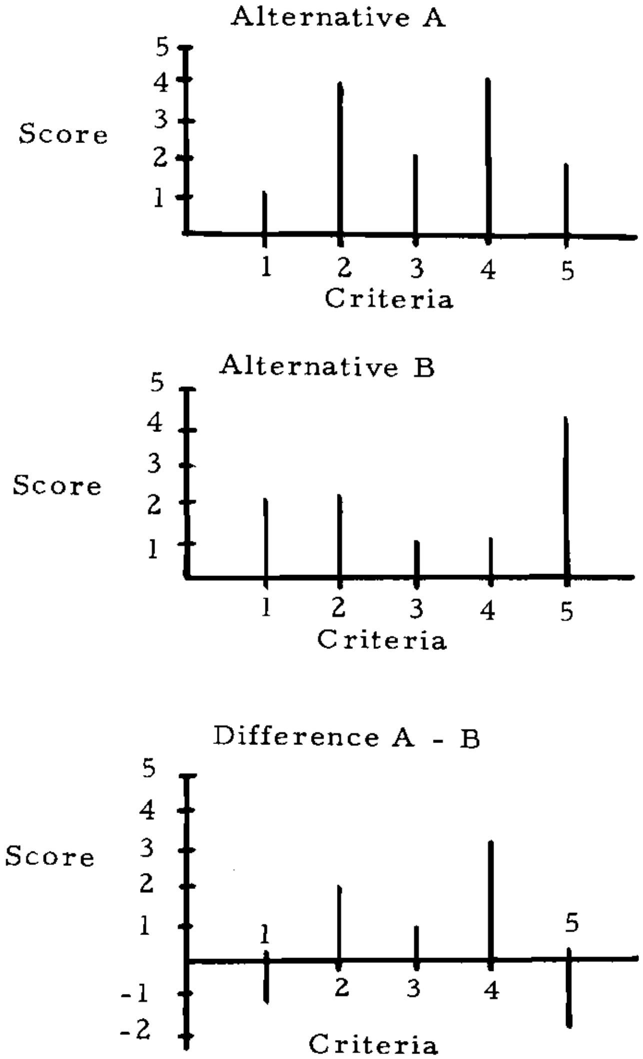


Figure 12. Constructing a Difference Profile.

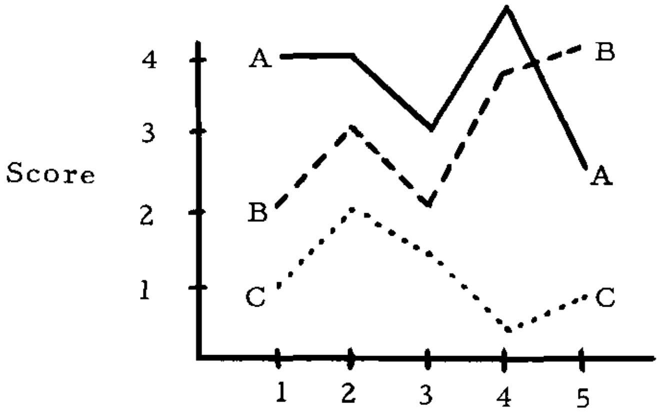


Figure 13. Dominance Graph.

In this example, both Alternatives A and B dominate Alternative C; thus, no further consideration need be given to Alternative C in the search for very important information.

Conclusion

This discussion of scoring models and scoring model design should provide sufficient background to allow the system designer to develop the pertinent details of a scoring model for use in an intelligence system.

CHAPTER V

AUTOMATING THE SYSTEM

Multi-Echelon Storage

The trend in the development of information/intelligence systems is to automation. The methods presented in Chapter III can be applied in an automated system; however, some modifications may be necessary.

Automated systems, which may or may not be computer-based, can include storage devices such as high-speed microfilm or microfiche retrieval systems. In computer-based intelligence systems, there are numerous methods of storing information. These include disc, tape, magnetic cards, drum, and perforated tape. Depending on the type of equipment used in the system, these methods have various retrieval times associated with them.

It is likely that the user of a system possessing multiple storage methods will wish to take advantage of the different retrieval times associated with each method. Frequently used material can be stored in rapid access devices, whereas material required less often can be placed in slower devices. If this technique is used, another problem is introduced into the file management procedures; namely, how to decide what level of storage is appropriate for each entry into the files.

In a large, complex intelligence system, many automated storage methods and several manual ones may be used. Thus, there are numerous levels of storage into which material may be entered, and a decision must be made as to which one is appropriate. Like the file management problems already addressed, this is a routine decision that should not require the repeated attention of management.

After applying the decision model used to determine the relative utility of elements in the file, a second decision model can be used to determine the appropriate level of storage. This second model is similar to the first and can also be designed to reflect the priorities and preferences of management.

During the processing stage, the criteria for both models would be rated, and the criterion scores would be associated with the item. The utility measure, U, would be computed and compared to the minimum acceptable utility (i.e., $U \geq U_{min}$?). If the comparison is favorable, the overall score, f, for the second model would be computed. This score, which would be called the file location score, would be compared with a minimum acceptable score for entry into the highest-level (i.e., fastest retrieval) storage device (i.e., $f \geq f_n$?). This comparison would be repeated for decreasing (i.e., slower retrieval) levels of storage until it equals or exceeds the cut-off score for some level. This is the highest level of storage in which the information would be kept. If that level of storage (e.g., magnetic core) is full, then either the item is stored in some lower level (e.g., tape) or an existing item is "bumped" to a lower level. The decision would be based on the utility. Then, the procedure would continue in a manner identical to that explained in Chapter III. This process is summarized below, and a complete representation for an n-level storage system is shown in Figure 14.

File Not Full

The item is collated for future use. It can now be removed for three reasons: (1) use or dissemination, (2) periodic file inspections, and (3) to determine whether it has passed its expiration date. In the first two cases, the item will be reviewed and considered for reclassification. If reclassification is not necessary, it will be recollated; otherwise, it will be reevaluated by both models, and the cycle will begin again. In the third case, its expiration date will be checked. If the item is still current, it will be recollated; otherwise, it will be identified as a candidate for purging. Items that are so identified will be reviewed as though they were a new item.

File Full

The utility of the lowest valued item in the file is compared with the utility of the new item. If the new item has a higher utility than one in the file, it is collated for future use, and the low utility item is treated as though it were a new item.

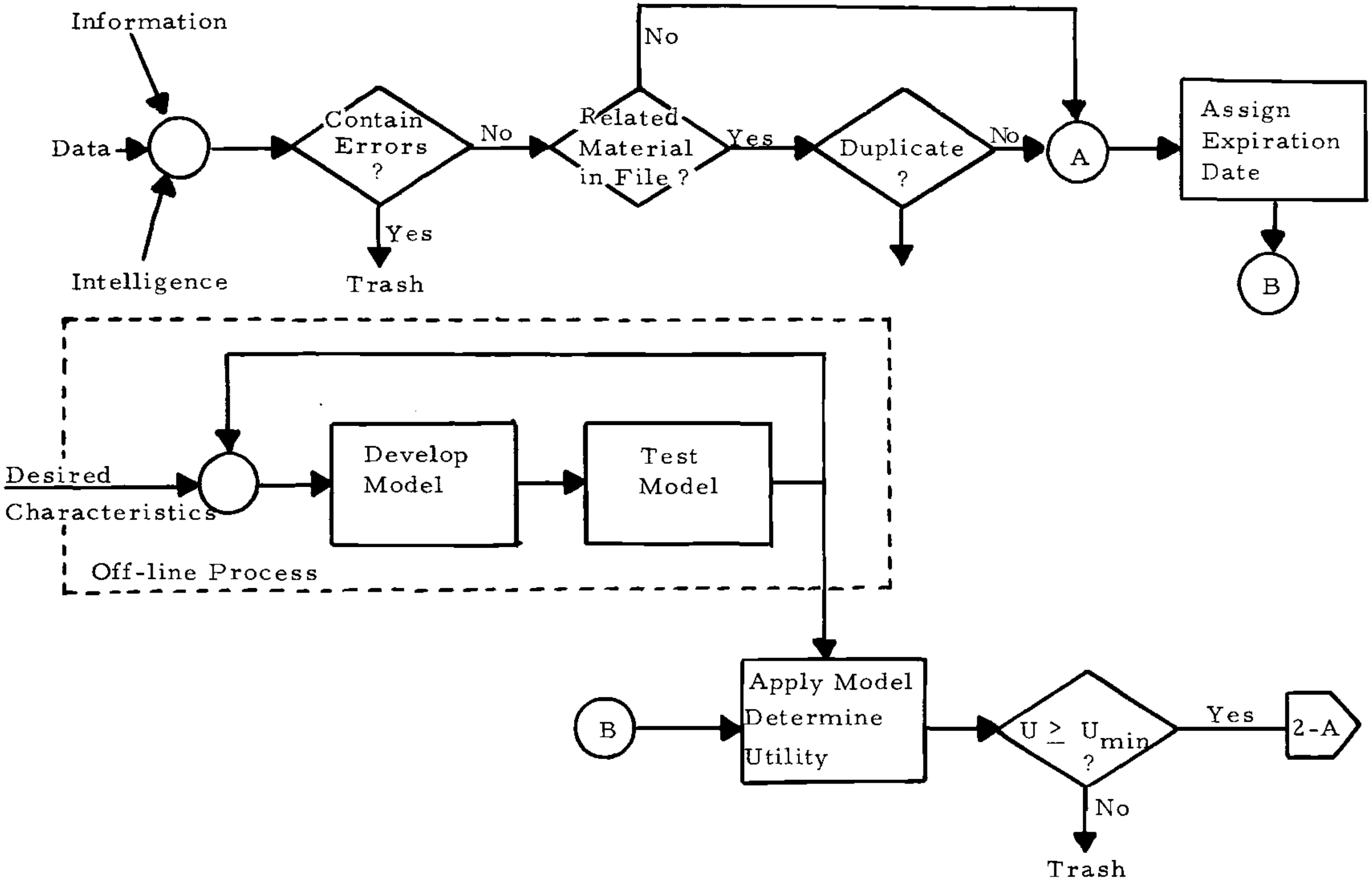


Figure 14. The Proposed Intelligence System with n Levels of Storage, Part 1.

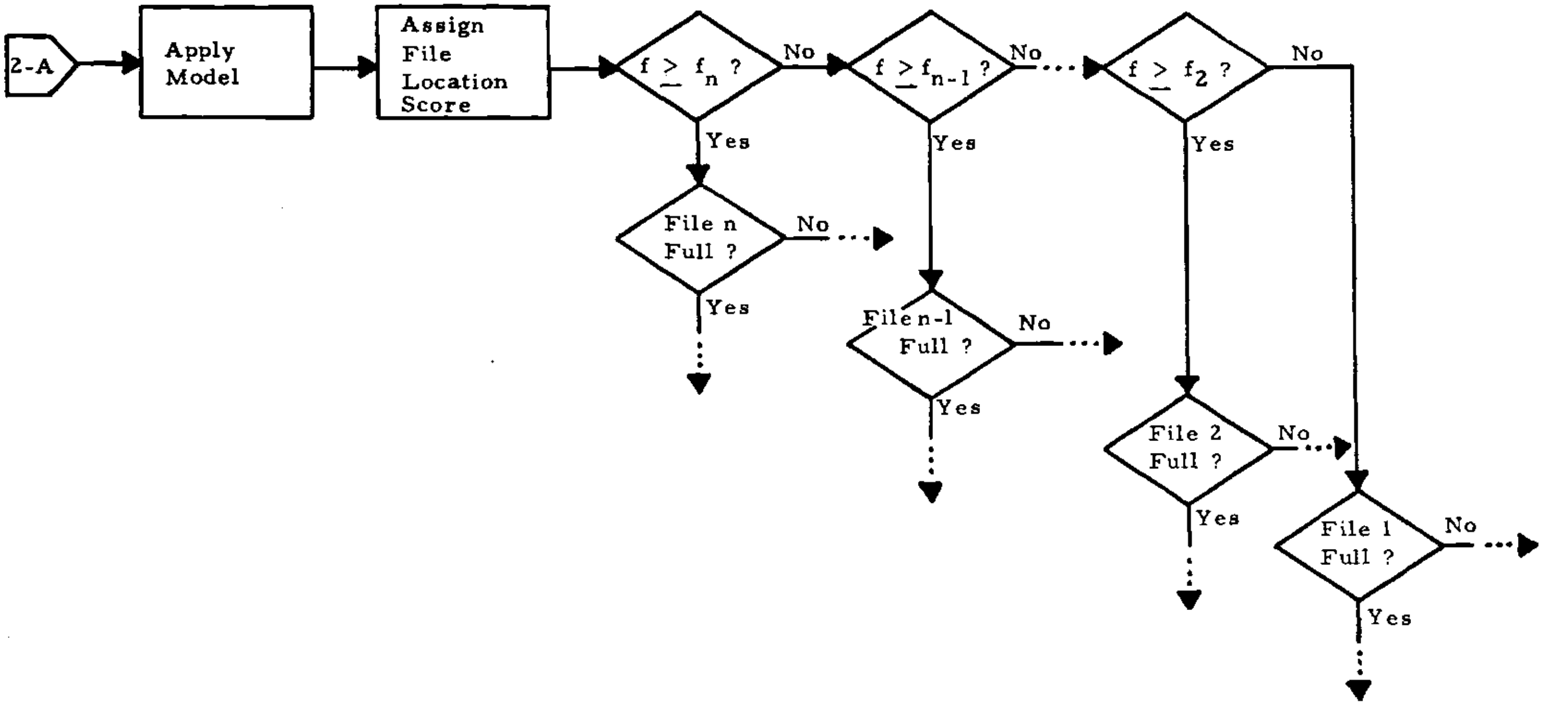


Figure 14 (continued). The Proposed Intelligence System with n Levels of Storage, Part 2.

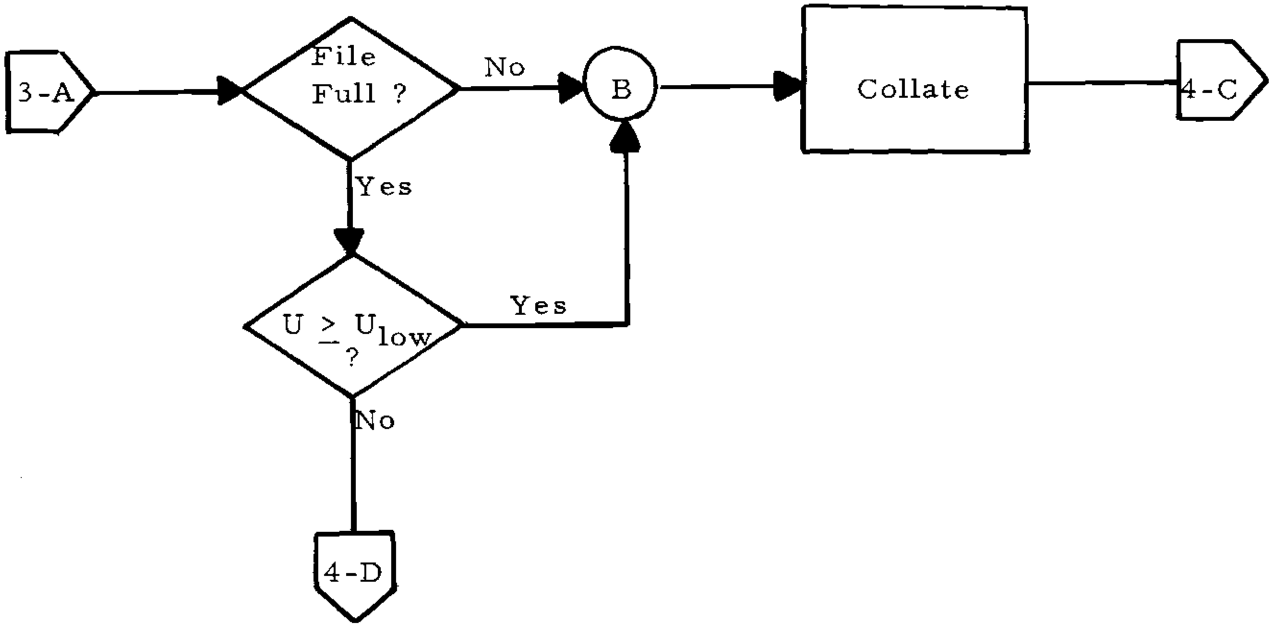


Figure 14 (continued). The Proposed Intelligence System with n Levels of Storage, Part 3.

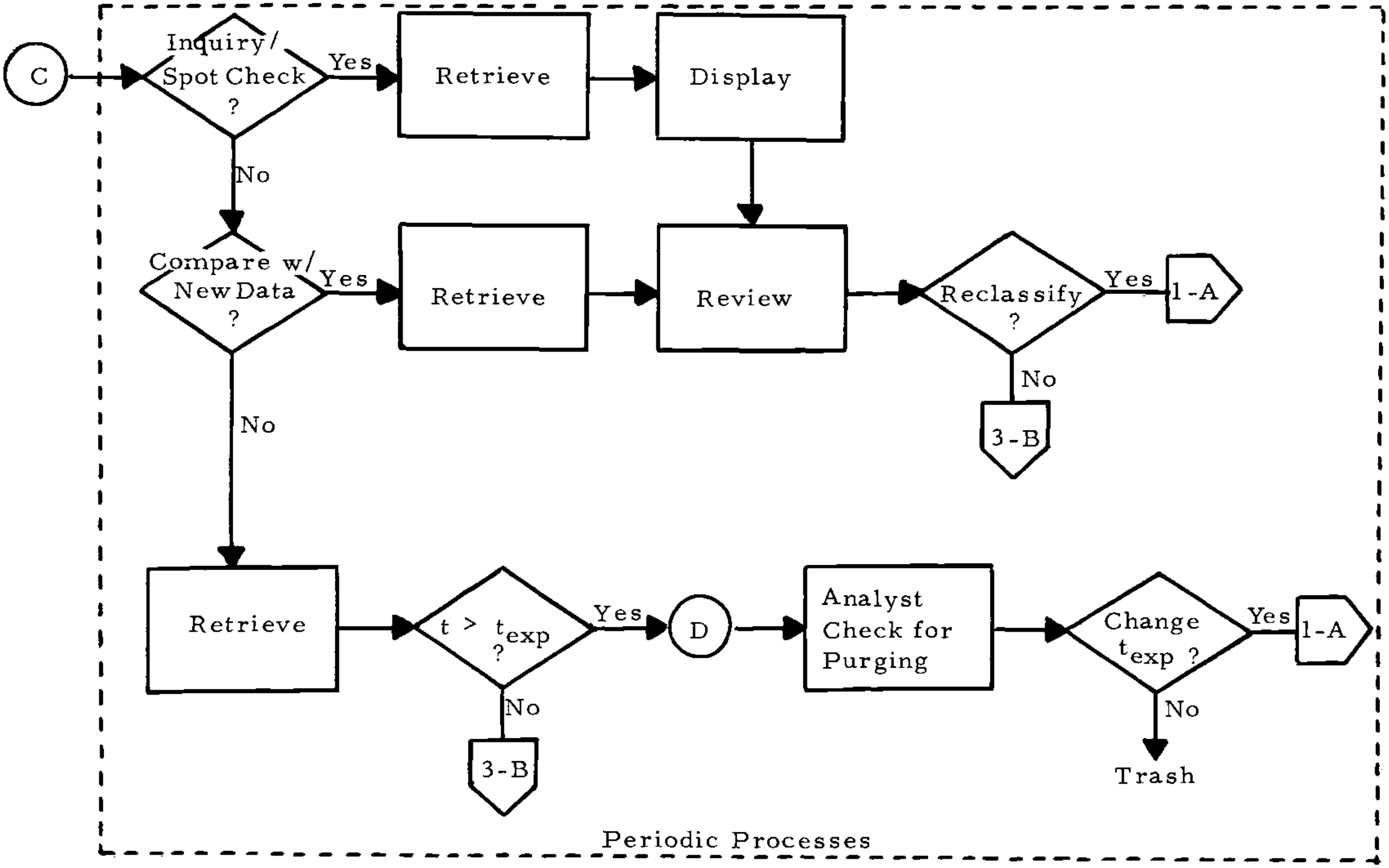


Figure 14 (continued). The Proposed Intelligence System with n Levels of Storage, Part 4.

A Computer-Based System Configuration

No attempt is made here to compare alternative configurations or competing brands of computer hardware. Rather, one possible configuration is presented. The only assertion is that the visual-display terminal is preferable to the typewriter terminal as the interface between the manager and the system. Morton's (28) work with management decision systems supports this assertion as does research dealing with the Army's tactical operations system (33).

A conceptual representation of a computer-based intelligence system is shown in Figure 15, and a summary of the actions occurring at each location is given below.

Intelligence Analyst

The intelligence analyst has the following duties:

1. Conduct preliminary screening for duplication and error.
2. Determine the expiration date of the item.
3. Evaluate material with respect to each criterion and assign criterion scores.
4. Request needed material from the files.
5. Review items that have been retrieved for use or identified as candidates for purging.
6. Determine whether or not previously filed material should be reevaluated.

Other Agencies

Other agencies may:

1. Request information from the intelligence files.
2. Generate data for input to the intelligence system.

Archives

The archives are a repository for low-valued or inactive information. Items that have been identified for purging and not reclassified will be forwarded to the archives.

Remote Access

The visual-display terminals and typewriter terminals will be in close proximity to the analysts. This display is the interface between intelligence personnel and the automated system. It is the primary method for input and output to the central computer.

Central Computer

The central computer performs the following functions:

1. Compute item utilities, U, and the file location scores, f.
2. Compare
   a. Item utility to minimum required utility; i.e., $U \geq U_{min}$?.
   b. File score to file cut-off scores; i.e., $f \geq f_i$?.
   c. Date to expiration date; i.e., $t > t_{exp}$?.
3. Sort the file and/or index entries based on their utilities.

Photo Center

The photo center produces microfiche files from the original

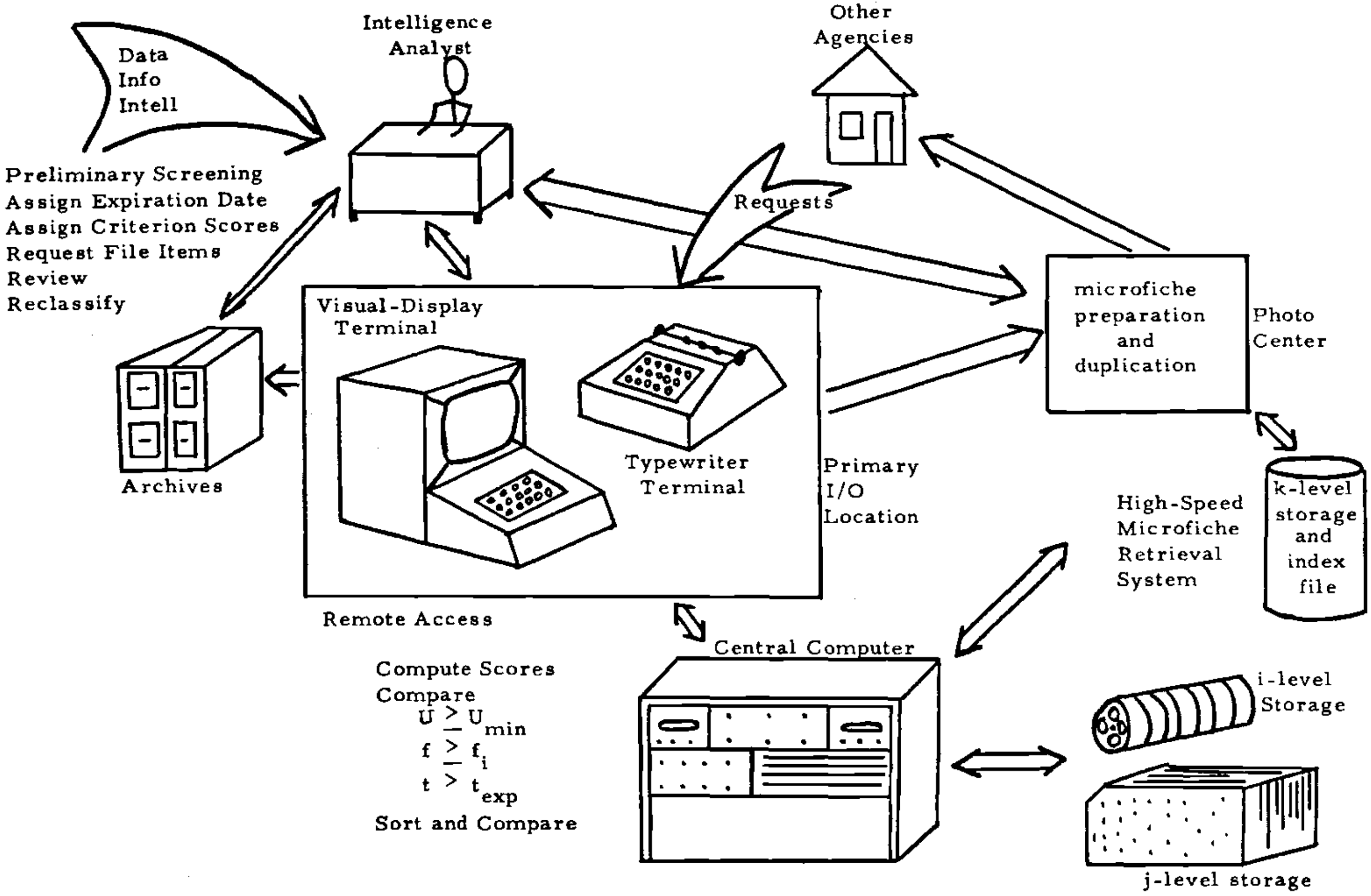


Figure 15. A Computer-Based Intelligence System.

documents, photographs, etc., and prepares duplicate microfiche for dissemination.

High-Speed Microfiche Retrieval

This storage device is used for the index to the intelligence files and the material entered in the kth level of storage.

Summary

The computer-based system described here is attainable with equipment that is available on the market today, and the software required to implement the proposed methods is routine. The necessary equipment is already in use in several local, state, and national police organizations and could easily be modified to implement the proposed intelligence system.

CHAPTER VI

DEVELOPING AN EXPERIMENTAL MODEL

This chapter presents an application of the model design procedure presented in Chapter IV. To develop the experimental model, it was necessary to select an organization that included an established intelligence system capable of providing data for at least limited testing of the model. The Georgia Bureau of Investigation (GBI)* satisfied this requirement.

Criteria Selection

Extensive interviews were conducted with personnel in the investigative division of the GBI. Initial interviews were devoted to a full explanation of the purpose of the model. Subsequent interviews dealt with the determination of criteria that determine the intelligence value of a particular piece of information. Emphasis was initially given to developing a complete list of criteria.

As the personnel involved in the interviews gained a better understanding of the model, the relevance and measurability of each criterion were discussed, and overlapping criteria were combined.

---

*Renamed "Criminal Investigation Division" in a recent reorganization of the Georgia state government.

The final list included the five criteria listed below.

1. Reliability of the source of the information.

2. Credibility of the information.

3. Value of the information in relation to other available information and in conjunction with its value as intelligence when considered by itself.

4. Relative importance of factors, other than an existing file, to which the information is related (e.g., a serious crime or a particular issue which reflects the priorities of the agency).

5. Timeliness of the information.

Criterion Weights

Forms were prepared to gather data for the initial determination of the criterion weights. Two forms were used--one for ranking the criteria and one for rating them. Thus, two sets of weights could be determined. (The forms are shown in the Appendix.)

The weights obtained by each of these methods are shown in Table 1. It will be noted that they are order consistent.

Initial Specification of the Model

Background gained during the interviews dealing with selection of criteria was used to construct a preliminary set of scales. During additional interviews, these preliminary scales were modified, and definitions were added to give increased clarity. Instructions for use of the scales were then prepared and added to the model. With the exception of the scale used to evaluate timeliness, the initial scales are the same as those presented in the Appendix. The initial scale for evaluating timeliness is shown in Figure 16. It will be noted that initially, "event in progress" received a score of 9 whereas "event soon to occur," "event to occur in the future," "event which has recently occurred," and "past event" received scores of 8, 7, 6, and 5, respectively.

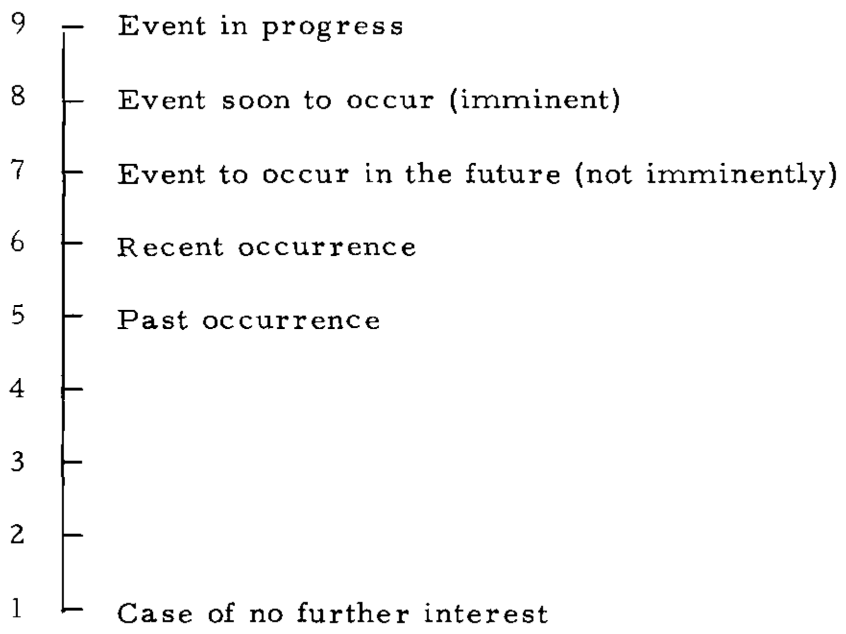


Figure 16. Initial Scale for Evaluating Timeliness.

Benchmark Determination

The benchmark chosen for comparison with the model rankings was the subjective ranking of the intelligence reports by the intelligence analysts of the GBI.

Initial Verification

GBI personnel preferred to test the model on new data rather than retrieving old data from the files. Because input to the intelligence system was slight during the testing period, initial verification was attempted with only one set of data. This set contained only nine reports instead of the recommended 15.

The nine reports were ranked subjectively by the analysts and were then evaluated using the initial model. Model rankings were obtained using both the weights derived by ranking and those derived by rating. The consistency between the subjective rankings and the rankings obtained from the model scores (using weights obtained by ranking) was measured using the Spearman Rank Correlation Coefficient (r_s). The results are shown in Table 2.

Tuning the Model

To attempt to achieve a higher correlation between the subjective and model rankings, the model weights were varied. The changes in correlation between these rankings for Analyst 1 are shown in Table 3.

A further study of the data indicated that all those reports having model ranks differing significantly from the subjective ranks had scores of 9 on the criterion of timeliness. This suggested that the timeliness scale did not accurately reflect the priorities of GBI personnel.

Table 1. Criterion Weights Obtained by Ranking and Rating

No.	Method Criterion	Ranking	Rating
1	Source Reliability	.323	.259
2	Information Credibility	.290	.259
3	Information Value	.194	.220
4	Related Factors	.065	.078
5	Timeliness	.129	.185

Table 2. Results Using the Initial Model

Analyst	Wts by Rank	Wts by Rate
	r_s	r_s
1	.804	.754
2	.854	.854
Average	.829	.804

r_s = Spearman Rank Correlation Coefficient.

Table 3. Effect on Correlation of Varying Criterion Weights

Change	Criterion Weights					r_s
Initial	.323	.290	.194	.065	.129	.804
Change 1	.28	.28	.22	.05	.17	.804
Change 2	.25	.25	.25	.05	.20	.804
Change 3	.20	.20	.20	.20	.20	.717
Change 4	.30	.30	.20	.05	.15	.804
Change 5	.30	.30	.20	.10	.10	.833
Change 6	.30	.20	.20	.10	.20	.730
Criterion No.	1	2	3	4	5	

Therefore, the scale for this criterion was changed to that shown in the Appendix. Adjusting the criterion scores to reflect this change in scale and reapplying the model with the initial weights (obtained by ranking) resulted in correlation coefficients of .979 and .854 for Analysts 1 and 2 respectively. Since this averaged to .916, new data was obtained and ranked both subjectively and by the altered model.

Data supplied the second time consisted of both subjective and model rankings of three sets (identified as A, B, and C) of five reports and a subjective ranking of a set of ten. This set of ten (labeled Combined Set) was made up of two of the previous three sets of five reports. The results are summarized in Table 4.

Since it had been previously decided that a correlation coefficient of 0.9 indicated an acceptable level of consistency between the subjective and model rankings, no additional changes were made in the model.

Sensitivity Analysis

An analysis was conducted to determine the sensitivity of the model to changes in the criterion weights. Data used for this test included both the subjective ranking and the criterion scores given by Analyst 2 to the ten reports comprising the combined set. The results are summarized in Table 5. This limited testing showed that the model is not very sensitive to moderate changes in the criterion weights.

The reasons for the behavior exhibited by the model when sub-

Table 4. Correlations Using the Altered Model

Analyst	Data Set A	Data Set B	Data Set C	Average	Combined Set
	r_s	r_s	r_s	r_s	r_s
1	.90	.90	.80	.867	.903
2	1.0	1.0	.90	.967	.900
Average	.95	.95	.85	.917	.902

r_s = Spearman Rank Correlation Coefficient.

Table 5. Sensitivity Analysis on the Final Data

Change	Criterion Weights					r_s
Initial	.323	.290	.194	.065	.129	.900
1	.35	.35	.10	.10	.10	.864
2	.20	.20	.20	.20	.20	.909
3	.30	.30	.20	.10	.10	.900
4	.30	.30	.05	.05	.30	.764
5	.34	.33	.33	0	0	.873
6	.34	.33	0	0	.33	.623
7	0	0	.50	.50	0	.905
Criterion #	1	2	3	4	5	

jected to gross changes in weights, such as those in Changes 5, 6, and 7 in Table 5, cannot be stated based on this limited testing. It may result from the data used in the sample or the covariance between the criteria may be high. Further testing with increased sample sizes and multiple sets of data to permit averaging results is necessary to verify the cause.

CHAPTER VII

CONCLUSIONS

The objective of this research as stated in Chapter I was accomplished. No claim is made that the methods presented here are optimal. However, they do represent a logical next step in synthesizing intelligence systems.

It is believed that the methods discussed can be integrated into existing intelligence systems and are adaptable to either manual or automated systems. The system presented in Chapter III is complete, and it is attainable without requiring any breakthroughs in the state of the art of management information systems, decision techniques, or computer technology. The computer-based system presented in Chapter VI is attainable with available hardware and routine software.

The major conclusions derived from this research are:

1. A system has been designed that may alleviate many of the file management problems associated with intelligence systems.

2. A weighted, additive scoring model has been designed to reflect the priorities of the manager of the intelligence system.

3. An experimental model has been tested using actual data and personnel of the Georgia Bureau of Investigation. The tests show the

model to be an appropriate solution and easily implementable.

4. The experimental model developed in consultation with the Georgia Bureau of Investigation is relatively insensitive to moderate variations in criterion weights. (See Table 5, p. 70.) This indicates that there is no requirement for frequent reassessment of the criterion weights and reduces the need for retuning the model. It further indicates that the use of a more sophisticated model will not necessarily produce significantly better results.

5. The scoring model described in the Appendix can be used to:

a. Screen irrelevant and low-valued inputs to prevent their entry into the system.

b. Integrate new material into the file so that the relative utility of most items is known.

c. Facilitate file update by forcing low-valued items to be either reevaluated or purged from the system.

d. Routinize file management procedures.

6. The system proposed in Chapters III and V provides for routine purging or reevaluation by associating an expiration date with each input to the system.

APPENDIX

EXPERIMENTAL MODEL

APPENDIX

EXPERIMENTAL MODEL

This appendix consists of the two forms used to determine criterion weights and an experimental linear scoring model which was developed in consultation with members of the investigative division of the GBI.

Enclosure 1 to this appendix is the form used to gather data for determining criterion weights based on a rank ordering of the criteria.

Enclosure 2 to this appendix is the form used to gather data for determining criterion weights based on a rating of the relative importance of each criterion.

Enclosure 3 to the appendix is the set of instructions and scales used to evaluate information in relation to each of the five criteria.

Enclosure 1

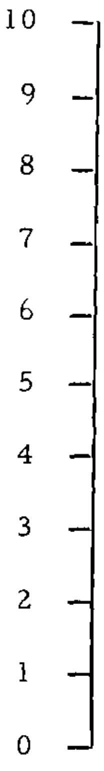
Rank each of the following criteria according to the relative importance it plays in determining the overall value of a piece of information. A rank of 1 indicates the most important criterion, and 5 indicates the least important.

Rank	Criterion
___	Timeliness of the information.
___	Value of the information in relation to other available information and in conjunction with its value as intelligence when considered by itself.
___	Reliability of the source of the information.
___	Relative importance of factors, other than an existing file, to which the information is related (e.g., a serious crime or a particular issue which reflects the priorities of the agency).
___	Credibility of the information.

Enclosure 2

Draw a line from each of the criteria listed on the right to a point on the scale on the left which reflects its relative importance to you in evaluating the importance of a piece of information. A value of 10 should be assigned to the most important criterion and smaller numbers to the other criteria depending on their relative importance.

It is permissible to relate the same point on the scale to more than one criterion, and it is not necessary to restrict the scale points to integer values (e.g., 6.3 is an acceptable rating).



Criterion

a. Timeliness of the information.

b. Value of the information in relation to other available information and in conjunction with its value as intelligence when considered by itself.

c. Reliability of the source of the information.

d. Relative importance of factors, other than an existing file, to which the information is related (e.g., a serious crime or a particular issue which reflects the priorities of the agency).

e. Credibility of the information.

Enclosure 3

INSTRUCTIONS FOR USE OF THE MODEL

There are five criteria which determine the value of information to the intelligence system; these are listed below.

1. Reliability of the source of the information.
2. Credibility of the information.
3. Value of the information in relation to other available information and in conjunction with its value as intelligence when considered by itself.
4. Relative importance of factors, other than an existing file, to which the information is related (e.g., a serious crime or a particular issue which reflects the priorities of the agency).
5. Timeliness of the information.

On the following pages, a spectrum of values ranging from 1 (lowest) to 9 (highest) is displayed for each criterion. Next to the scale are descriptions which illustrate the intended meaning of certain numerical scores for a criterion.

The scale is continuous; you are not limited to choosing a number which has a definition, nor are you limited to choosing an integer value for a criterion. For example, 9, 8, and 7.5 are all acceptable ratings for reliability of the source.

Each piece of information is to be evaluated in regard to each of the five criteria. When evaluating a particular criterion, it should be examined independently of the other criteria; e.g., credibility of the information should be evaluated without regard to the reliability of the source and independently of its timeliness.

RELIABILITY OF THE SOURCE

9	Highly reliable
8	
7	Usually reliable
6	Sometimes reliable
5	
4	Unknown
3	
2	Not often reliable
1	Unreliable

This factor is intended to measure the reliability of the source of the information; the description of each of the points defined on the scale provides the intended meaning.

Highly reliable - source has furnished information on numerous occasions, and investigation has verified that the information is accurate on almost every occasion.

Usually reliable - source has furnished information on numerous occasions, and investigation has verified that approximately 75% of the information is accurate.

Sometimes reliable - source has provided information on a number of occasions; the information is as likely to be erroneous as it is to be accurate.

RELIABILITY OF THE SOURCE cont'd

Unknown - source has provided little or no information before, and his reliability cannot be judged.

Not often reliable - source has furnished information before, and it is more likely to be erroneous than it is to be accurate.

Unreliable - source has furnished information on a number of occasions, and it is almost always erroneous.

CREDIBILITY OF THE INFORMATION

9	Factual
8	Observations
7	Opinion based on observations
6	Hearsay (primary source can be checked)
5	
4	Unsubstantiated opinion
3	Hearsay (cannot be checked further)
2	Unknown
1	Unlikely

This factor is intended to measure only the credibility of the information; in assessing this factor, no consideration should be given to the reliability of the source of the information. Descriptions of the points which are defined on the scale provide the intended meaning.

Factual - the information is a matter of public record, common knowledge, or documented by indisputable evidence known to the evaluator.

Observations - the information is a report only of observations made by the source--no attempt has been made to interpret them.

Opinion based on observations - the information is an interpretation of observations made by the source.

Hearsay (primary source can be checked) - the primary source

CREDIBILITY OF THE INFORMATION cont'd

of the information is someone other than the individual supplying the information, but the primary source can be checked.

Unsubstantiated opinion - the information is strictly the opinion of the source.

Hearsay (cannot be checked further) - the primary source of the information is someone other than the individual supplying the information, and the primary source cannot be checked.

Unknown - the credibility of the information cannot be judged.

Unlikely - based on evidence available to the evaluator, the information is probably not correct.

RELATED INFORMATION AVAILABLE AND RELATIVE IMPORTANCE OF INFORMATION

9 Additional information to an existing file/key information
8
7 No related file/key information
6
5 Additional information to an existing file/minor information
4
3 No related file/minor information
2
1 Duplication of available information/unimportant information

This factor is intended to measure the value of information based on two things--its value as a part of a larger body of information, and its value as intelligence by itself.

Additional information to an existing file/key information - the information is related to information in an existing file, and it constitutes a significant piece of intelligence by itself.

No related file/key information - there is no existing file to which the information relates, but it is an important piece of intelligence on its own merit.

Additional information to an existing file/minor information - the information is related to information in an existing file but is not

RELATED INFORMATION AVAILABLE AND RELATIVE IMPORTANCE OF INFORMATION cont'd

significant intelligence on its own merit; however, the information is worth keeping.

No related file/minor information - there is no existing file to which the information relates, and it is not important information by itself, but it is worth keeping.

Duplication of available information/unimportant information - the information is already available in an existing file or it has no importance as intelligence.

RELATIVE IMPORTANCE OF RELATED FACTORS

9 Serious crime related to an important issue
8 Serious crime related to organized crime
7 Other serious crime
6 Minor crime related to an important issue
5 Minor crime related to organized crime
4 Other minor crime
3 Information related to an important issue
2 Information related to organized crime
1 Other information

This factor is intended to evaluate the importance of factors related to the information; viz., the nature of the crime and the affiliation with organized crime or other key issues.

Serious crime - for example, murder, rape, or arson.

Minor crime - any crime not classified as a serious crime.

Information - not directly related to crime; e.g., reports on the day-to-day activities of known felons.

Important issue - an issue which reflects the day-to-day priorities of the agency.

Organized crime - for example, the Cosa Nostra or a gambling syndicate.

TIMELINESS

9	Event soon to occur (imminent)
8	Event to occur in the future (not imminently)
7	Recent occurrence
6	Past occurrence
5	Event in progress
4	
3	
2	
1	Case of no further interest

This factor measures the currentness of the information. The points defined on the scale provide the intended meaning.

Event soon to occur (imminent) - the activity described may occur within *

Event to occur in the future (not imminently) - the activity described may occur in the future but not for at least *

Recent occurrence - the event has occurred within the recent past.*

Past occurrence - the event has occurred in excess of * ago.

Event in progress - the activity with which the information deals is taking place at the time of evaluation.

*NOTE: These time periods will vary depending on the type of information that is being evaluated.

TIMELINESS cont'd

Case of no further interest - for example, the statute of limitations has expired in the case to which the information relates, or the case has been solved.

LITERATURE CITED

1. Ackoff, Russell L., "Management Misinformation Systems," Management Science, vol. 14, no. 4 (Dec 1967), pp. 147-156.

2. Anonymous, Intelligence Handbook, United States Army Infantry School, Ft. Benning, Ga., 1970.

3. __________, Organized Crime Law Enforcement Conference, Law Enforcement Assistance Administration, Williamsburg, Va., 1970.

4. Bristow, A. P., Decision-Making in Police Administration, Charles C. Thomas, Springfield, Illinois, 1961.

5. Conroy, George E., "A Field Interview System Using EAM Equipment," Law Enforcement Science and Technology II, ed. S. I. Cohn, Port City Press, Inc., 1968.

6. Dean, B. V., and Nishry, M. J., "Scoring and Profitability Models for Evaluating and Selecting Engineering Projects," Operations Research, vol. 13, no. 4 (July-Aug 1965), pp. 550-569.

7. Dean, J., Capital Budgeting, Columbia University Press, New York, 1951.

8. Eckenrode, R. T., "Weighting Multiple Criteria," Management Science, vol. 12, no. 3 (Nov 1965), pp. 180-192.

9. Eilon, Samuel, "Some Notes on Information Processing," The Journal of Management Studies, vol. 5 (May 1968), pp. 139-153.

10. Einhorn, H. J., "The Use of Nonlinear, Noncompensatory Models in Decision Making," Psychological Bulletin, vol. 73, no. 3 (1970), pp. 221-230.

11. __________, "Use of Nonlinear, Noncompensatory Models as a Function of Task and Amount of Information," Organizational Behavior and Human Performance, vol. 6, no. 1 (1971), pp. 1-27.

12. ______________, "Expert Measurement and Mechanical Combination," *Organizational Behavior and Human Performance*, vol. 7, no. 1 (1972), pp. 86-106.

13. Engstrom, E. W., "System Engineering--A Growing Concept," *Electrical Engineering*, vol. 76 (1957), pp. 113-116.

14. Goldberg, L. R., "Simple Models or Simple Processes," *American Psychologist*, vol. 23 (1968), pp. 483-496.

15. Goode, H. H., and Machol, R. E., *System Engineering--An Introduction to the Design of Large-Scale Systems*, McGraw-Hill Book Co., Inc., New York, 1957.

16. Goodwin, P. G., "A Method for Evaluation of Subsystem Alternate Designs," *IEEE Transactions on Engineering Management*, vol. EM-19, no. 1 (Feb 1972), pp. 12-21.

17. Hall, Arthur D., *A Methodology for Systems Engineering*, Van Nostrand Reinhold Co., New York, 1962.

18. Hammond, K. R., and Summers, D. A., "Cognitive Dependence on Linear and Nonlinear Cues," *Psychological Review*, vol. 72 (1965), pp. 215-224.

19. Heany, Donald F., *Development of Management Information Systems: What Management Needs to Know*, The Ronald Press Co., New York, 1968.

20. Hertz, D. B., "Risk Analysis in Capital Investment," *Harvard Business Review*, vol. 42, no. 1 (Jan-Feb 1964), pp. 95-106.

21. Howard, R. A., "Information Value Theory," *IEEE Transactions on Systems Science and Cybernetics*, vol. SSC-2, no. 1 (Aug 1966), pp. 22-34.

22. ______________, "Value of Information Lotteries," *IEEE Transactions on Systems Science and Cybernetics*, vol. SSC-3, no. 1 (June 1967), pp. 54-60.

23. ______________, "Bayesian Decision Models for Systems Engineering," *IEEE Transactions on Systems Science and Cybernetics*, vol. SSC-1, no. 1 (Nov 1965), pp. 36-40.

24. Kelly, Joseph F., Computerized Management Information Systems, The MacMillan Company, London, 1970.

25. Kendall, M. G., Rank Correlation Methods, Charles Griffin & Co. Ltd., London, 1955.

26. Moore, J. R., and Baker, N. R., "Computational Analysis of Scoring Models for R&D Project Selection," Management Science, vol. 16, No. 4 (Dec 1969), pp. B-212-B232.

27. ____________, "An Analytical Approach to Scoring Model Design--Application to Research and Development Project Selection," IEEE Transactions on Engineering Management, vol. EM-16, no. 3 (Aug 1969), pp. 90-98.

28. Morton, Michael S. Scott, Management Decision Systems: Computer-Based Support for Decision Making, Graduate School of Business Administration, Harvard University, Boston, 1971.

29. Mottley, C. M., and Newton, R. D., "The Selection of Projects for Industrial Research," Operations Research, vol. 7 (Nov-Dec 1959), pp. 740-751.

30. Murrish, J. B., Trends in Management Information Systems and Their Impact on Management in the 1970's, Stanford Research Institute, Menlo Park, California, 1969.

31. Prince, Thomas R., Information Systems for Management Planning and Control, Richard D. Irwin, Inc., Homewood, Illinois, 1970.

32. Raiffa, H., Decision Analysis: Introductory Lectures on Choices under Uncertainty, Addison-Wesley, Reading, Massachusetts, 1970.

33. Ringel, Seymour, et al., Human Factors Research in Command Information Processing Systems--Summary of Recent Studies (Research Study 69-6), Behavioral Science Research Laboratory, 1969.

34. Sawyer, J., "Measurement and Prediction: Clinical and Statistical," Psychological Bulletin, vol. 66 (1966), pp. 178-200.

35. Schultz, Donald O., and Norton, Loran A., Police Operational Intelligence, Charles C. Thomas, Springfield, Ill., 1968.

36. Solomon, E., "The Arithmetic of Capital Budgeting Decisions," Journal of Business, vol. 29 (1956), p. 240.

37. Whisenand, Paul M., and Tamaru, Tug T., Automated Police Information Systems, John Wiley and Sons, Inc., New York, 1970.

38. Wiggins, N., and Hoffman, P. J., "Three Models of Clinical Judgment," Journal of Abnormal Psychology, vol. 73 (1968), pp. 70-77.

39. Williams, William F., Principles of Automated Information Retrieval, The Business Press, Elmhurst, Illinois, 1968.

40. Yntema, D. B., and Torgerson, W. S., "Man-Computer Cooperation in Decisions Requiring Common Sense," IRE Transactions of the Professional Group on Human Factors in Electronics, vol. HFE-2 (1961), pp. 20-26.

41. Symonds, Curtis W., A Design for Business Intelligence, American Management Association, Inc., 1971.