Proceedings of the 1987 Winter Simulation Conference
A. Thesen, H. Grant, W. David Kelton (eds.)

A SIMULATION APPROACH TO ANALYSIS OF EMERGENCY SERVICES AND TRAUMA CENTER MANAGEMENT

Thomas D. Clark, Jr.
Craig W. Waring
Department of Information & Management Sciences
The Florida State University
Tallahassee, FL 32306-1042

ABSTRACT

The operation of a hospital trauma center and associated transportation facility is discussed in this paper. A SLAM simulation model of the system is explained and example experiments with it discussed. Results indicate that significant improvement in patient service can be made through alternative staff scheduling patterns. Patient flow is particularly sensitive to the shift scheduling of physicians.

1. INTRODUCTION

The scheduling of diagnostic and clinical treatment for emergency patients has received some attention in the management science literature. Bolling (1972) developed a classical queuing model of an emergency room and many authors (Baker, 1976) address workforce scheduling under conditions of variable stochasticity. For example, Hershey and others (1974) employed simulation to compare nurse allocation. More recently, Carlson and others (1979) use a combination of optimization models and simulation in a walk-in clinic and Vassilacopoulos (1985) developed a dynamic programming formulation to address the allocation of doctors in emergency rooms. Each of these contribute to the theory of staffing and scheduling in stochastic systems but none provides the requisite flexibility for extensive experimentation. This is especially true when the extended system, including transportation, is addressed.

With simulation modeling, flexibility exists to approach the simultaneous scheduling of doctors, nurses and technicians in a variety of ways. This flexibility is required because of the complexity of the system created by the variety of variables present and the interactions among them. Since with emergency care, medical services are extended beyond the limits of the hospital, the transportation system must be included as must the dimensions of the physical facilities and availability of professional staff. The purpose of this paper is to detail a simulation approach to the problem with a specific objective of suggesting alternatives that will improve the scheduling of professional staff delivering emergency care.

A difficulty in scheduling emergency care is created because the exact nature of the care required for a specific patient is most difficult to predict. The inherent stochasticity of the system means that the emergency care system requires, to a much greater extent than almost any service system of similar complexity, the flexibility offered by idle or reserve resources. Diverse resources must be provided to insure adequate care of emergency patients suffering from sudden injury or serious illness. The costs of providing such care are significant, so a balance between unlimited flexibility and adequate flexibility to insure reasonable care must be found.

2. SYSTEM STRUCTURE

The Trauma Center of the Tallahassee Florida Regional Medical Center is the referent system for the discussion in this section. The hospital is a major public nonprofit medical facility with 771 beds serving a region of nine counties in north Florida and three in south Georgia. The area has a combined population of about 400,000 people. The trauma center includes a state-of-the-art ground ambulance service and a "life flight" helicopter ambulance service staffed by aeromedical paramedics. Seven experienced physicians and thirty-four nurses provide the staffing flexibility necessary for twenty-four hour operation.

The system has two major subsystems. One is composed of the interacting variables that produce the behaviors outside the trauma center and the other of those variables that produce behavior outside the center. Stabilization and transport are the primary activities of the first subsystem and patient evaluation, preliminary care and transition treatment are the primary activities of the second. This latter subsystem is focused upon in this paper.

The system's variables may be classified into four major subsets that can be used to guide model development and research design. The first set of variables labeled "policy variables," includes those which are controllable by the trauma center managers. The second set includes those stochastic variables or processes outside direct managerial control; the third set includes parameters or system constants; and the fourth set includes output variables or performance measures that are labeled "measures of merit." The variables for the system, shown in Table 1, will be used to discuss system behavior and measurement.

Table 1: System Variables

STOCHASTIC	POLICY	PARAMETERS	MEASURES OF MERIT
• Demand - Quantity - Type - Timing • Patient Treatment Time • Ambulance Response Time • Facility Preparation Time	• Physician - Scheduling - Staffing • Nurse - Scheduling - Staffing • Paramedical Personnel - Scheduling - Staffing • Service Disciplines • # of vehicles - air - ground	• Facility Structure	• Patient Total System Time • Medical Personnel Utilization • Patient Waiting Time to First Physician Contact

The demand for emergency services is quite complex but can be addressed easily by collapsing the range of possible patient conditions into four broad categories. Each patient is assigned one of four severity conditions (non-critical walk-ins, air transport critical, ground transport critical, critical walk-ins). These categories are established with the assumption that non-critical patients will not be transported by air or ground ambulance. Within each category, information about required treatment times was derived from hospital records. It is obvious that these categories can be decomposed to their lowest level if more detailed information was needed. This is not required for this phase of the research as the objective is to focus on personnel scheduling. Requisite variety in the model is achieved for the study objectives with only these four classifications (Beer, 1966).

Patients enter the system with exponential interarrival times. The mean of the arrival distribution shifts from 40 minutes in the pre-dawn hours to 15 minutes in the busier parts of the day. First contact takes one of the four forms previously mentioned with the probability that a given arrival will be of a given type also dependent on the time of the day. Records indicate that a higher proportion of patients are in need of immediate response in the late hours and are more likely to be transported by hospital paramedics and ambulance. Routine and critical walk-ins occur more in the evening than at other times of the day.

Some patients, because of critical need, may preempt currently working doctors or nurses. After the patient is cared for, he or she either is dismissed from the center or sent into the hospital admissions process. The system's structure may be represented by a series of queus in a network. Such structures are amenable to modeling with the Simulation Language for Alternative Modeling (SLAM) (Pritsker, 1986) using a process orientation or world view. The SLAM system model is discussed in the next section.

3. MODEL STRUCTURE

The structure of the model is shown in the SLAM network diagrams in Figures 1, 2, and 3. There are three major sectors in the model. One sector (Figure 2) manages patients that because of their condition preempt currently working doctors or nurses. A second sector (Figure 1) manages patients that do not preempt servers and the third (Figure 3) manages the work scheduling of the professional staff. Doctors and nurses are modeled with the SLAM "resources" feature. All patients enter the system in the same manner and are appropriately routed. The following discussion of patient flow will best be followed by referring to the appropriate figure.

There are two doctors available for nine hours of the day and only one on duty for the remaining fifteen hours. To properly route patients, given the physician availability, there are three branches emanating from the node labeled WDW at the left of Figure 1. The first branch is taken to await a doctor who has already seen the patient at least once; the second branch is taken if the patient has not yet seen a doctor and two doctors are on duty; and the third branch is taken if the patient has already seen that doctor or if that doctor is the only one on duty.

When a patient is seen by either doctor, an entity representing the service is processed to the node labeled C8, indicating the beginning of the first interaction of patient and doctor. The C8 node collects statistics on the amount of time between when a patient clears transportation and triage and the time when he or she sees a doctor for the first time.

From C8, entities are simultaneously sent to the routing node G01, shown at the bottom left of the figure. From there, a maximum of two of five branches will be taken. One of the first two branches will always be taken to free the appropriate doctor; the third branch will be taken to the node labeled WNX if the patient has been assigned a nurse from the pool of nurses, and if there have been fewer than three doctor interactions thus far; the fourth branch is taken if the patient has not yet seen a nurse; and the fifth branch is taken if the patient has been assigned the last available nurse and has not seen her more than once before. When patients being seen by a nurse complete service, they are routed to assign nodes for status updates and then to free nodes to release the appropriate nurse. From there the patients leave the system.

Entities which are routed from G01 to free doctors are then routed to the node labeled G02 from which one of three routes

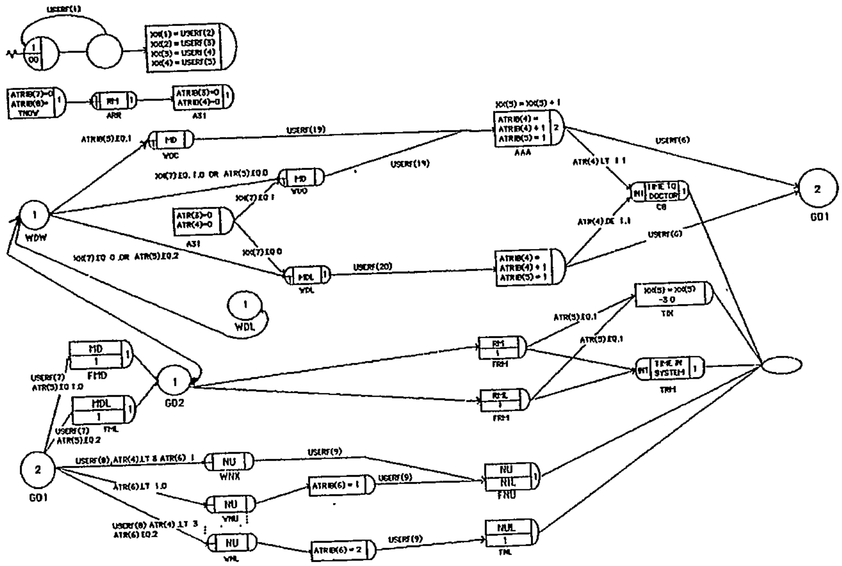

Figure 1. SLAM Model Structure: Normal Flow

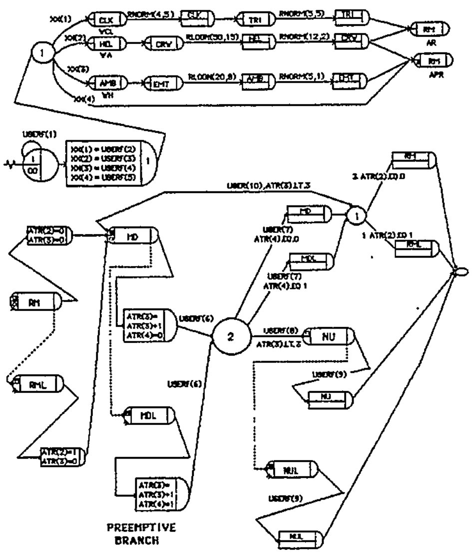

Figure 2. SLAM Model Structure Preemptive Branch

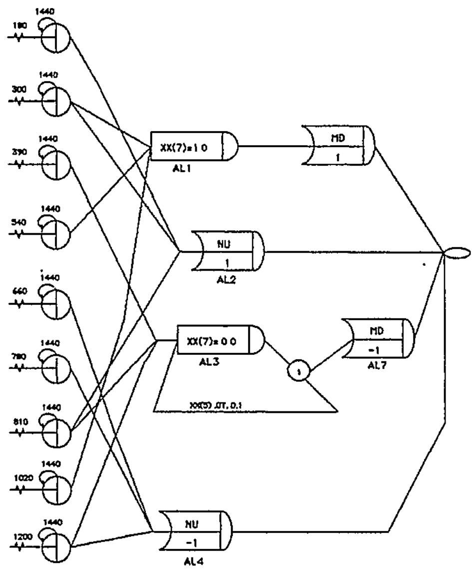

Figure 3. SLAM Model Structure: Staff Management Process

is taken. The first route is back to WDW for patients which have not had three interactions with doctors. The second and third routes are taken for patients who have completed service. The treatment rooms are freed and appropriate statistics collected when a patient completes service. System time for a patient reflects the time between triage and departure while doctor waiting time reflects the time from initial entry into the system until the first encounter with a physician.

Figure 2 contains the network that manages patient routing when immediate care is required. If a physician or nurse is busy, he or she is preempted (current service stopped) and assigned to the critical patient. Of course, this only occurs when the staff member is not already serving a critically ill patient. The network essentially is the same as before except "preempt" rather than "await" nodes are used to represent waiting patients. Figure 3 contains the network used to control the scheduling of doctors and nurses. These schedules are the subjects of a later discussion.

The model was extensively tested to insure its behavior matched that of the current system. Since adequate data about patient type and flow existed, the major test used was behavior reproduction using an ultra empirical approach (Naylor and others, 1966). The behavior reproduction tests were supplemented by discussions with practicing managers about system and model behavior. The verification and validation process produced confidence that a useful experimental model had been developed. Representative experiments performed with the model will be discussed in the next section.

4. EXPERIMENTAL DESIGN

The measures of merit introduced in Table 1 indicate that there are a number of possible ways to judge the effectiveness of this system. One proposed by the Medical Director of TMRMC was the time a patient spends waiting for initial contact with a doctor. Not surprisingly, review of the literature (Groom and Au, 1969; Vassilacopoulos, 1985; Willemain, 1977) indicated that patient satisfaction was more sensitive to the initial response times than total system time. That is, if total waiting time is the same, patients are more satisfied if out of three encounters (an average number) with the doctor, the first wait is shortest and the last longest. This is a good measure of behavior if only one output variable is employed.

There are other variables, and perhaps a combination of variables, that would actually provide an indication of the best policy for system management. In the case of multiple response variables, use of Multiple Analysis of Variance (MANOVA) would define the best policy given the linear combination of output measures. To illustrate the possible alternatives available with various policy structures, examples will be discussed.

Two representative policies for scheduling doctors and three for scheduling nurses will be employed. Currently, doctors are scheduled as follows:

<u>Doctor Scheduling</u> (Policy 1)

2400-1300:	one doctor
1330-1900:	two doctors
1900-2100:	one doctor
2100-2400:	two doctors

An alternative scheduling policy discussed with management and derived from observation of patient arrival structure is:

<u>Doctor Scheduling</u> (Policy 2)

2400-0900:	one doctor
0900-1030:	two doctors
1030-1300:	one doctor
1300-1730:	two doctors
1730-2100:	one doctor
2100-2400:	two doctors

The current schedule for nurses is:

<u>Nurse Scheduling</u> (Policy 1)

1730-0130:	three nurses
0130-0700:	two nurses
0700-0900:	three nurses
0900-1500:	four nurses
1500-1700:	three nurses
1700-1730:	two nurses

Two alternative scheduling policies were developed from observation of the system and initial model behavior:

<u>Nurse Scheduling</u> (Policy 2)

2300-0700:	two nurses
0700-1500:	four nurses
1500-2300:	three nurses

<u>Nurse Scheduling</u> (Policy 3)

2300-0700:	one nurse
0700-2300:	four nurses

This policy structure produces an experimental design matrix of the following form:

		Nurse Policy		
		1	2	3
Doctor Policy	1	measure(s) of merit		
	2			

As suggested, several alternatives are available for testing policy choices within this design. First, a completely randomized two factor design will be used to determine if staffing policy will have an effect on the waiting times patients experience. The form of the model tested is:

$$Y_{ijk} = \mu + \beta_j + \theta_k + \beta_j\theta_k + \varepsilon_{ijk}$$

Where: y is the average time a patient waits to see a doctor after transportation and triage.

μ is the grand mean average waiting time for all sets of conditions.

β_j is the treatment effect of the jth scheduling policy for doctors.

Θ_k is the treatment effect of the kth scheduling policy for nurses.

$\beta\Theta_{jk}$ is the interaction term for the joint effects of the two policies at the j and k levels.

ε_{ijk} is the error effect for replication in i treatment combination jk.

This single response variable design may also be used for the other candidate output variables. In addition to the time a patient initially spends waiting for a doctor, the design can be employed for the time a patient spends in the system, the utilization rate for doctors and the utilization rate for nurses. If a series of univariate measures are employed, the following set of null hypotheses would be tested:

Ho(1) Doctor scheduling will make no difference in the time patients spend waiting to see a doctor.

Ho(2) Doctor scheduling will make no difference in the total time patients spend in the system.

Ho(3) Doctor scheduling will make no difference in the utilization rate of doctors.

Ho(4) Doctor scheduling will make no difference in the utilization rate of nurses.

Ho(5) Nurse scheduling will make no difference in the time patients spend waiting to see a doctor.

Ho(6) Nurse scheduling will make no difference in the total time patients spend in the system.

Ho(7) Nurse scheduling will make no difference in the utilization rate of doctors.

Ho(8) Nurse scheduling will make no difference in the utilization rate of nurses.

Ho(9) The interaction between doctor and nurse scheduling will make no difference in the time patients spend waiting to see a doctor.

Ho(10) The interaction between doctor and nurse scheduling will make no difference in the total time patients spend in the system.

Ho(11) The interaction between doctor and nurse scheduling will make no difference in the utilization rate of doctors.

Ho(12) The interaction between doctor and nurse scheduling will make no difference in the utilization rate of nurses.

These hypotheses were tested using a two-way ANOVA for each one. A sample of the results are shown in Table 2 for the first, fifth and ninth hypotheses. These results are typical of all of the two-way tests, with the scheduling of physicians having a significant effect on waiting time and nurse scheduling and the interaction between doctor and nurse scheduling not having a significant effect. The amount of time a patient spends waiting to see a doctor for the first time will be reduced by changing the physician scheduling policy from its current structure to the one proposed.

Table 2:
ANOVA Results For Ho(1), Ho(5), Ho(7)

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	484.832	3	161.611	2.470	.081
DOCPOL	475.822	1	475.822	7.273	.011
NURPOL	9.010	2	4.505	.069	.934
2-WAY INTERACTIONS	38.530	2	19.265	.294	.747
DOCPOL NURPOL	38.530	2	19.265	.294	.747
EXPLAINED	523.361	5	104.672	1.600	.190
RESIDUAL	1962.763	30	65.425		
TOTAL	2486.124	35	71.032		

The results of the other tests of hypotheses are shown in Tables 3, 4, and 5. These indicate that the utilization rate of nurses may be improved by the scheduling change for both doctors and nurses, but that little improvement in physician utilization would occur. Physicians are busy about 35% of the time and nurses between 40 and 45% of the time. These rates indicate the nature of the work and likely could not be improved no matter what combination of scheduling policies was followed.

Table 3:
ANOVA Results For Ho(2), Ho(6), Ho(10)

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	802.448	3	267.483	5.423	.004
DOCPOL	790.753	1	790.753	16.033	.001
NURPOL	11.695	2	5.848	.119	.889
2-WAY INTERACTIONS	16.764	2	8.382	.170	.845
DOCPOL NURPOL	16.764	2	8.382	.170	.845
EXPLAINED	819.212	5	163.842	3.322	.017
RESIDUAL	1479.638	30	49.321		
TOTAL	2298.850	35	65.681		

Table 4:
ANOVA Results For Ho(3), Ho(7), Ho(11)

SOURCE OF VARIATION	SUM OF SQUARES	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	.002	3	.001	.410	.747
DOCPOL	.002	1	.002	1.230	.276
NURPOL	.000	2	.000	.000	.999
2-WAY INTERACTIONS	.000	2	.000	.004	.996
DOCPOL NURPOL	.000	2	.000	.004	.996
EXPLAINED	.002	5	.000	.248	.938
RESIDUAL	.052	30	.002		
TOTAL	.054	35	.002		

Table 5:
ANOVA Results For Ho(4), Ho(8), Ho(12)

SOURCE OF VARIATION	SUM OF SQUARE	DF	MEAN SQUARE	F	SIGNIF OF F
MAIN EFFECTS	.080	3	.027	10.674	.001
DOCPOL	.011	1	.011	4:388	.045
NURPOL	.069	2	.035	13.817	.001
2-WAY INTERACTIONS	.008	2	.004	1.656	.208
DOCPOL NURPOL	.008	2	.004	1.656	.208
EXPLAINED	.089	5	.018	7.067	.001
RESIDUAL	.075	30	.003		
TOTAL	.164	35	.005		

As mentioned earlier it is possible to test a multivariate response model of the form:

$$Y_{ijk} + W_{ijk} + X_{ijk} + Z_{ijk} = \mu + \beta_j + \theta_k + \beta_j\theta_k + \varepsilon_{ijk}$$

where:

Y is the average time a patient waits to see a doctor after transportation and triage.

W is the average system time for a patient.

X is the utilization rate for doctors.

Z is the utilization rate for nurses.

and the other variables are as before.

With this formulation, the hypothesis tested is:

$$H_o: \underline{\mu}_1 = \underline{\mu}_2 = \underline{\mu}_3 = \underline{\mu}_4 = \underline{\mu}_5 = \underline{\mu}_6$$

where $\underline{\mu}_i$ is a vector of the means of the response variables for each possible policy in the design matrix. Six replications of each policy were conducted. The values of Hotelling's Trace, Wilk's lamda, and Pillais' statistic indicate both nurse and doctor scheduling policy have a significant effect on system behavior. The Eigenvalues for the roots indicate that the second nurse scheduling policy is significantly different than the others. This means that following a scheduling policy for doctors and nurses different from the current policies will improve system behavior. It appears the combination of doctor scheduling policy number two and nurse scheduling policy number two is the best course to follow.

5. SUMMARY AND CONCLUSIONS

A SLAM model of a typical Trauma Center was presented and discussed. Operation of the model and experiments with it were illustrated. The model provides the flexibility to test a number of staff scheduling alternatives and provides data that are useful for a variety of different statistical analysis approaches. To continue development, experiments with alternative transportation structures would be beneficial.

REFERENCES

Baker, K.R. (1976). Workforce Allocation in Cyclical Scheduling Problems: A Survey. Operational Research Quarterly 27, 155-167.

Beer, S. (1966). Decision and Control. New York: John Wiley & Sons.

Bolling, W.B. (1972). Queuing Model of a Hospital Emergency Room. Industrial Engineering 4, 26-31.

Bloom, B.S. and Fendrick, A.M. (1987). Waiting for Care: Queuing and Resource Allocation. Medical Care 25, 61-85.

Carlson, R.C., Hershey, J.C. and Kropp, D.N. (1979). Use of Optimization and Simulation Models to Analyze Outpatient Health Care Settings. Decision Sciences 10, 412-433.

Groom, H.R. and Au, T. (1969). The Simulation of Hospital Patient Treatment. Third Conference on Application of Simulation. Los Angeles, CA, 269-278.

Hershey, J.C., Abernathy, W.J. and Baloff, N. (1974). Comparison of Nurse Allocation Policies - A Monte Carlo Model. Decision Sciences 5, 58-72.

Ingram, J.A. and Asbury, A.J. (1983). Patient Administration-II. British Medical Journal 287, 667-670.

Jones, A.W., Sahney, V.K. and Kurtoglue, A. (1983). A Discrete Event Simulation for the Management of Surgical Suite Scheduling. Record of Proceedings, the 16th Annual Simulation Symposium. Ed. L.A. Holbrook, Silver Springs, MD., 263-278.

Kamenetzky, R.D., Shuman, L.J. and Wolfe, H. (1982). Estimating Needs and Demand for Prehospital Care. Operations Research 30, 1148-1166.

Law, D.K. and others (1982). Trauma Operating Room in Conjunction with an Air Ambulance System: Indications, Interventions, and Outcomes. The Journal of Trauma 22, 759-766.

Naylor, T.H., Balintfy, J.L., Burdick, D.S. and Chu, K. (1966). Computer Simulation Techniques. New York: John Wiley & Sons.

Pritsker, A.A.B. (1986). Introduction to Simulation and SLAM II. (3rd Ed.). West Lafayette, IN: Systems Publishing Corporation.

Shahein, H.I.H. and Zaky, M.M. (1983). ESMIS-A Computer-Based Emergency Medical Services Management Information System. Part 1: Design Procedure. International Journal of Bio-Medical Computing 14, 451-462.

Statistical Digest 1985, Leon County Planning Department, Tallahassee, FL.

Tallahassee Memorial Regional Medical Center Annual Report, 1983-1984.

Vassilacopoulos, G. (1985). Allocating Doctors to Shifts in an Accident and Emergency Department. Journal of the Operational Research Society 36, 517-523.

Willemain, T.R. (1977). Emergency Medical Systems Analysis. Lexington, MA: Lexington Books.

AUTHORS' BIOGRAPHIES

THOMAS D. CLARK, JR. is an Associate Professor of Information and Management Sciences and director of the Center for Information Systems Research at The Florida State University. He has been active in the Winter Simulation Conference for ten years as an author, session chairman, and track chairman. He has taught a variety of computer simulation courses at the graduate level and has supervised over sixty masters and doctoral theses and dissertations that have employed simulation analysis. He had extensive experience in logistics management before joining a university faculty and has published a variety of papers using simulation in study of logistics and operations management systems. His current research interests include investigation of a methodology to value management information, the structure of multi-echelon service operations and the nature of decision support system evolution. He is a member of ACM, ORSA, TIMS, SIM, SCS, DSI and SGSR.

Thomas D. Clark, Jr.
Department of Information & Management Sciences
College of Business
The Florida State University
Tallahassee, Florida 32306-1042
(904) 644-5508

CRAIG W. WARING is a Ph.D. student in the Information and Management Sciences Department at The Florida State University. His dissertation deals with the relationship of variables in the design and implementation process for decision support systems. His research interests also include the operation of service systems especially those where complex staff scheduling problems exist. He is a member of SCS and DSI.

Craig W. Waring
Department of Information & Management Sciences
College of Business
The Florida State University
Tallahassee, Florida 32306-1042
(904) 644-5508