Relationship Between Risk and Complexity in System Using Connection Based Metric

João R. B. Paiva[1,2]
Viviane M. Gomes[1,2,3]
[1]Electrical, Mechanical & Computer Engineering School
University of Goias, Brazil
Email: jricardobpaiva@yahoo.com.br

Márcio R. C. Reis[1,2,3]
Gabriel Wainer[3]
[2]Experimental & Technological Research and Study Group
Federal Institute of Goias, Brazil

Wesley P. Calixto[1,2,3]
[3]Department of Systems and Computer Engineering
Carleton University, Canada.

***Abstract*—This paper proposes methodology to relate risk in systems using complexity metric based on active connections as a criterion for selection of decision support methods. A case study of application of complexity metric is presented with emphasis on choosing of the most appropriated decision support method for approach. Obtained results indicate that the treated problem can be classified in simple solutions area. It is concluded that system performance measure and complexity can be used in an associated way as a criterion for selection of decision and risk support methods.**

***Index Terms* - risk, complexity, decision support methods.**

I. Introduction

The concept of complexity from the observer's point of view or the individual who performs some interaction with the system, is subjective and associated with the perception of it and not only the object properties in question [1]. Thus, if the amount of knowledge about the object of study increases, it becomes less complex. Similarly, its perceived complexity decreases (being considered simpler) as knowledge about the object of study increases [2].

To measure the complexity of a system in mathematical and independent terms of subjectivity, metrics are required for using as a measure the main characteristics of this system. By using complexity metrics, it is possible to establish relations between two measured values. In this way, there is sense in using the metric only if the goal is to compare between two or more systems. Therefore, obtained value for the complexity of a system aggregates knowledge only if it is compared to another value of obtained complexity for another system or for the same system, towards different conditions. The complexity of a system can be linked to its operational risk [3], [4].

Risk is the combination between the possibility of occurrence of undesirable situations and their respective repercussion, in case of this situation occurs [5]. If the possibility of occurrence of a particular situation is remote, but the repercussion of its occurrence is negative, there is a high risk. If the repercussion of a certain situation is not negative, but its possibility of occurrence is high, it is also considered high risk [6].

Decision support methods can be used as tools in risk management. In [7] it is defined that there are four types of decision support methods: deterministic, heuristic, stochastic and inferential, which can be selected according to the strategy adopted to solve the problem.

This paper presents a methodology for selecting decision support methods considering the system risk from the perspective of complexity metric, based on active connections.

In the section II used methodology is presented, with emphasis on aspects of complexity and decision support methods, as well as the used metric and the presented case study. The results obtained are exposed and discussed in the section III. Reached conclusions are presented in section IV.

II. Methodology

A. Complexity and Decision Support Methods

Stacey's study [8] uses two parameters as a reference to distinguish the level of complexity of a subject or an object. The first is the level of disagreement on a given subject and it is related to the definition of the object treated and what the requirements to describe it. The second parameter is the level of uncertainty that exists on this same object and it is related to how it can be developed, realized or constructed and the existing technology for it. The Figure 1 illustrates the organization of these ideas in the scheme known as Stacey's Matrix.

In Figure 1 it is possible to observe that in the proximity of (0.25;0.25) is the zone that aligns agreement with the description of what the object is treated and certainty that there are familiar technologies to manipulate it (region of simplicity). When there is no agreement on the description of the object being treated and the technologies to manipulate it are few or nonexistent, it works in the zone of anarchy or chaos, found in the proximity of (1.00;1.00). Situations where it is far from the agreement but close to certainty or far from certainty but close to the agreement are considered

The authors would like to thank the National Council for Scientific and Technological Development (CNPq), the Foundation for Research Support in the State of Goiás (FAPEG) and the Coordination for the Improvement of Higher Education Personnel (CAPES) for financial assistance to this research.

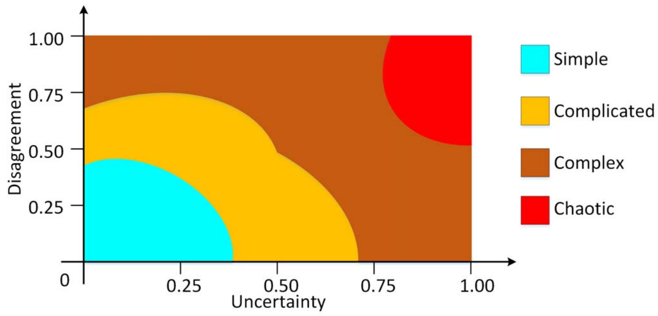

Figure 1. Stacey's Matrix.

complicated, and are located close to the (0.50;0.50). In this way, complexity arises between the complicated and the chaos, in the zone (0.75;0.75). The complexity in Stacey's Matrix is the area where problems are difficult to fit, for which cause and effect relationships are unclear and the problem evolves while being treated [8].

According [7] there is a correlation among complexity, uncertainty and decision support methods, as shown in Figure 2. In abscissas's axis, in the points near the 0 the uncertainty is low, and at points close to 1.00 the uncertainty is high. On the ordinate's axis, in points near 0 the complexity is low, and in the points near 1.00 the complexity is high. The relationship between complexity and uncertainty defines the most appropriate decision support method for solving the problem.

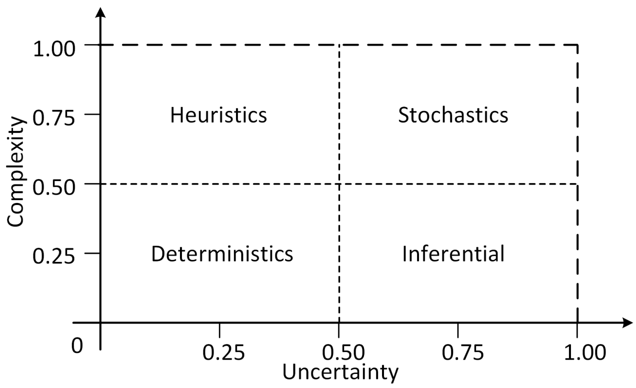

Figure 2. Correlation among complexity, uncertainty and decision support methods.

B. Complexity Metric Based on Connections

Several metrics for calculating complexity were developed based on system size, entropy, information, cost, hierarchy, organization, among other criteria [9]. Complexity metrics system can be grouped according to various criteria, such as the size of the system (in terms of number of elements or connections) or the amount of functional requirements it must meet.

One of the ways to measure the complexity of a system is by using the number of connections of its elements. For a k element system, each element can connect with up to $k-1$ elements. Therefore, the maximum number of connections in a system grows proportionally to the number of system elements. The method used to measure system complexity considers the active connections during its dynamics, considering the relation among entities, resources and queues, according to the expression (1) [10].

$$\gamma(s) = -\sum_{i=1}^{\rho} p(c)_i \cdot \log_2 p(c)_i \tag{1}$$

where $\gamma(s)$ is the system complexity, ρ is the number of active connections in the instant t and $p(c)_i$ is the probability that a given connection i occurs, given by:

$$p(c)_i = \frac{1}{n_e \cdot (n_r + n_f)} \tag{2}$$

where n_e is the number of entities of the system in the instant and t, n_r is the number of system resources and n_f is the number of queues.

The log_2 of expression (1) arises from the probability ψ that the connection is active or inactive. This is, $\psi(active, inactive) = (\frac{1}{2}, \frac{1}{2})$. Considering normalized values, if $\gamma(s) = 1$, so $\gamma(s) = -(\frac{1}{2} \cdot log_x \frac{1}{2} + \frac{1}{2} \cdot log_x \frac{1}{2}) = log_x 2$ and $log_x 2 = 1$ if and only if $x = 2$. The total number of possible system states for the configuration in instant t will be $w = 2^N$, which N is the number of elements in M and $log_2 2^N = N$. The number of active connections ρ at time t is given by:

$$\rho = \sum_{j=1}^{k} n_{c_j} \cdot n_{e_j} \tag{3}$$

where k is the number of entity states, n_{c_j} is the number of active connections per entity in the state j and n_{e_j} is the number of entities in the state j.

The metric application starts next to the system simulation process. During the simulation events occur, which cause system state change and change the active connections between its elements. At each system configuration, at time t, complexity will be calculated. This measurement cycle will be repeated with each change of state during the simulation. At the end of the simulation, as a result, the behavior of complexity relative to all states will be obtained.

C. Case Study: Distribution Center

The Distribution Center problem consists of order delivery logistics on demand. After arriving at the Distribution Center, the request waits in queue for loading. When resources are available (dock, truck and team of four loaders) the first order of the queue is loaded [11]. When the load is completed, the truck leaves for delivery, and the dock and loaders are released for re-loading. After delivery of the order to the consignee and return to the Distribution Center, the empty truck is again available for loadings. Figure 3 illustrates the Distribution Center.

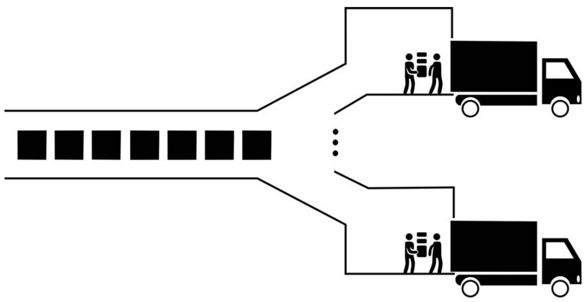

Figure 3. Distribution Center.

The Distribution Center is a typical discrete event system, which features entities, queues, and resources. Entities are the requests, which wait in queue for availability by resources: docks, trucks and loaders. The set of discrete states concerning the requests are: i) waiting in queue, ii) being loaded and iii) being transported. Based on the states, it is possible to determine how many resources (docks, trucks and loaders) are being used at every instant of time t. From the viewpoint of the distribution center, there are the following events: a) receiving new request, b) allocating group of loaders to perform charging, c) starting using the dock, d) allocating truck, e) beginning loading truck, f) finishing loading truck, g) finishing using dock, h) deallocating group of loaders, i) transporting request to the recipient, and j) deallocating truck.

The Distribution Center is an open system, as new entities can be integrated into the system during its operation. Thus, the number of entities (requests) and demand for resources (docks, trucks and loaders) may vary over time. The delivery time T_e of the orders, which corresponds to the time spent from the arrival of the order to the Distribution Center, until the moment it is delivered to the recipient, is composed by the sum of the times: i) waiting in queue until loaded t_{p_f}, ii) order upload t_{p_c} and iii) transportation of the order from the Distribution Center to the recipient t_{p_t}, given by (4):

$$T_e = t_{p_f} + t_{p_c} + t_{p_t} \tag{4}$$

The number of active connections is determined by the number of requests, and each queued request will add one connection to the system (with the request in front of it), each request being loaded will add three connections (one with the dock, one with the truck and one with the team of loaders) and each order being transported will contribute with one connection in the system (with the truck). In this way:

$$\rho = 1 \cdot n_{p_f} + 3 \cdot n_{p_c} + 1 \cdot n_{p_t} \tag{5}$$

where n_{p_f}, n_{p_c} and n_{p_t} are the number of requests in queue, of orders being loaded and being transported, respectively.

The number of entities n_e that each resource can match will be equal the total number of requests in the system at time t, assuming that each entity (request) can be served by any system resource. Thus n_e is the sum of n_{p_f}, n_{p_c} and n_{p_t}. The value of n_r is equal to the sum of the number of docks, trucks, and loader teams in the system. The value of n_f corresponds to the number of queues adopted in the system modeling, in this case, one queue.

III. RESULTS

The Distribution Center model was simulated with the following dynamics: i) the arrival of orders occurs in time intervals t_1, ii) the loading of each truck lasts the time t_2, iii) when a truck leaves to deliver, the dock and the team of loaders are released for reloading, iv) the transportation of each order to the recipient takes time period t_3 and v) after delivery, the empty truck returns to the Distribution Center during a period of time t_4.

The probabilistic distributions adopted for each time t_n of this case study and the respective values for its parameters in minutes are arranged in the Table I. The choice of these probabilistic distributions is based on the literature of discrete events systems, in which they are often used to model the processes of arrival in queue of load and delivery. In Table I the variables $\overline{m}$ and σ are the means and the standard deviation, respectively.

Table I. PROBABILISTIC DISTRIBUTIONS ADOPTED FOR THE DISTRIBUTION CENTER.

Time	Distribution	Parameters
t_1	Exponential	$\overline{m}$: 120
t_2	Normal	$\overline{m}$: 100, σ: 30
t_3	Uniform	Interval: $[120 - 240]$
t_4	Uniform	Interval: $[120 - 240]$

The number of docks used in the simulation ranged from 1 to 10, trucks from 1 to 15, and shipper groups from 1 to 10, producing 1500 scenarios. Simulation was done for 180 days for each scenario, considering 24 daily operation hours and each truck loaded with only one order at a time. In the 1500 simulated scenarios were calculated the delivery time T_e, in minutes, and the complexity $\gamma(s)$ of the system, presented in normalized form, in Figure 4.

As the simulation is a combination of resources: docks, trucks and loader groups, in this order, in Figure 4, the T_e peaks (in blue) correspond to combinations where only one truck was available. With each change in the number of docks, one truck was used for 10 scenarios, causing the highest values of T_e and $\gamma(s)$. The oscillations that occur in the range between 0.1 and 0.3 of $\gamma(s)$ are derived from the variations in the number of trucks, indicating greater sensitivity of this parameter of the system.

The lowest complexity found in the 1500 scenarios was $\gamma(s) = 0,2697$, which corresponds to the third lowest delivery time and the scenario with the maximum number of available resources. The worst delivery time $T_e \approx 340$ times greater than the shortest time, occurred in 100 different scenarios, as illustrated in Figure 4, in the 10 peaks, where 10 scenarios occur. The highest complexity obtained $\gamma(s) = 2.7906$, was for the worst time, since the calculation of both the delivery time and the complexity, considers the requests remain in the queue. Since each queued request corresponds to a connection, the larger the queue, the greater the time T_e and the greater will be $\gamma(s)$.

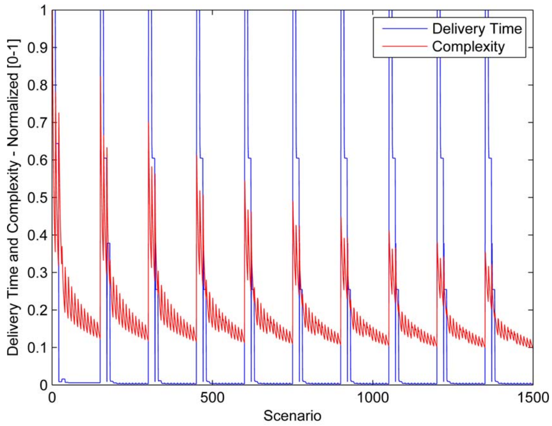

Figure 4. Relationship between Delivery Time and Complexity.

Considering the complexity and performance measurement adopted, T_e, can be developed metric that indicates the cost in terms of complexity for each minute of permanence of request in the system (R_C), given by the expression (6).

$$R_C = \frac{T_e}{\gamma(s)} \qquad (6)$$

On Table II, the values are found for R_C. It is observed that the lowest value found refers to the time of 815.7777 minutes, whose presented scenario reduced amount of resources. Even so, the delivery time was 115 times less than the worst case and only 3 times higher than the best case.

Table II. RELATIONSHIP BETWEEN DELIVERY TIME × COMPLEXITY.

T_e	$\gamma(s)$	R_C	Docks	Trucks	Groups
276.9357	0.2963	934.6463	10	15	6
291.6034	0.4083	714.1128	2	11	7
297.8970	0.5892	505.5827	3	7	2
322.1149	0.4690	686.8121	2	6	9
815.7777	2.0263	402.5947	1	3	1

One way to measure risk is by analyzing the maximum, average and minimum values of complexity. If the average value of the complexity is close to the minimum value the system has low risk and if the average value of the complexity is close to the maximum value, the system is high risk. It should be taken into account that this metric measures the risk of failure between the connections.

Analyzing the average values of complexity and the relation R_C, it is noted that both are closer to the minimum values than the maximum values obtained. This indicates that the operated system , during the simulations, with complexity close to the minimum in most scenarios. In this way, the operational risk of the system is considered reduced, as well as the number of active connections during its operation. Also, it is observed that the system worked on most of the simulation with relation between performance measure and complexity (R_C) reduced, indicating the use of optimized scenarios in most simulated configurations.

The choice of the most appropriate decision support method to solve the optimization problem of any model can be performed by normalizing the mean values of the complexity and measure of system performance and analyzing the Figure 5, obtained with the overlap of Figure 2 with Figure 1.

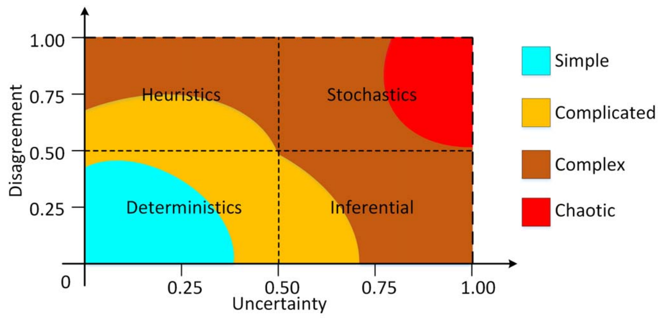

Figure 5. Overlap Decision Support Methods and Stacey's Matrix.

If the system is in the region of simplicity in Figure 5, methods such as deterministic are able to perform optimization. As the system becomes more complex, more elaborate methods, which generally employ artificial intelligence, are needed to perform the optimization. Analyzing the Figure 5 it is observed that for problems located in the other regions (complicated, complex and chaotic) heuristic, stochastic and inferential methods are recommended, depending on the degree of complexity.

For the Distribution Center problem, normalized mean values are 0.2 for complexity and 0.1 for R_C. Adopting correspondence in Figure 5, it is observed that the point (0.2;0.1), referring to the Distribution Center problem, is located in the region of simple problems, for which deterministic methods are indicated for resolution.

IV. CONCLUSION

Measure of performance and complexity, associated, may be used as criteria for selection of decision support methods. As system complexity increases, more robust decision support methods are demanded. The positioning of the system in the different levels of complexity of the Stacey's Matrix might help to choose the most appropriate method for its resolution. This positioning can be performed based on results obtained by applying the presented complexity metric, in order to minimize risks. Thus, in the case of operational risk, undesirable situations are those related to the possibility of failure in the operation of the system. The higher the number of connections in the system, the more fault points there are. Since the complexity metric used is based on connections, it can be used as a risk indicator in the operation of systems. Operational risk depends on the system dynamics. The system that has the highest number of connections during its operation, presents greater operational risk. In this way, the complexity metric can help in choosing the best operational approach of this system. The financial risk is related to situations where there is possibility of economic loss. When system resources are underutilized, there is cost disadvantage as it would be possible to maintain the same operating dynamics and performance criteria with fewer resources. In the Distribution Center problem, the complexity metric can be used in conjunction with the system performance measurement for resource sizing (number of docks, trucks and loader teams) and valuation of both operational and service costs.

REFERENCES

[1] H. A. Simon, "How complex are complex systems?," in *PSA: Proceedings of the biennial meeting of the Philosophy of Science Association*, pp. 507–522, JSTOR, 1976.

[2] Y. Bar-Yam, "Complexity rising: From human beings to human civilization, a complexity profile," *Encyclopedia of Life Support Systems (EOLSS), UNESCO, EOLSS Publishers, Oxford, UK*, 2002.

[3] R. Baldam, R. Vale, and H. Rozenfeld, *Gerenciamento de Processos de Negócio-BPM: Uma referência para implantação prática*. Campus, 2014.

[4] J. H. Holland, *Complexity: A very short introduction*. Oxford University Press, 2014.

[5] T. L. Norman, *Risk analysis and security countermeasure selection*. CRC Press, 2016.

[6] B. Fischhoff, S. Lichtenstein, P. Slovic, S. Derby, and R. Keeney, *Acceptable risk*. Cambridge University Press, 1984.

[7] A. A. Abramczuk, *The practice of decision making (in portuguese)*. Atlas, 2000.

[8] R. D. Stacey, *Strategic Management and Organisational Dynamics: The challenge of complexity to ways of thinking about organisations*. Prentice Hall, 6 ed., 2011.

[9] S. Lloyd, "Measures of complexity: a nonexhaustive list," *IEEE Control Systems Magazine*, vol. 21, no. 4, pp. 7–8, 2001.

[10] J. Paiva, V. Gomes, S. Oliveira, E. Domingues, A. Alves, G. Cruz Junior, and W. Calixto, "Calculation of system complexity based on the connections: methodology and applications," in *16 IEEE International Conference on Environment and Electrical Engineering*, 2016.

[11] L. Chwif and A. C. Medina, *Modeling and simulation of discrete events (in portuguese)*. Campus-Elsevier, 4 ed., 2014.