

Fuzzy Multiple Criteria Workflow Robustness and Resiliency Modeling with Petri Nets

Madjid Tavana, La Salle University, USA

Timothy E. Busch, Air Force Research Laboratory, USA

Eleanor L. Davis, University of Texas at Austin, USA

ABSTRACT

The increasing complexity and tight coupling between people and computer systems in military operations has led to improved efficiency, as well as greater vulnerability due to system failure. Careful management of workflow systems can minimize operational vulnerability in command and control. Tavana et al. (2011) developed a workflow management framework capable of both modeling structure and providing a wide range of quantitative analysis with high-level Petri nets (PNs). The framework is based on a sustainability index that captures the concepts of self-protecting and self-healing systems. This index uses crisp numerical values to measure the robustness and resiliency of the system. However, the observed values of data in real-world military operations are often imprecise or vague. These inexact data can be represented by fuzzy numbers to reflect the decision makers' intuition and subjective judgments. In this paper, the authors extend this model to a fuzzy framework by proposing a new fuzzy workflow modeling system with PNs. The new model plots the fuzzy robustness and resiliency measures in a Cartesian coordinate system and derives an overall fuzzy sustainability index for the system based on the theory of displaced ideals. The proposed model also considers multiple criteria to produce this fuzzy index.

Keywords: Fuzzy Sets Theory, High-Level Petri Net, Resiliency, Robustness, Workflow Management System

INTRODUCTION

Robustness, the ability of a system to avoid failure, and resiliency, the ability to recover from failure once it occurs, are essential elements of good workflow management systems. Nevertheless, there has been very little discussion of these properties in the workflow management

literature. The necessity for readiness and the ability to cope with the possibility and reality of failure in complex systems makes this an important area for future workflow management studies, especially in highly critical areas such as military operations.

Tavana et al. (2011) proposed measures of *robustness* and *resiliency* for Petri nets (PNs) expanded with alternate paths and repair times. These measures require the use of crisp values

DOI: 10.4018/ijkbo.2011100105

for times and probabilities. However, in real-world situations, workflow management systems are characterized by a high level of imprecision, vague parameters, and ill-defined relationships. Imprecision reduction must occur to find the measures of robustness and resiliency for PNs. Few approaches are suggested in the literature to adequately represent imprecision, and formally reduce it to precise values. Fuzzy set theory has considerable potential for addressing the imprecision in workflow management systems.

In this paper, we propose the use of fuzzy sets in the calculation of resiliency and robustness for workflow management systems. These fuzzy values will account for the variability of data while still enabling the calculation of robustness and resiliency. Fuzzy logic and approximate reasoning enable computation in the face of uncertainty, generating approximate results (Nedjah & De Macedo Mourelle, 2005). Fuzzy triangular numbers are used for each property. Tavana et al. (2011) based their calculations of robustness and resiliency on three factors: arc breakage probability, transition completion time, and transition repair time. We incorporate additional cost-based properties in the model proposed in this study. We also make use of available protection, as proposed by Zammori et al. (2009). Probabilities of failure are approximated by verbal expressions. The calculations of robustness and resiliency (κ and γ) are modified to accommodate these new and modified values. We extend the pragmatism and efficacy of the model proposed by Tavana et al. (2011). We will incorporate additional factors contributing to the resiliency and robustness of systems and allow for more flexible input of uncertain, vague, and ambiguous data through the use of fuzzy numbers.

This paper is organized as follows. We review the work on PN-based workflow management systems in the next section followed by a review of the literature on imperfect data representation methods. We next discuss the workflow management extensions considered in this study (i.e., probability of occurrence, time, cost, available protection). We next introduce our measures of robustness and resiliency for

the PNs in fuzzy environments and provide a graphical representation of the model. We then present a numerical example and an application of the proposed model to an air tasking order generation process of the U.S. Air Force. Finally, we present our conclusions and future research directions.

PN-BASED WORKFLOW MANAGEMENT

A workflow management system is a set of activities involving the coordinated implementation of various tasks performed by different processing entities (Casati et al., 1995; van der Aalst & van Hee, 2002). Different techniques may be used for workflow modeling depending on the goals and objectives. While workflow management systems are popular, with widespread applications, they still suffer from lack of standards and an agreed-upon modeling method (Salimifard & Wright, 2001). These systems are complex artifacts; they are difficult and expensive to build and validate, especially when the components of the system exhibit complicated properties such as sequential synchronization, merging, or prioritization (Balduzzi et al., 2000; Mehrez et al., 1995).

PNs have gained popularity recently as a tool for workflow modeling (Hsieh & Chiang, 2010; van der Aalst et al., 1994, 1998, 2000). However, the majority of existing PN models put their focus on the process aspect and do not consider important characteristics of the workflow management systems (Julia et al., 2008). PNs are the only formal technique that provides a wide range of qualitative and quantitative analysis in addition to modeling structure. They provide a graphical representation to enhance the understanding of the modeled system. PNs can also be used to formally analyze, verify, and validate the model (van der Aalst, 1997; Desel, 2000).

Each PN is defined by sets of transitions, places, and arcs. Transitions (rectangles) and places (circles) are two types of nodes, connected by arcs. Transitions represent some sort

of action or process, while places represent the input and output of these processes. Some PNs include concepts of time, color, or other attributes. For further details, see Tavana et al. (2011).

In this study, we use high-level timed PNs expanded in several dimensions. The ability of PNs to model the workflow primitives identified by the Workflow Management Coalition was shown by vanderAalst (1996). We expand those primitives to include alternate splits, where the dotted arcs serve as secondary paths to be taken if there is a fault in the primary route (e.g., a broken arc or a failed transition). The modeling of these “back-up plans” is useful in the study of the robustness and resiliency of the PNs (Figure 1).

IMPERFECT DATA REPRESENTATION

Several theories have been developed to deal with imperfect data. Imperfect data can be characterized as being uncertain, imprecise, or both. Other imperfect data such as vague or incomplete data can be described as a special form of imprecision and (or) uncertainty (Smets, 1997). Bayesian theory deals with both uncertainty and imprecision (Fienberg, 2006; Howson & Urbach, 1993; Jaynes, 2003). Rough sets theory is used to handle imprecision when uncertainty is involved but not quantified (Pawlak, 1991). The theory of evidence deals with data that contains both imprecision and uncertainty at the same time (Shafer, 1976; Dempster, 1967). The theory of possibility handles incomplete data, which is a combination of imprecise and uncertain data (Zadeh, 1978). The theory of fuzzy sets deals with vague data which is a particular case of both imprecise and uncertain data (Zadeh, 1965). Although these theories are used to handle only one type of imperfection, random sets and the conditional event algebra are proposed to cope with all types of imperfection (Goodman et al., 1997).

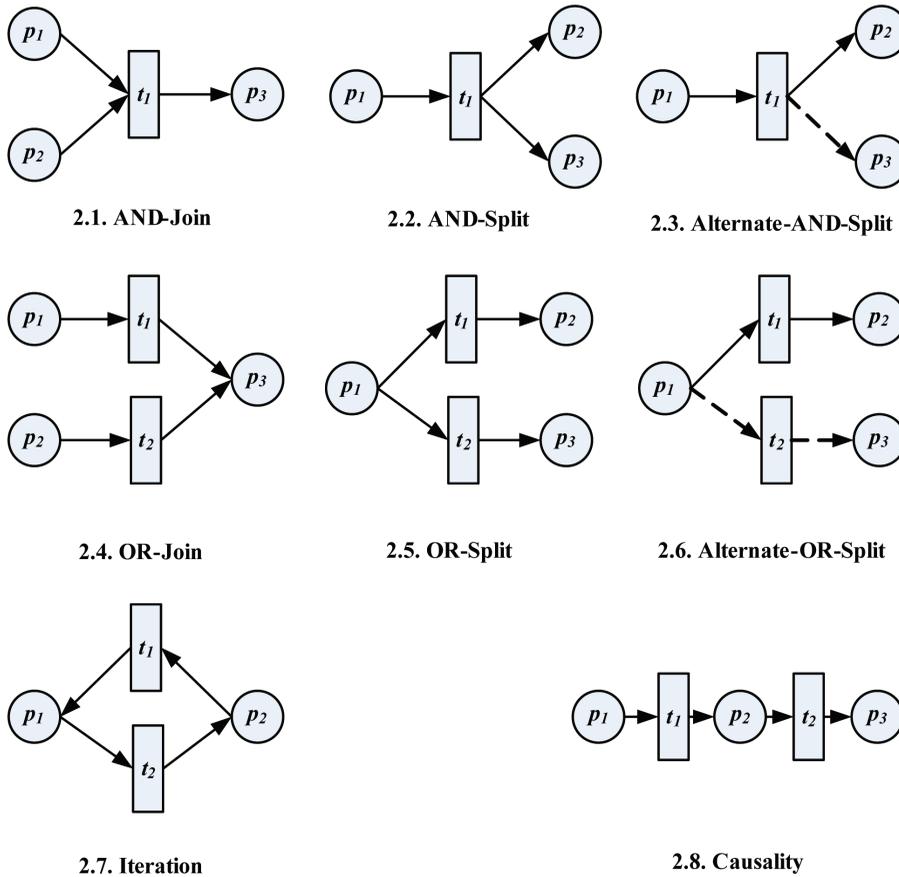
Uncertainty represents the state of knowledge about a piece of data, while imprecision

is the characteristic of a piece of data that cannot be expressed with a single value. When uncertainty is used to characterize chance, it is objective data; however, when uncertainty is used to characterize the state of knowledge, it is subjective data. Imprecision is a characteristic of objective data only. Vagueness is generally due to the limitation of the vocabulary and is often used to characterize subjective data. The phrase “Jim died young” is a definite piece of information but vague because we neither know how old Jim was when he died, or what the limits of “young” is. Does young mean 20 years old or 40 years old? This vague data is imprecise and uncertain and can be better represented with the theory of fuzzy sets.

In classical set theory, an element either belongs to a set or it does not. This strict division is not always practical. For example, if we consider the time spent in a dentist’s waiting room, we may define the crisp set of waiting times that are “long” to include all waits longer than 30 minutes. According to this definition, a wait of 30 minutes is considered “long,” while one of 29 minutes and 59 seconds is “not long.” In practice, both would seem to be about the same length. We would likely distinguish between degrees of length: the wait may be “somewhat long” or “extremely long”; there would be a gradual change from “not long” rather than a strict division between sets. In other words, some time values will be members of the set of long waits with more strength than others.

The membership function of a fuzzy set defines the mapping of inputs to the degree or strength of membership, ranging from 0 to 1. The shape of this membership function can vary, as any function whose image is between 0 and 1 is a possible membership function. The simplest of these functions are those represented by straight lines, such as triangular and trapezoidal member functions. In this study, we use triangular fuzzy numbers to represent and quantify the vagueness associated with the decision variables and the input and output data. Fuzzy sets have been used to account for the variability of data in workflow management systems (Lin et al., 2007; Tsai & Wang, 2008).

Figure 1. Workflow primitives



The choice of triangular fuzzy numbers in this study is made due to their simplicity and the ease of interpretation in modeling. Other types of fuzzy numbers, including trapezoidal fuzzy numbers, may increase the computational complexity without substantially affecting the accuracy of the results (Wang & Elhag, 2006; Yang & Hung, 2007).

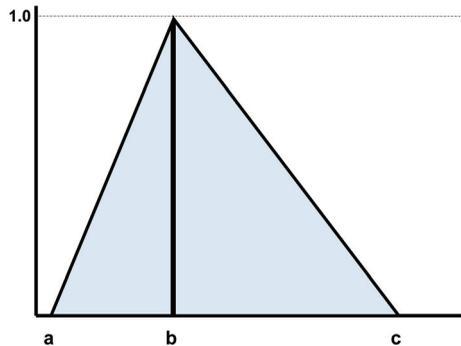
Fuzzy triangular numbers (FTNs) are comprised of an ordered triplet $\{a, b, c\}$ defining the lower value, core (i.e., most likely), and the upper value that the given property may

take on (Figure 2). FTNs will be used in this study because they are both simple to define and easy to manipulate while still providing the flexibility of data desired of fuzzy numbers (Zammori et al., 2009).

Fuzzy Arithmetic

In order to measure system properties, it is necessary to manipulate the FTNs and perform basic arithmetic operations on them. The relevant operations are defined as follows (Zammori et al., 2009):

Figure 2. Fuzzy triangular number $\{a,b,c\}$



Fuzzy Maximum $\left(\begin{matrix} \sim \\ max \end{matrix} \right)$

$$\begin{matrix} \sim \\ max \end{matrix} (FTN_1, FTN_2) \equiv \{ \max(a_1, a_2), \max(b_1, b_2), \max(c_1, c_2) \}$$

Fuzzy Division ($/$)

$$FTN_1 (/) FTN_2 = \left(\frac{a_1}{c_2}, \frac{b_1}{b_2}, \frac{a_1}{a_2} \right)$$

Fuzzy Subtraction ($-$)

$$FTN_1 (-) FTN_2 = (a_1 - c_2, b_1 - b_2, c_1 - a_2)$$

Fuzzy Addition ($+$)

$$FTN_1 (+) FTN_2 = (a_1 + a_2, b_1 + b_2, c_1 + c_2)$$

Fuzzy Multiplication (\times)

An α -cut, defined for some value α ($0 \leq \alpha \leq 1$), is an interval containing the members whose values are greater than α .

$$FTN_1 (\times) FTN_2 = (a_1 \times a_2, b_1 \times b_2, c_1 \times c_2)$$

It should be noted that while using these operations, $a \leq b \leq c$ at all times.

PN WORKFLOW EXTENSIONS

Probability of Occurrence Consideration

The existence of alternate paths leading to an output increases the robustness of the network. When part of the network fails, an alternate arc can be taken; when there are multiple paths that must all be taken for the system to function (i.e., concurrent transitions), the existence of back-ups can create different possible routes. We define a route as a possible set of paths that form a functioning system; a path is a set of arcs, transitions, and places leading from the input to the output.

Failures of proper behavior in transitions or places can be modeled by the incapacitation of arcs. An arc from place i to transition j is denoted l_{p_i, t_j} , while one in the opposite direction would be written as l_{t_j, p_i} . We do not distinguish between arcs from places to transitions and arcs from transitions to places; the notation P_{p_i, t_j} is considered to cover both cases. In general, the smaller the overall expected incapacitation probability, the more robust the network. Each arc is assigned a

breakage probability FTN P_{p_i, t_j} based on the likelihood of that arc breaking.

We represent the fuzzy set of arcs that will fail with S_F . An arc is only sure to break if it is a member of that set with degree 1; if it is a member with degree 0, it will never fail. The degree of membership corresponds directly to the probability that the arc will fail.

It is often difficult to assign a numerical value to something like the chance of failure. Instead, these probabilities can be represented by linguistic terms characterized by fuzzy numbers to reflect the decision makers' intuition and subjective judgments. Many people prefer to give verbal expressions of uncertainty but to receive numerical probability judgments (Tavana et al., 1997). While a number of studies have proposed various mappings of verbal phrases to numerical values, most have dealt with equally spaced ranges of probabilities. In this paper, we believe that the ability to distinguish between small probabilities of failure is more important than the maintenance of uniform ranges; the difference a 0.001 chance of failure and a 0.01 chance is much more likely to be useful than the difference between chances of 0.601 and 0.6001, for example. We therefore utilize the set of verbal descriptors proposed by Goodarzi et al. (2010). We expand the specific probability assigned to each term into an FTN, as shown in Figure 3, with the proposed probability functioning as the most-likely core value. Since it is difficult and in fact impractical to draw a strong division between the expressions, the probability ranges overlap. The specific verbal phrases used and the corresponding probability ranges can be altered at the discretion of the user.

Each arc will be assigned a verbal evaluation of the breakage probability. The corresponding FTN can be found in Table 1. In situations where sufficient data are available to provide more precise probability values or intervals, these values should be used instead of verbal estimates.

The incapacitation probability of a network is the probability that the network will be rendered nonfunctional. It can be determined by

examining the structure of the network and the breakage probabilities of its constituent arcs. If every arc must be complete for the network to function, then the incapacitation probability is calculated by subtracting the probability that none of the arcs are broken from 1 (i.e., certainty). The probability that none of the arcs are broken is the product of $1(-)P_{p_i, t_j}$ for all expected breakage probabilities P_{p_i, t_j} :

$$P_{\text{incapacitation}} = 1(-) \prod \left(1(-) P_{p_i, t_j} \right) \quad (1)$$

It should be noted that an integer can be treated as an FTN: $1 = \{1, 1, 1\}$.

The calculation of the incapacitation probability becomes more complex when there are alternate routes to the output and so the breakage of one arc does not necessarily result in the failure of the entire network. Instead, the network is incapacitated only when all of the possible routes are broken. In other words, at least one of the paths in all of the routes must be broken. To simplify this calculation, we recommend dividing the PN into zones surrounding each transition. Zones should contain either primary or back-up (alternate) arcs, but not both. A failure in any of the arcs in each zone will incapacitate at least one path. By considering the probability of failure or success for each individual zone and various combinations of zones, the incapacitation probability can be found (see the practical application that follows). It should be noted that through traditional fuzzy arithmetic, the sum of all possible cases may be more than 1. By using linear optimization and forcing the sum to equal 1, it is possible to calculate the fuzzy probability of failure (Buckley, 2006).

Transition Properties

Time Consideration

The amount of time it takes to complete a process is vital information when studying workflow systems. In a timed PN, each transi-

Figure 3. Fuzzy verbal expression membership functions (not to scale)

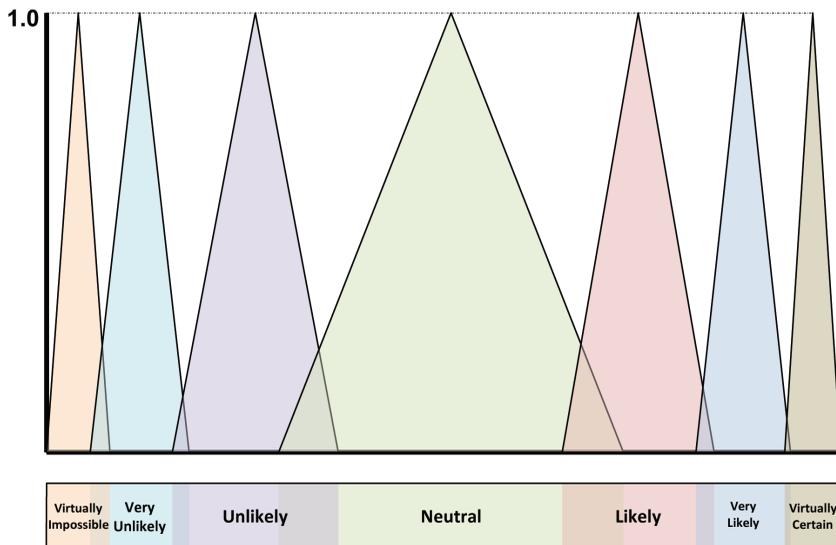


Table 1. Verbal expressions and corresponding probability estimates

Verbal expressions	Probability
Virtually Certain	{0.992025, 0.999, 1}
Very Likely	{0.93275, 0.99, .994775}
Likely	{0.68, 0.9, .947475}
Neutral	{.287525, 0.5, .71225}
Unlikely	{.04825, 0.1, .32}
Very Unlikely	{0.005225, 0.01, .062975}
Virtually Impossible	{0, 0.001, 0.00775}

tion takes a certain amount of time. However, the length of time required by a process is often not constant. The specific people or computer systems involved and even outside factors such as the weather can greatly influence the time involved. Because of this uncertainty, we represent the completion times with FTNs. The completion time for transition i is denoted T_{t_i} .

The system completion time is the time that it takes to progress from the initial input to the final output. Each route will have its own completion time Tr . When a route consists of a

single path, the route completion time is the sum of the completion times of the transitions along that path and when a route consists of multiple paths (i.e., concurrent transitions) the route completion time is calculated using the fuzzy maximum operation (the longest path dictates the time that it will take the system to run).

When the system has only one route, the system completion time is the route completion time. Otherwise, a weighted average of the resulting route completion times can be taken to find the overall system completion time. In

this calculation, it is assumed that some route to the output will be found. Network incapacitation is not considered, since that would imply an infinite amount of time until completion. The probability of taking each route is an FTN derived from the breakage probabilities; the core value of each FTN is multiplied by its route completion time. These values are then added together, giving the overall system completion time:

$$Ts = Tr_1 (\times) P_{core} (\text{route}_1) (+) \dots (+) Tr_n (\times) P_{core} (\text{route}_n) \quad (2)$$

When a transition fails, a certain amount of time is required to repair it. While this repair time will most likely vary depending on the specific type of failure, it is possible to consider a range of probable times. Each transition is therefore assigned a repair time R_{t_i} that is also an FTN. Route and system repair times are defined similarly to the route and system completion times. The overall expected system repair time is:

$$Rs = Rp_1 (\times) P_{core} (\text{route}_1) (+) \dots (+) Rp_n (\times) P_{core} (\text{route}_n) \quad (3)$$

Cost Consideration

There are costs associated with each step of a process. These costs may characterize man-hours, materials, equipment, and resources; they may also consist of non-financial costs such as opportunity cost, tax credit, and so on. We assign a completion cost (C_{t_i}) and a repair cost (D_{t_i}) to each transition. Route costs are the sum of the costs along that route. If there are no alternate routes, then system completion and repair costs are the sum of all the transition completion and repair costs. If alternate routes do exist, the system costs are the sum of the route costs weighted according to the core probability of taking them, assuming that the system works.

Available Protection

It is often possible to reduce the completion time of a transition by increasing its cost. This may involve assigning additional man-hours to a project, outsourcing a task, or other methods. Each transition is therefore assigned a maximum possible reduction in time (ΔT_{t_i}) and the corresponding increase in cost (ΔC_{t_i}) as proposed by Zammori et al. (2009). For each route, we can calculate the Available Protection (AP). The AP gives the potential reduction in route time per unit cost:

$$AP_{r_i} = \frac{\sum_i \Delta T_{t_i}}{\sum_i \Delta C_{t_i}} \quad (4)$$

where the numerator and denominator give the route values ΔT_{r_i} and ΔC_{r_i} , respectively.

The system APs are calculated similarly to the system costs. It measures the total flexibility in time rather than the flexibility of the system completion time. The APs are not derived from the route AP, but from the route changes in time and cost.

$$AP_s = \frac{\sum_i \left[P_{core} (\text{route}_i) (\times) \Delta T_{r_i} \right]}{\sum_i \left[P_{core} (\text{route}_i) (\times) \Delta C_{r_i} \right]} = \frac{\Delta T}{\Delta C} \quad (5)$$

ROBUSTNESS AND RESILIENCY MEASURES

Robustness Measure

Robustness is the ability of a system to avoid failure. The less vulnerable the system time and cost are to problems, the more robust the network will be. Therefore, a comparison of the optimistic and pessimistic scenarios is relevant. We make an α - cut with $\alpha = 0.1$ in our FTNs Ts and Cs . These cuts define the interval $[a^*, c^*]$

in which membership is greater than 0.1. The lowest values of time or cost occur between a and a^* , while the highest values occur between c^* and c . We define a best-case FTN with the ordered triplet $\left\{ a, \frac{a + a^*}{2}, a^* \right\}$ and

a worst-case FTN with the ordered triplet $\left\{ c, \frac{c + c^*}{2}, c^* \right\}$. The ratios of the best-case

system time to the worst-case system time and the best-case system cost to the worst-case system cost are both incorporated into our robustness measure. The smaller the value of these ratios, the greater the potential ramifications of disruption to the system, and so the less robust the system is.

The APs are also incorporated. Since the decrease in time ΔT will always be smaller than the system time Ts but the increase in cost ΔC could be either smaller or greater than the system cost Cs , a straightforward ratio is not practical. Instead, we multiply the ratio of the decrease in time to the system time by the ratio of the system cost to the total cost $\frac{\Delta T}{Ts}(\times) \frac{Cs}{\Delta C(+)Cs}$. Since the traditional FTN

division would treat the values of Cs in the numerator and denominator as independent variables, an alternate method must be used.

The minimum of $\frac{Cs}{\Delta C(+)Cs}$ will occur when

Cs is at its minimum and ΔC is at its maximum, while the maximum will occur when ΔC is minimized and Cs is maximized. These two values, combined with the traditionally calculated middle value, will provide the appropriate FTN as a result. The quantity $\frac{\Delta T}{Ts}(\times) \frac{Cs}{\Delta C(+)Cs}$ is bounded above by 1 and

below by 0; it will be the greatest when the potential to decrease the time is high but the corresponding increase of cost is low.

We define κ as the combination of time, cost, and probability used to measure robustness:

$$\kappa = \left(\frac{T_{s_{best}}(+) C_{s_{best}}(+) \frac{\Delta T}{Ts}(\times)}{T_{s_{worst}} C_{s_{worst}} \frac{Cs}{\Delta C(+)Cs}(-) P_{incapacitation}(-) 1} \right) / 2 \tag{6}$$

The final subtraction term is included simply so that κ will have a symmetrical range. The denominator is added to provide a range from -1 to 1. A system is least robust when $\kappa = -1$ (the worst-case time is infinitely greater than the best-case time, the worst-case cost is infinitely greater than the best-case cost, time cannot be decreased without infinite cost, and failure is certain). It is most robust when $\kappa = 1$ (the best-case time and worst-case time are equal, the best-case cost and worst-case cost are equal, the process can take zero time with no added cost, and failure is impossible). Practically, neither of these extreme scenarios will actually occur, so the measured robustness will lie in the open interval from -1 to 1.

Resiliency Measure

Resiliency is the ability of the system to recover from failure. We compare the system completion time Ts and the system completion cost Cs with their corresponding repair values to measure resiliency (γ). The system repair time and cost assume failure in each transition, regardless of the probability of failure associated with it. This is reasonable, since resiliency is concerned with the ability to recover after disaster has already occurred, no matter how likely that disaster was. Resiliency is defined as:

$$\gamma = \left(\frac{T_s(-) R_s(+) T_s(-) R_s(+)}{T_s(+) R_s(+) T_s(+) R_s(+)} \right) / 2 \tag{7}$$

Again, the numerators and denominators cannot be treated traditionally as separate FTNs since this would treat the repeated values as independent variables. Instead, the minimum and maximum values for each division should be used for the minimum and maximum results

of the FTN. Similar to our robustness measure κ , γ ranges from -1 (the repair time and cost are infinitely greater than the system time and cost) to 1 (the system time and cost are infinitely greater than the repair time and cost).

A Graphical Perspective

When both robustness and resiliency values have been calculated, the intervals κ and γ can be plotted on the plane with range and domain $[-1,1]$. By choosing the threshold values $\hat{\kappa}$ and $\hat{\gamma}$, the plane can be divided into four quadrants. These threshold values are left to the discretion of the user, as the classification of some results as “resilient” or “robust” is situation-dependent.

After the thresholds have been chosen, the plane will be divided into four quadrants, identified as the Possession Quadrant, the Preservation Quadrant, the Restoration Quadrant, and the Devastation Quadrant (Figure 4).

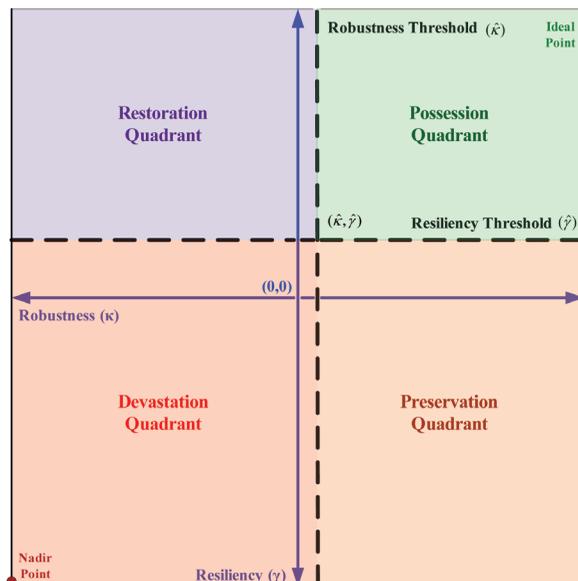
- **Possession Quadrant:** Networks in this quadrant are both robust and resilient. They are unlikely to encounter obstacles that will

disable the system; if they do, the system will recover without significant difficulty.

- **Preservation Quadrant:** In this quadrant, networks are robust but not resilient. These networks are unlikely to fail. However, if an unforeseen disaster occurs, the road to recovery will be long and hard.
- **Restoration Quadrant:** Networks in this quadrant are resilient but not robust. They may fail again and again, but repairs are quick. Before long, they will be back on their feet.
- **Devastation Quadrant:** These networks are neither robust nor resilient. They are very susceptible to failure, and once they have failed, recovery is difficult. Significant changes are needed to improve these systems.

Since both γ and κ are intervals, it is possible that one or both of the threshold values will run through them, causing the network to reside in multiple quadrants. We define the sub-ideal point and the sub-nadir for each robustness/resiliency pair. The sub-ideal point is the ordered pair made up of the highest pos-

Figure 4. Robustness-resiliency plane



sible values for robustness and resiliency (i.e., the upper right-hand corner of the plane), while the sub-nadir is the combination of the lowest possible values for robustness and resiliency, located in the lower left-hand corner of the plane. The overall ideal point for networks is (1,1), where both robustness and resiliency are at their maximum; while the overall nadir is (-1,-1), at which both are at their minimum values. The displacement of a given point from the ideal point is given by the Euclidean distance formula, $Distance = \sqrt{(\gamma - \gamma_0)^2 + (\kappa - \kappa_0)^2}$. We define $(\kappa_0, \gamma_0) = (1, 1)$. The overall sustainability index (ω) is the ratio of the distance between the core robustness and the resiliency values and the ideal point to the maximum distance possible (i.e., the distance between the ideal and the nadir points).

$$\omega = \frac{\sqrt{(\gamma_{core} - 1)^2 + (\kappa_{core} - 1)^2}}{2\sqrt{2}} \quad (8)$$

The lower the sustainability index, the closer the system resiliency and robustness are to the ideal. In addition to the intersection of the core values, a best-case and worst-case sustainability index can be determined using the sub-ideal point and the sub-nadir, respectively.

PRACTICAL APPLICATION

Let us consider the PN presented in Figure 5. The execution of this system is dependent on two parallel sequences. The arc breakage probabilities are given in Table 2, while the properties associated with each transition are shown in Table 3.

Since there are no alternate transitions, if any arc is broken, the network will fail. The incapacitation probability can be calculated as:

$$P_{incapacitation} = 1(-) \prod (1(-) P_{p_i, t_j}) = 1(-) \{0.074, 0.353, 0.605\} = \{0.395, 0.647, 0.926\}$$

The system completion and repair times are found by taking the fuzzy maximum of the values for the parallel tasks:

$$Ts = \widetilde{max}(T_{t_1}(+)T_{t_2}, T_{t_3}(+)T_{t_4}) = \{21, 25, 37\}$$

$$Rs = \widetilde{max}(R_{t_1}(+)R_{t_2}, R_{t_3}(+)R_{t_4}) = \{27, 35, 45\}$$

The costs for all transitions are added to calculate the system completion and repair costs:

$$Cs = \{200, 330, 385\}(+) \{260, 360, 450\}(+) \{100, 125, 165\}(+) \{100, 125, 265\} = \{910, 1190, 1400\}$$

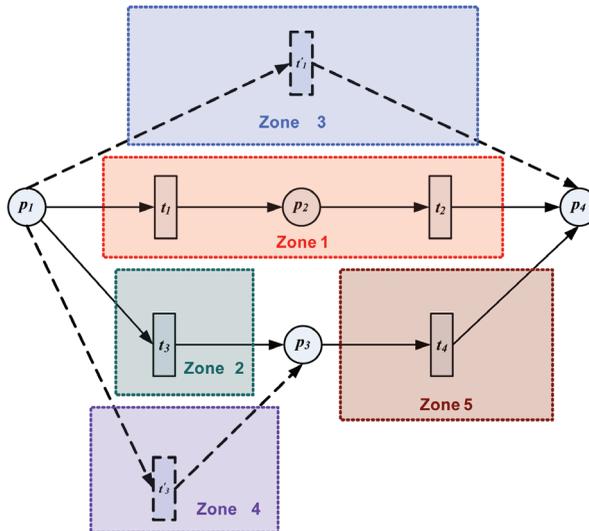
$$Rs = \{975, 1600, 2700\}(+) \{210, 225, 240\}(+) \{50, 55, 70\}(+) \{435, 450, 475\} = \{1670, 2330, 3485\}$$

The available protection is calculated similarly, with fuzzy division performed in the last step:

$$APs = \frac{\Delta T}{\Delta C} = \frac{\{1.5, 1.9, 2.4\}(+) \{2, 3, 3.25\}(+) \{1, 1.5, 1.75\}(+) \{.85, 1, 1.4\}}{\{289, 330, 370\}(+) \{90, 100, 125\}(+) \{80, 120, 130\}(+) \{60, 95, 110\}} = \{5.35, 7.4, 8.8\} \{0.007, 0.011, 0.0170\}$$

In order to compute the network robustness, best- and worst-case values for Ts and Cs must be determined. Making an α - cut in each FTN with $\alpha = .1$ and assuming that the core value is at the center of the interval, we find:

Figure 5. Simple workflow example (with alternate transitions)



$$T_{s_{best}} = \{21, 21.2, 21.4\},$$

$$T_{s_{worst}} = \{35.8, 36.4, 37\},$$

$$R_{s_{best}} = \{910, 924, 938\}, \text{ and}$$

$$R_{s_{worst}} = \{1379, 1389.5, 1400\}.$$

Next, we find the lowest value for resiliency by combining the minimum values of T_s and C_s with the maximum values of R_s and D_s . The maximum resiliency value is achieved with the greatest values of T_s and C_s and the minimum values of R_s and D_s . The core resiliency value is derived from the core time and cost values:

The robustness measure can then be calculated as:

$$\kappa = \left(\begin{array}{l} \frac{\{21, 21.2, 21.4\}}{\{35.8, 36.4, 37\}} (+) \\ \frac{\{910, 924, 938\}}{\{1379, 1389, 1400\}} (+) \\ \frac{\{5.35, 7.4, 8.8\}}{\{21, 25, 37\}} (\times) \\ \frac{\{0.553, 0.649, 0.730\}}{-\{0.395, 0.647, 0.926\}} (-) 1 \end{array} \right)^2$$

$$\gamma = \left(\begin{array}{l} \frac{21(-)45}{21(+)45}, \\ \frac{25(-)35}{25(+)35}, \\ \frac{37(-)27}{37(+)27} \end{array} \right)^2 + \left(\begin{array}{l} \frac{910(-)3485}{910(+)3485}, \\ \frac{1190(-)2330}{1190(+)2330}, \\ \frac{1400(-)1670}{1400(+)1670} \end{array} \right)^2$$

$$= \{-0.475, -0.245, 0.034\}$$

Table 2. Breakage probabilities for the example

Arc	Verbal Breakage Probability
l_{p_1, t_1}	Unlikely
l_{p_1, t_3}	Very Unlikely
l_{t_1, p_2}	Very Unlikely
l_{t_3, p_3}	Virtually Impossible
l_{p_2, t_2}	Virtually Impossible
l_{p_3, t_4}	Unlikely
l_{t_2, p_4}	Neutral
l_{t_4, p_4}	Very Unlikely
l_{p_1, t'_1}	Very Unlikely
l_{p_1, t'_3}	Unlikely
$l_{t'_1, p_4}$	Virtually Impossible
$l_{t'_3, p_3}$	Virtually Impossible

The overall sustainability index can then be calculated using the core values for robustness and resiliency:

$$\omega = \frac{\sqrt{(-0.245 - 1)^2 + (-0.104 - 1)^2}}{2\sqrt{2}} = 0.588$$

Figure 4 shows the network expanded with alternate paths. Transition t'_1 is an alternative to transitions t_1 and t_2 . Transition t'_3 is a backup for transition t_3 ; there is no alternative to transition t_4 . Five zones are defined which may either succeed or fail, giving a total of 32 possible states for the network (Table 4). Since transition 4 has no backup, if zone 5 fails, the

network fails regardless of the state of the other zones. Nine of the 32 possible states (identified as “Yes” in the fourth column of Table 4) allow the network to function.

In order to find the probability of failure, the relevant probabilities cannot simply be added, since this would generate results greater than 1. Instead, linear optimization is used. Since there are 32 possible network states, each state occurs with a probability p_i , $1 \leq i \leq 32$. We define a_i as the lowest value given by the probability FTN for state i , and b_i as the highest value. The probabilities are then constrained as follows:

$$a_i \leq p_i \leq b_i$$

$$p_1 + \dots + p_{32} = 1$$

Table 3. Transition properties for the example

Transition	Completion Time T_t	Repair Time R_t	Completion Cost C_t	Repair Cost D_t	Change in Completion Time ΔT_t	Change in Completion Cost ΔC_t
t_1	{12, 15, 22}	{23, 28, 35}	{200, 330, 385}	{975, 1600, 2700}	{1.5, 1.9, 2.4}	{289, 330, 370}
t_2	{9, 10, 13}	{4, 7, 10}	{260, 360, 450}	{210, 225, 240}	{2, 3, 2 3.25}	{90, 100, 125}
t_3	{4, 6, 9}	{6, 8, 8.5}	{100, 125, 165}	{50, 55, 70}	{1, 1.5, 1.75}	{80, 120, 130}
t_4	{3.5, 4, 4.75}	{1.5, 3, 4.25}	{350, 375, 400}	{435, 450, 475}	{.85, 1, 1.4}	{60, 95, 110}
t'_1	{6.5, 8, 9}	{.5, 2, 4}	{500, 850, 925}	{30, 40, 45}	{0, 0, 0}	---
t'_3	{11, 11.5, 13}	{6, 6.5, 7}	{85, 90, 100}	{40, 50, 60}	{6, 10, 14}	{10, 15, 18}

In order to generate an FTN for the probability of success, we wish to find the maximum and minimum values for the sum of the successful states:

$$(p_1 + p_3 + p_5 + p_9 + p_{11} + p_{13} + p_{17} + p_{19} + p_{21})$$

Because this is a linear equation, its optimization can be calculated quickly by Maple or Excel Solver. The maximum and minimum values, combined with the sum of the corresponding core values (which do not require optimization), define an FTN for the probability of success. When subtracted from 1, this gives a system incapacitation probability of {0.017, 0.076, 0.525}.

In order to calculate the system times and costs along with the Available Protection, each possible route value must be weighted according to the probability of taking that route. These probabilities are different from those given in Table 4, since it is assumed that some working route will be found. Additionally, only the core values of these probabilities are used in the calculations of these quantities.

There are four possible routes from input to output. Route 1 goes through zones 1, 2, and 5; route 2 goes through zones 1, 4, and 5; route 3 goes through zones 2, 3, and 5; and route 4 goes through zones 3, 4, and 5. The probability of taking each route is based on the probability that the alternate transitions are not needed. The probability of failure of alternate transitions or transitions with no back-up is not taken into consideration.

The probability of taking route 1 is the probability that both zones 1 and 2 work:

$$P(\text{route}_1) = \{0.079, 0.357, 0.608\}$$

The probability of taking route 2 is the probability that zone 1 works and zone 2 does not work:

$$P(\text{route}_2) = \{0.007, 0.044, 0.233\}$$

The probability of taking route 3 is the probability that zone 1 does not work and zone 2 does work:

$$P(\text{route}_3) = \{0.228, 0.534, 0.829\}$$

The probability of taking route 4 is the probability that neither zone 1 nor zone 2 works:

$$P(\text{route}_4) = \{0.019, 0.065, 0.318\}$$

The route completion times are found in the same way that the system completion time was found in the unmodified system. The new system completion time is the weighted average of these values:

$$\begin{aligned} Ts &= Tp_1(\times)P(\text{route}_1)(+) \\ Tp_2(\times)P(\text{route}_2)(+) &+ Tp_3(\times) \\ P(\text{route}_3)(+) &+ Tp_4(\times)P(\text{route}_4) \\ &= \{21, 25, 27\}(\times)0.357(+) \{21, 25, 37\}(\times) \\ &0.044(+) \{7.5, 10, 13.75\}(\times)0.534(+) \\ &\{13.5, 15.5, 17.75\}(\times)0.065 \\ &= \{13.305, 16.373, 23.333\} \end{aligned}$$

The system repair time, system completion cost, and system repair cost are calculated similarly, as are ΔT and ΔC (Table 5).

We use these values to find the robustness and resiliency measures and the overall sustainability index:

$$\gamma = \{-0.310, -0.052, 0.201\}$$

$$\kappa = \{-0.074, 0.249, 0.383\}$$

$$\omega = 0.457$$

Let us further assume that the threshold values of $\hat{\sigma} = 0$ and $\hat{\sigma} = 0$ are chosen. While each threshold value may be any number between -1 and 1, these values are chosen for the sake of simplicity. As shown in Figure 6, the initial system was likely to be neither robust nor resilient. The robustness and resiliency measures both increased with the addition of backup

transitions, moving the core values of the system used for the sustainability index from the Devastation Quadrant to the Possession Quadrant.

CONCLUSION AND FUTURE RESEARCH DIRECTIONS

Organizations use workflow management systems to coordinate their business environment, facilitate instant access to accurate and up-to-date data, schedule and synchronize changes and test the effectiveness of business processes. Despite the importance of developing and maintaining self-protecting and self-healing business processes, the study of resiliency and robustness has received little attention in the workflow management systems literature. Tavana et al. (2011) proposed measures of *robustness* and *resiliency* for PNs expanded with alternate paths and repair times. Their measures required the use of crisp values for times and probabilities. However, in real-world problems, workflow management systems are characterized by a high level of imprecision, vague parameters, and ill-defined relationships. Imprecision reduction must occur to find the measures of robustness and resiliency for PNs. Few approaches are suggested in the literature to adequately represent imprecision, and formally reduce it to precise values. Fuzzy set theory has considerable potential for addressing the imprecision in workflow management systems.

This study increases the practical application of the workflow management model of Tavana et al. (2011) by incorporating additional factors contributing to the resiliency and robustness of systems and allowing for more flexible input of uncertain data through the use of fuzzy numbers. Fuzzy measures for robustness and resiliency are also generated, showing the range that each value may take on, as well as the most likely value. This expanded model allows broader and more flexible inputs while still generating well-defined and useful results. The framework proposed in this paper could be expanded to:

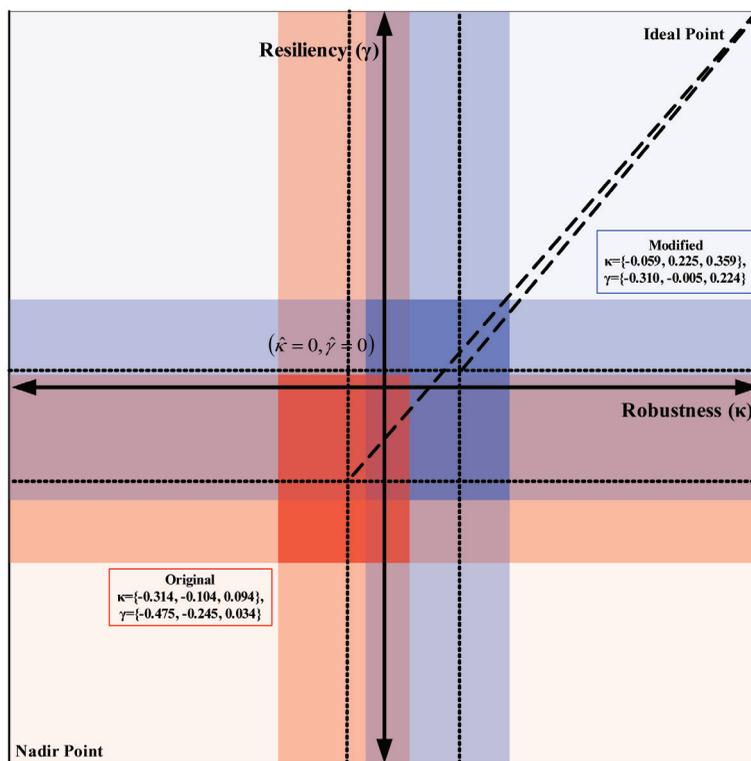
Table 4. Possible network conditions and probabilities of occurrence

Path	Functioning Zones	Failing Zones	Network Functions	Probability
1	1, 2, 3, 4, 5	-----	Yes	{0.0638, 0.346, 0.598}
2	1, 2, 3, 4	5	No	{3.588E-4, 0.004, 0.042}
3	1, 2, 3, 5	4	Yes	{3.588E-4, 0.004, 0.042}
4	1, 2, 3	4, 5	No	{2.016E-06, 4.267E-05, 0.003}
5	1, 2, 4, 5	3	Yes	{3.588E-4, 0.004, 0.042}
6	1, 2, 4	3, 5	No	{2.016E-06, 4.267E-05, 0.003}
7	1, 2, 5	3, 4	No	{2.016E-06, 4.267E-05, 0.003}
8	1, 2	3, 4, 5	No	{1.133E-08, 4.742E-07, 2.106E-4}
9	1, 3, 4, 5	2	Yes	{0.005, 0.042, 0.229}
10	1, 3, 4	2, 5	No	{2.997E-05, 4.698E-4, 0.016}
11	1, 3, 5	2, 4	Yes	{2.997E-05, 4.698E-4, 0.016}
12	1, 3	2, 4, 5	No	{1.684E-07, 5.221E-06, 0.001}
13	1, 4, 5	2, 3	Yes	{2.997E-05, 4.698E-4, 0.016}
14	1, 4	2, 3, 5	No	{1.684E-07, 5.221E-06, 0.001}
15	1,5	2, 3, 4	No	{1.684E-07, 5.221E-06, 0.001}
16	1	2, 3, 4, 5	No	{9.465E-10, 5.801E-08, 8.072E-05}
17	2, 3, 4, 5	1	Yes	{0.183, 0.516, 0.816}
18	2, 3, 4	1, 5	No	{0.001, 0.006, 0.058}
19	2, 3, 5	1, 4	Yes	{0.001, 0.006, 0.058}
20	2, 3	1, 4, 5	No	{5.790E-06, 6.376E-05, 0.004}
21	2, 4, 5	1, 3	Yes	{1.684E-07, 5.221E-06, 0.001}
22	2, 4	1, 3, 5	No	{5.790E-06, 6.376E-05, 0.004}
23	2, 5	1, 3, 4	No	{5.790E-06, 6.376E-05, 0.004}
24	2	1, 3, 4, 5	No	{3.254E-08, 7.085E-07, 2.872E-4}
25	3, 4, 5	1, 2	No	{0.015, 0.0632, 0.313}
26	3, 4	1, 2, 5	No	{8.606E-05, 7.019E-4, 0.022}
27	3, 5	1, 2, 4	No	{8.606E-05, 7.019E-4, 0.022}
28	3	1, 2, 4, 5	No	{4.836E-07, 7.800E-06, 0.002}
29	4, 5	1, 2, 3	No	{8.606E-05, 7.019E-4, 0.022}
30	4	1, 2, 3, 5	No	{4.836E-07, 7.800E-06, 0.002}
31	5	1, 2, 3, 4	No	{4.836E-07, 7.800E-06, 0.002}
32	----	1, 2, 3, 4, 5	No	{2.718E-09, 8.66728E-08, 1.101E-4}

Table 5. Example system results

System	Original	Modified
Probability of Failure $P_{\text{Incapacitation}}$	{0.395, 0.647, 0.926}	{0.017, 0.076, 0.525}
Completion Time T_s	{21.000, 25.000, 37.000}	{13.305, 16.373, 23.333}
Repair Time R_s	{27.000, 35.000, 45.000}	{15.319, 20.525, 25.583}
Completion Cost C_s	{910.000, 1190.000, 1400.000}	{932.327, 1282.033, 1446.830}
Repair Cost D_s	{1670.000, 2330.000, 3485.000}	{977.007, 1260.151, 1749.660}
Available Protection AP_s	{0.007, 0.011, 0.017}	{0.009, 0.015, 0.024}
Robustness κ	{-0.314, -0.104, 0.094}	{-0.074, 0.249, 0.383}
Resiliency γ	{-0.475, -0.245, 0.034}	{-0.310, 0.052, 0.201}
Overall Sustainability Index ω	.588	.457

Figure 6. Original and modified case study robustness-resiliency plane



- Include a third dimension in the resiliency, robustness and sustainability index to track changes to a system over time.
- Consider the interdependency of resiliency and robustness. These two features may not necessarily function independently, as changes to one may affect the other.
- Encompass a computer implementation of the proposed framework. An automated system will provide the capability for continuous monitoring of the resiliency and robustness in large systems.

In this study, we have built upon the ground-work for the consideration of resiliency and robustness in workflow management systems. We hope that the concepts introduced and expanded upon here will provide inspiration for future research.

ACKNOWLEDGMENT

This research was supported by the U.S. Air Force Research Laboratory grant number FA8750-09-2-0155.

REFERENCES

Balduzzi, F., Giua, A., & Menga, G. (2000). First-order hybrid Petri nets: a model for optimization and control. *IEEE Transactions on Robotics and Automation*, 16(4), 382–399. doi:10.1109/70.864231

Buckley, J. J. (2006). *Fuzzy Probability and Statistics*. Berlin, Germany: Springer-Verlag.

Casati, F., Ceri, S., Pernici, B., & Pozzi, G. (1995). Conceptual modelling of workflows. In *Proceedings of the 14th International Object-Oriented and Entity-Relationship Modelling Conference* (LNCS 1021, pp. 341-354).

Dempster, A. (1967). Upper and lower probabilities induced by multivalued mapping. *Annals of Mathematical Statistics*, 38, 325–339. doi:10.1214/aoms/1177698950

Desel, J. (2000). Validation of process models by construction of process nets. In W. M. P. van der Aalst, J. Desel, & A. Oberweis (Eds.), *Business process management: models, techniques and empirical studies* (LNCS 1806, pp. 110-128).

Fienberg, S. E. (2006). When Did Bayesian Inference Become “Bayesian”? *Bayesian Analysis*, 1, 1–40. doi:10.1214/06-BA101

Goodman, I., Mahler, R. P. S., & Nguyen, H. (1997). *Mathematics of data fusion*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Howson, C., & Urbach, P. (1993). *Scientific reasoning: The Bayesian approach*. Chicago, IL: Open Court.

Hsieh, F.-S., & Chiang, C. Y. (2010). Collaborative composition of processes in holonic manufacturing systems. *Computers in Industry*.

Janssens, G. K., Verelst, J., & Weyn, B. (2000). Techniques for Modelling Workflows and their Support of Reuse Definition of Basic Workflow Concepts Workflow Concepts Translated into Petri Nets. In W. M. P. van der Aalst, J. Desel, & A. Oberweis (Eds.), *Business process management: Models, techniques and empirical studies* (LNCS 1806, pp. 1-15).

Jaynes, E. T. (2003). Probability theory: The logic of science. In Bretthorst, G. L. (Ed.), *The art of scientific computing*. New York, NY: Cambridge University Press.

Julia, S., de Oliveira, F. F., & Valette, R. (2008). Real time scheduling of workflow management systems based on a p-time Petri net model with hybrid resources. *Simulation Modelling Practice and Theory*, 16(4), 462–482. doi:10.1016/j.simpat.2008.01.006

Lin, H.-Y., Hsu, P.-Y., & Sheen, G.-J. (2007). A fuzzy-based decision-making procedure for data warehouse system selection. *Expert Systems with Applications*, 32(3), 939–953. doi:10.1016/j.eswa.2006.01.031

Mehrez, A., Muzumdar, M., Acar, W., & Weinroth, G. (1995). A Petri-net model view of decision making: an operational analysis. *Omega*, 23(1), 63–78. doi:10.1016/0305-0483(94)00049-G

Nedjah, N., & De Macedo Mourelle, L. (2005). Introducing you to fuzziness. *Studies in Fuzziness and Soft Computing*, 181, 3–21.

Pawlak, Z. (1991). *Rough sets - Theoretical aspects of reasoning about data*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

- Salimifard, K., & Wright, M. (2001). Petri net-based modelling of workflow systems: An overview. *European Journal of Operational Research*, 134(3), 664–676. doi:10.1016/S0377-2217(00)00292-7
- Shafer, G. (1976). *A mathematical theory of evidence*. Princeton, NJ: Princeton University Press.
- Smets, P. (1997). Imperfect information: Imprecision - uncertainty . In Motro, A., & Smets, P. (Eds.), *Uncertainty management in information systems: From Needs to Solutions* (pp. 225–254). Dordrecht, The Netherlands: Kluwer Academic Publishers.
- Tavana , M., Busch, T.E. & Davis, E.L. (2011) Modeling Operational Robustness and Resiliency with High-Level Petri Nets' *International Journal of Knowledge-Based Organizations*, Vol. 1, No. 2, pp. 17-38.
- Tavana, M., Kennedy, D., & Mohebbi, B. (1997). An applied study using the analytic hierarchy process to translate common verbal phrases to numerical probabilities. *Journal of Behavioral Decision Making*, 10(2), 133–150. doi:10.1002/(SICI)1099-0771(199706)10:2<133::AID-BDM255>3.0.CO;2-5
- Tsai, M.-J., & Wang, C.-S. (2008). A computing coordination based fuzzy group decision-making (CC-FGDM) for web service oriented architecture. *Expert Systems with Applications*, 34(4), 2921–2936. doi:10.1016/j.eswa.2007.05.017
- van der Aalst, W. M. P. (1996). Petri-net-based workflow management software. In A. Sheth (Ed.), *Proceedings of the NFS Workshop on Workflow and Process Automation in Information Systems*, Athens, GA (pp. 114-118).
- van der Aalst, W. M. P. (1997). Verification of workflow nets. In *Application and Theory of Petri Nets* (LNCS 1248, pp. 407-426).
- van der Aalst, W. M. P., De Michelis, G., & Ellis, C. A. (Eds.). (1998). *Proceedings of workflow management: net-based concepts, models, techniques and tools*.
- van der Aalst, W. M. P., & van Hee, K. (2002). *Workflow management: Models, methods, and systems*. Cambridge, MA: MIT Press.
- van der Aalst, W. M. P., van Hee, K. M., & Houben, G. J. (1994). Modelling workflow management systems with high-level Petri nets. In *Proceedings of the 2nd Workshop on Computer-Supported Cooperative Work, Petri nets and related formalisms* (pp. 31-50).
- Wang, Y. M., & Elhag, T. M. S. (2006). Fuzzy TOPSIS method based on alpha level sets with an application to bridge risk assessment. *Expert Systems with Applications*, 31(2), 309–319. doi:10.1016/j.eswa.2005.09.040
- Yang, T., & Hung, C. C. (2007). Multiple-attribute decision making methods for plant layout design problem. *Robotics and Computer-integrated Manufacturing*, 23, 126–137. doi:10.1016/j.rcim.2005.12.002
- Zadeh, L. A. (1978). Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems*, 1, 3–28. doi:10.1016/0165-0114(78)90029-5
- Zammori, F. A., Braglia, M., & Frosolini, M. (2009). A fuzzy multi-criteria approach for critical path definition. *International Journal of Project Management*, 27(3), 278–291. doi:10.1016/j.ijproman.2008.03.006