Modeling the U.S. Military Intelligence Process

Dr. J.O. Miller
Department of Operational Sciences
Graduate School of Engineering and Management
Air Force Institute of Technology
2950 Hobson Way
Wright-Patterson AFB, OH 45433
Voice: 937-255-6565 x4326; Fax: 937-656-4948
John.Miller@afit.edu

Capt Carl R. Pawling, USAF
ACC/DRYF
Langley AFB, VA 23665
Voice: 757-764-2069; Fax: 757-764-7217
Carl.Pawling@langley.af.mil

Maj Stephen P. Chambal, Ph.D.
Studies & Analyses Group
NRO DDSE, 4GB18C
14675 Lee Road
Chantilly, VA 20151-1715
Voice: (703)808-3211; Fax: (703)808-1310
Stephen.Chambal@nro.mil

Abstract

Military use of the intelligence process is vital both in and of itself and as a valuable input to our Command and Control, enhancing military power through information superiority. The intelligence process begins when a need for information or intelligence is identified and encompasses how these information needs are met. As such, this process includes all of the satellites, aircraft, and communications systems used to gather and transmit data as well as the people, organizations, and resources involved in turning raw data into useful information. To support current and future Intelligence, Surveillance, and Reconnaissance (ISR) systems more detailed analysis of the intelligence process becomes critical to determine what pieces may need to be improved or expanded. Our approach involves construction of a modularized top level computer simulation model of a generalized military intelligence process using the *Arena* process oriented simulation software. The model provides the ability to perform quick turn analysis for comparing structural modifications to the intelligence process using typical measures of performance (quality, quantity, timeliness, and information needs satisfaction). The study also includes a statistical analysis of various configurations of this intelligence process.

Introduction

The importance of assessing the information flow of the military intelligence process has been brought to light in recent years. Although, the need for this assessment is not new, it becomes more critical as both capability and demand increases. According to the National Security Strategy, "We must transform our intelligence capabilities and build new ones to keep pace with" terrorist and other threats [White House, 2002:30]. The intelligence process begins when a need for information or intelligence is identified and encompasses how these information needs are met. As such, it includes all of the satellites, aircraft, communications, and other systems used to gather and transmit data as well as the people, organizations, and resources involved in turning raw data into useful information. The intelligence process can take on many forms, two of which are Task, Process, Exploit, Disseminate (TPED) and Task, Process, Post, Use (TPPU).

To illustrate the TPED and TPPU processes, take for example a person with a standard 35mm camera, lets call him Bob. The need for information about a particular place has arisen. Once this need is realized and it is determined that a picture can satisfy the need, Bob is directed or *tasked* to go take a picture of this place. Once Bob has taken the picture and returned, the picture is still of no use. A series of steps must be taken to put the picture into a usable form. Specifically the picture must be developed or *processed*. Now that the image is in a usable form, the remainder of the intelligence process can be carried out in several ways. For the TPED approach, an analyst would take additional steps to *exploit* the picture. These steps might include marking important aspects of the image and adding notes to describe what those aspects are. Once this is complete, the picture with the markings and additional information would be sent out or *disseminated* to the person or organization that needed the picture. Alternatively, the TPPU approach would bypass the exploitation up front and send or *post* the picture to a web page. Then the person or organization that needed the picture could retrieve it and use it without the overhead of exploitation. Although this is a simplified example, it illustrates how TPED and

TPPU can differ. In reality, the intelligence process is more complex and is dependent on several systems, organizations, and user requirements.

In developing a simulation model of this intelligence process the researcher needs to ensure that the model is structured for specific measures of interest. Those measures should provide the required information for the intended uses of the model. Four measures of interest for analyzing the intelligence process are quality, quantity, timeliness, and information needs satisfaction (QQTI). The first three measures; quality, quantity, and timeliness (QQT), have historically been used when assessing various aspects of the intelligence process. As such previous models may be designed to provide those quantitative measures for a specific implementation of the intelligence process. Where applicable some of the concepts in these models may be used in model development. More importantly examination of open source and unclassified descriptions of the intelligence process provide the primary source for model development. In addition to basic descriptions, factors that influence the process and consequently the QQTI measures are taken into account based on subject matter expert (SME) discussions. The use of *Arena* to model the military intelligence process in a modularized design allows the fidelity of individual portions of the model to be easily expanded for future simulation studies.

Construction of a simulation model of the intelligence process requires a number of simplifying assumptions that dictate the level of abstraction of the model from the real world. Based on the intended uses of our model, a moderate level of detail should be sufficient for this initial study. Comparing the impact of changes to various portions of the intelligence process for quick look studies does not require the detail of engineering level models. Furthermore, the moderate level of detail extends to all portions of the intelligence process model including an embedded communications model. By maintaining the focus of building a top level framework for simulation studies, only open source and unclassified information are required to develop the structure of the model.

The remainder of this paper provides a brief discussion of the intelligence process and some previous work in modeling and assessing this process. We then provide more detail on the development of our *Arena* model of a generalized intelligence process. The paper continues with a discussion of our model input data and an analysis of some baseline results. We conclude with some insights and conclusions based on this study and recommendations for further efforts. A much more detailed presentation of this study can be found in Pawling [2004].

Background

The intelligence process used for this study is described in Joint Publication 2-0, *Doctrine for Intelligence Support to Joint Operations*, as the intelligence cycle. The intelligence cycle has six phases shown in Figure 1. These phases are Planning and Direction, Collection, Processing and Exploitation, Analysis and Production, Dissemination and Integration, and Mission Evaluation and Feedback. Planning and Direction involves planning for future contingencies in theaters and determining what resources might be required for those contingencies. Existing information and previously scheduled information gathering is used when possible, but the remaining information requirements are turned into requests for information (RFIs) which can lead to either production

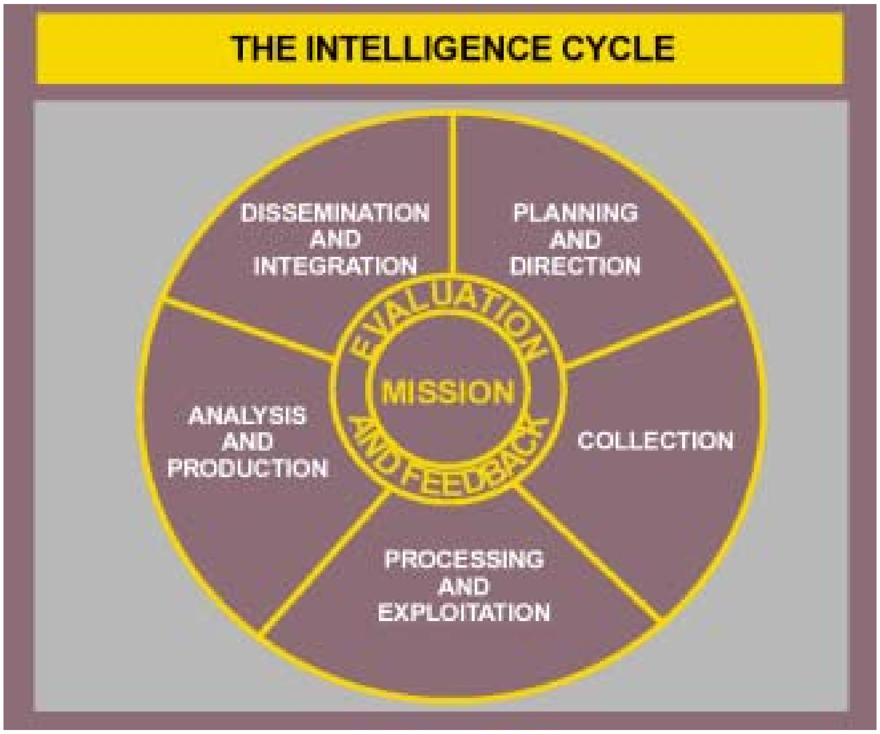

Figure 1. The Intelligence Cycle [JP 2-0, 2000:II-1]

or collection requirements [JP 2-0. 2000:II-4]. The second phase, Collection, carries out the plan for the gathering of information. In this phase, organizations or agencies that operate collection assets such as satellites or surveillance equipment task those assets to gather information at specified times and places. The means and methods of collection are highly dependent on the source of the information and these sources are generally categorized into various intelligence disciplines.

After Collection, the Processing and Exploitation phase takes the raw data gathered and transforms it so that it can be used for analysis and production [JP 2-0, 2000:II-7]. The amount and type of work involved depends on the type of intelligence that has been gathered. Analysis and Production uses processed and/or exploited information to generate intelligence products to meet the RFIs. The intelligence products are usually categorized by their primary use: indications and warning, current intelligence, general military intelligence, target intelligence, scientific and technical intelligence, and counterintelligence [JP 2-0, 2000:II-10-12]. These categories may overlap and are not confined to any particular source of intelligence. Dissemination and Integration involves sending intelligence products to the user and the user incorporating these products into their efforts [JP 2-0, 2000:II-12]. This phase of the intelligence cycle is heavily dependent on communications systems due to the likelihood of significant geographic separation between the intelligence production center and the user. The dissemination of information can be either “pushed” to the user to answer a request or “pulled” by the user from databases and other centralized sources of information.

The final phase, Mission Evaluation and Feedback, is integral to all of the other phases and is appropriately placed in the center of Figure 1. It is not conducted independently but must be accomplished throughout each phase to ensure that the process is working as expected. Qualitative attributes that are used to evaluate the quality of intelligence are timeliness, accuracy, usability, completeness, relevance, objectiveness, and availability [JP 2-0, 2000:II-14]. Of these, timeliness can be easily quantified by comparing the time of delivery with the time of need. The other attributes may be difficult to quantify, but they are nonetheless vital when evaluating the overall quality of intelligence. In addition to these individual measures, a single aggregated measure can be defined. One example is information needs satisfaction (INS). Although this measure can be viewed as the process of meeting information needs, as a measure it can describe either the proportion of needs that are met or the degree to which needs are met. When specifically related to QQT, INS would simply be the proportion of intelligence requests that meet both of the quality and timeliness requirements. Alternatively, INS could be rated on a scale with items that meet both of the quality and timeliness requirements at the top with decreasing scores based on how poor or how late the response to a request was received.

Before constructing our simulation model of the intelligence process, we conducted a review of previous simulation studies and existing models. The models reviewed ranged from high fidelity to low fidelity and implemented either a single architecture or multiple architectures. For a more detailed discussion on specific models reviewed, see Chapter 2 of Pawling [2004]. For quick look studies a low fidelity or high level model that describes multiple implementations is ideal. The Quick ISR Concept of Operations Modeler (QUICM) comes closest of the models examined to meeting the needs of our analysis [Kanewske, 2003]. However, QUICM is limited in only allowing two very structured architectural implementations of the intelligence process. The ability to model hybrid implementations allows for a better representation of reality and analysis of transitional stages between complete TPED or TPPU implementations. For example, such a transitional stage might be altering the process for one type of information or for a specific user and examining the impact of the overall system. Rather than modifying one of these existing models we decided to develop our own tailored model of the intelligence cycle.

Methodology

The Intelligence Process Model (IPM) consists of seven *Arena* submodels, six of which are taken directly from the intelligence cycle shown in Figure 1. The last submodel is named Communications and contains the logic to tie together the other submodels. The entity modeled in the IPM represents an RFI. These entities are related to real world RFIs, but they are not exactly the same. RFIs in the IPM are only gathered from a single information source. As such the simulated RFIs more closely resemble a tracking sheet that follows a real information request through the process. The rationale for choosing this abstraction is that assessment of the intelligence process at the top level does not require actual information, only the status of the requests. As the RFIs flow through the IPM a number of attributes keep track of each entity's specific request characteristics and requirements and the associated timeliness and information quality as these are updated throughout the model. These attributes provide the data used in the generation of statistics to evaluate the measures of performance. Figure 2 provides the top level view of the IPM. This modular approach allows any submodel to be easily modified or replaced

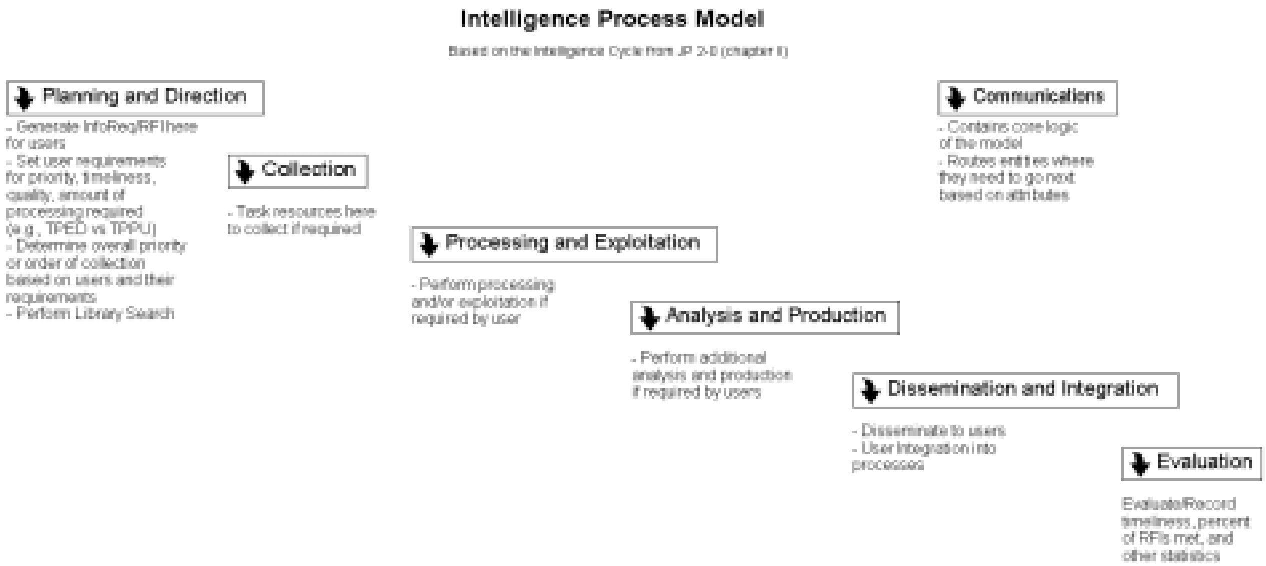

Figure 2. IPM Top Level View

and keeps the various portions of the process distinct. Each of the submodels is briefly described in the following paragraphs.

The *Planning and Direction* submodel indicates the beginning of the intelligence process and corresponds directly with the process of the same name from the intelligence cycle. The purpose of this submodel is to generate user information requirements and prioritize them globally. As an additional part of the planning process, users perform a library search to determine if existing information may meet their needs. This submodel consists of five separate user requirement submodels representing different generic intelligence users. Each of these submodels generates standing requirements (RFIs) and additional requirements (RFIs) according to input parameters selected. Each RFI is also assigned attributes that determine which of the remaining submodels it is required to pass through as it leaves. This allows the IPM to capture a TPED like architecture (all steps of the intelligence cycle carried out) or different implementations of a TPPU hybrid with some selected steps left out.

The *Collection* submodel corresponds directly to the collection phase of the intelligence cycle. This submodel simulates collection from various sources and determines the quality of the information collected. When arriving to this submodel RFIs are placed into a ranked queue based on their priority attribute until an appropriate collection resource becomes available for use. Upon leaving the queue each RFI is immediately checked for timeliness. If the RFI fails this check it is immediately discarded from the simulation and the next RFI in the queue is assigned the available collection resource. A similar timeliness check is included in all the remaining submodels. From here the RFI undergoes a delay based on its associated required quality and information source attributes. This delay represents the time taken to collect information and is taken from an array of expressions (discrete value, draw from a random probability distribution, mathematical function, etc.) indexed on the value of these attributes.

After completing the delay the RFI is assigned a value (based upon input parameters) representing the actual quality of information collected.

The *Processing and Exploitation* submodel contains an independent set of logic for processing and for exploitation, allowing RFIs to proceed through either or both of these processes. In the processing portion RFIs wait in a ranked queue based on priority for an appropriate processing resource identified by the RFIs information source attribute. The RFI then undergoes a delay based upon the information source and quality attributes. In the exploitation portion various information specialists (different *Arena* resources) are seized to represent the exploitation of the information represented by each RFI and to determine the effect on the information quality. RFIs wait for appropriate exploitation resources in another ranked queue based on priority. The exploitation delay is determined based upon attributes defining the RFI's information source and the required and actual information quality.

The *Analysis and Production* submodel consists of two separate sets of logic to allow RFIs to proceed through either or both of these processes. The analysis portion of the model is logically structured like processing and exploitation. Resources in this submodel represent analysts that can handle RFIs from any information source (all source analysts) or analyst assigned to only process RFIs from one of the available thirteen information sources. As in previous submodels RFIs wait in a ranked queue based on priority. The appropriate delay for analysis is determined based upon attributes defining the RFI's information source and the required and actual information quality. After the delay, the assessed effect of analysis on actual information quality is used to update the actual quality attribute. In the production portion of this submodel RFIs wait in a ranked queue based on priority for an appropriate resource as defined for the analysis process. As in the analysis logic, the appropriate delay for production is determined based upon attributes defining the RFI's information source and the required and actual information quality. After the delay, the assessed effect of production on actual information quality is used to update the actual quality attribute.

The *Dissemination and Integration* submodel consist of two separate sets of logic to allow RFIs to proceed through either or both of these processes. The dissemination portion of this model reflects the user process of acquiring information that has been through some portion of the intelligence process. RFIs wait in a ranked queue based on user priority for a resource from a set of user analysts. The appropriate delay for dissemination is determined based upon the user type. After the delay, the assessed effect of the dissemination on actual information quality based on the information source is used to update the actual quality attribute. Integration in this context refers to the individual users integrating information into their processes, not information integration as related to exploitation or analysis. If this portion of the model is not used, then timeliness is determined by when a user receives requested information. If this portion of the model is used, then timeliness includes a delay to account for user integration. RFIs wait in a ranked queue based on user priority for a resource from a set of user analysts. The appropriate delay for integration is determined based upon the user type. After the delay, the assessed effect of the integration on actual information quality based on the information source is used to update the actual quality attribute.

In addition to these five submodels mimicking the outer intelligence cycle, the *Communications* submodel routes all RFIs between the other submodels and provides appropriate communication delays and resources. The *Evaluation* submodel collects RFI and system statistics. The primary measures of performance for our system are once again quality, quantity, timeliness, and information needs satisfaction (where INS is defined in our context as meeting both quality and timeliness). Upon examination of the raw data for any of these measures, it becomes clear that a composite measure of some type would be more useful than a measure such as the average information quality. In addition, a simple count of the number of requests that are processed provides limited insight. To address these issues we combine the quantity of requests with the other measures to come up proportions of the requests that meet defined quality and/or timeliness requirements. The use of these proportions allows clear comparisons between competing system configurations. Grouping these proportions into various categories, such as by user type or information source, also provides additional insight into the system performance. The following section briefing describes the input data requirements for the IPM and presents results and analysis using these measures of system performance.

Analysis

Before performing a simulation study with the IPM we clearly need to populate our model with data. Even in a high level model such as the IPM the amount of data required to characterize the system is extensive. The data required can be divided into the following basic categories: processing times, resources, effects on quality, and request properties. Everywhere in the model that a delay occurs a resource is required. In addition each time an RFI passes through some simulated process an update is made to the actual quality. The request properties are used to determine when and what type of RFIs are generated. For a complete listing of the required data for a simulation study see Pawling [2004]. As designed with a maximum of five users, thirteen information sources, and thirteen analyst specialties, there are over 2700 data entries available. Real world data needed to populate the IPM was not available due to its sensitivity and classification. So notional data was generated as a baseline for comparative study. For our sample study 830 data entries, many of these probability distributions, were selected. Another important aspect of our baseline model was distinguishing our five generic user types by the specific steps of the intelligence cycle required for each. User 1 required all steps, User 2 never required exploitation, User 3 never required analysis, User 4 never required analysis or production, and User 5 never required exploitation or production. In this way we were able to approximate different hybrid TPPU architectures, considering User 1 to represent TPED.

We considered our model of the intelligence process as a steady state simulation. Therefore, before performing production runs of the IPM for this study, pilot replications were analyzed to determine an appropriate truncation point (utilizing the deletion approach to reduce initialization bias) and a sufficient replication length to reach steady state performance. See Chapter 4 of Pawling [2004] for details. Production runs for the study consisted of 25 replications with a replication length of 1580 days with the first 120 days deleted. Results are discussed for our baseline system and a system with a 50% increase in exploitation times over the baseline.

Figure 3 shows the average proportion of requirements that were met in terms of quality, timeliness, or both for our baseline configuration by priority. As would be expected with a

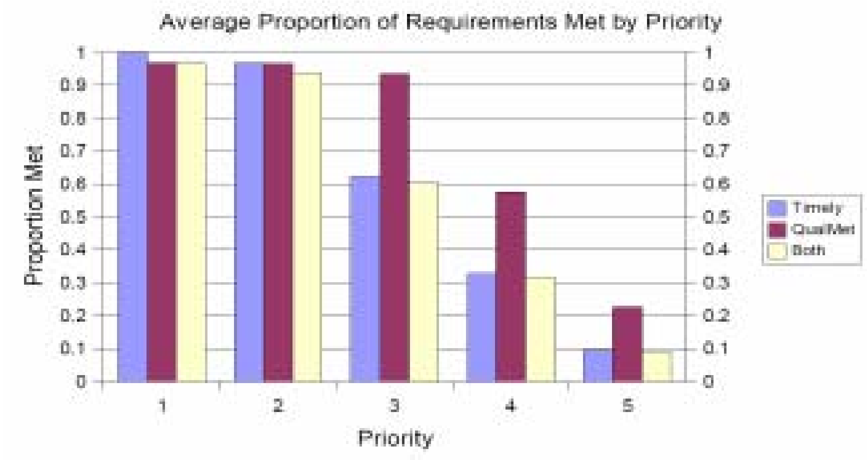

Figure 3. Baseline Average Proportion Met by Priority

priority based system, the highest priority requirements (with 1 being the highest) are on average almost always on time. As the priority goes down, so does the proportion that meet the timeliness requirement. Figure 4 shows the average proportion of requirements that were met in terms of quality, timeliness, or both for our baseline configuration by user type. This graph shows some variation in the proportion of quality requirements met but it is fairly consistent across users. Since User 1 requires all steps it is not surprising that there is a smaller proportion of timeliness requirements met.

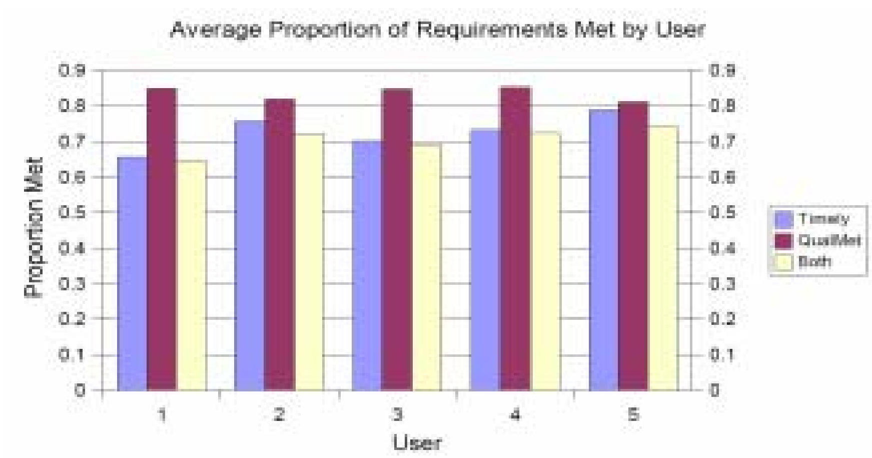

Figure 4. Baseline Average Proportion Met by User Type

In addition to our baseline model we also considered seven different cases varying different input parameters to examine various aspects of the IPM. See Pawling [2004] for more details. We include results below from our case 6 (labeled C6). This case was defined by increasing all the exploitation times by 50%. This was implemented by using a multiplicative factor of 1.5 when exploitation times were assigned in the model. Figures 5 and 6 show the proportion of both quality and timeliness requirements met by priority and user type respectively for C6 compared with the baseline (labeled BL). Figure 5 clearly shows that the system can only keep up with the top priority items for case 6. Priority 2 and 3 items have suffered a practically significant decrease in the proportion of requirements met, whereas the proportion of priority 4 and 5 items did not change significantly from the baseline. Figure 6 illustrates some significant differences with case 6 based upon the different steps of the intelligence cycle modeled for the different users. In both modeled systems Users 2 and 5 do not require exploitation except for possibly a small portion of additional requirements. This distinction explains why Users 1, 3, and 4 show a practically significant reduction in their proportion of requirements met for case 6 while Users 2 and 5 show no significant change.

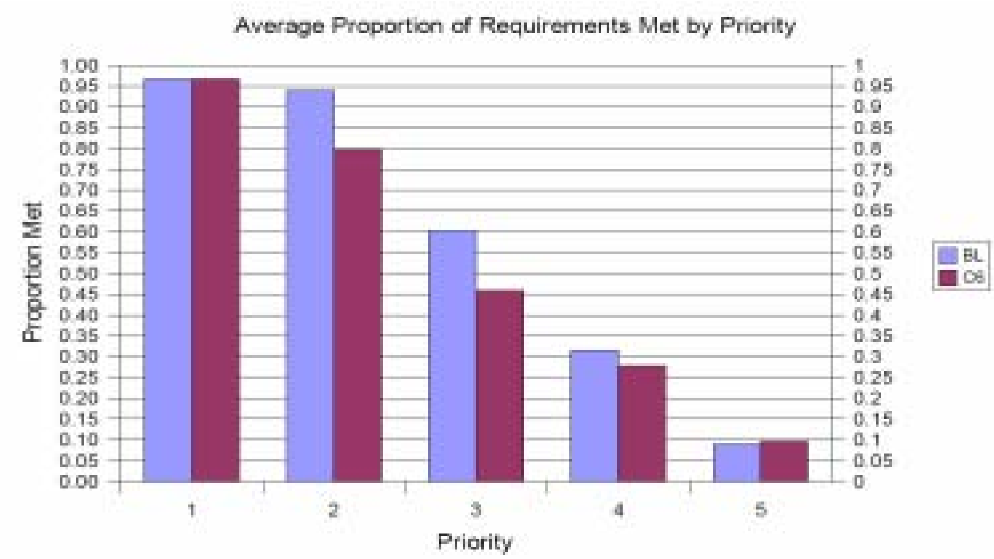

Figure 5. BL vs C6: Average Proportion Met by Priority

Conclusion

The focus of this effort was to develop a high level simulation model of the intelligence process. A review of prior efforts in modeling the intelligence process generally focused on the TPED architecture or required too much detail for a high level model. We developed the IPM from the high level perspective of the intelligence cycle presented in Joint Publications as well as subject matter expert input from U.S. national intelligence organizations. The benefit of this approach is that it is grounded on documented policy and procedures while taking a top down view of the process. The modular design of the IPM allows additional detail to be added to any particular submodel for future studies with little or no changes required for the rest of the model.

Verification and validation efforts were undertaken throughout the development of the model and included many conversations with subject matter experts and detailed walk-throughs of the IPM. The use of a notional baseline system and case studies designed to stress various aspects of the model lends additional credibility to the model given that implemented changes induced expected outcomes in the system performance. The two top areas for future enhancement to the model involve development of a more robust and realistic communications model and the incorporation of an information fusion capability.

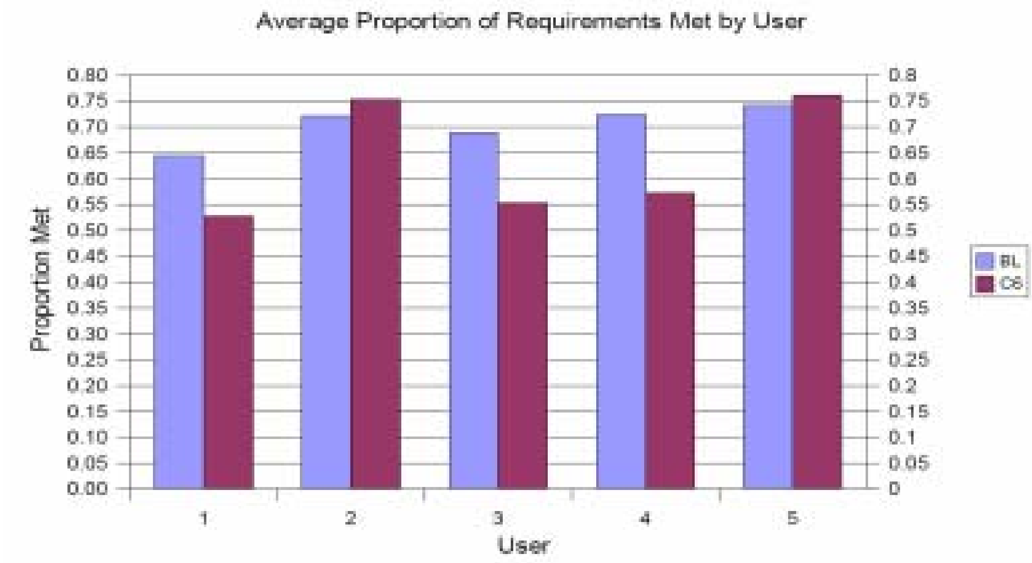

Figure 6. BL vs C6: Average Proportion Met by User Type

DISCLAIMER

The views expressed in this paper are those of the authors and do not reflect the official policy or position of the United States Air Force, Department of Defense or the U. S. Government.

References

Department of the Defense. *Doctrine for Intelligence Support to Joint Operations*. Joint Publication 2-0. Washington: GPO, 9 March 2000.

Kanewske, Charlie, Program Manager. *QUICM Software User's Manual*. Contract No. NRO000-98-D-2185. National Reconnaissance Office. April 2003.

Pawling, Carl R. "Modeling and Simulation of the Military Intelligence Process." MS Thesis. AFIT/GOR/ENS/04-09, Graduate School of Engineering and Management, Air Force Institute of Technology, Wright-Patterson AFB, OH. March 2004.

White House. The National Security Strategy of the United States of America. Washington: GPO, September 2002.