Precise Determination of Non-additive Measures

Hai-Feng Guo, Wen Zheng
Deparment of Computer Science
University of Nebraskat at Omaha
Omaha, NE 68182-0500, USA
{haifengguo, wzheng}@mail.unomaha.edu

Song Ci
Dept. of Computer and Electronics Engineering
University of Nebraska at Lincoln
Omaha, NE 68182-0572
sci@engr.unl.edu

***Abstract*—The Choquet integral model has been shown useful in many practical applicaitons due to its distinguished feature that the interaction among predictive attributes toward the objective attribute can be properly reflected through a set of non-additive measures. Non-additive measures are traditionally determined using heuristic strategies or genetic algorithms; however, their solutions to Choquet integrals are often not unique or deterministic. Besides, raw attribute data are often not good enough to identify all coefficients, which leads to an incomplete Choquet model. In this paper, we present an efficient approach to determine the complete non-additive measures. The approach adopts a new data normalization, based on *median alignment*, for the Choquet model to overcome the above practical deficiencies raised in traditional methods. We show how the Choquet model is applied to the significance analysis for wireless cross-layer network design.**

I. INTRODUCTION

The Choquet integral [3] is a generalization of the Lebesgue integral, defined over a set of non-additive measures (also called fuzzy measures). Let $X = \{x_1, x_2, \cdots, x_n\}$ be a set of attributes, $f(x)$ be the observed or partially evaluated value on each attribute $x \in X$, f is a tuple of observed or partially evaluated values on X, and z be an objective. The linear/additive multiregression model is traditionally represented as a weighted sum $z = \sum_{x \in X} w_x f(x)$, where the weight w_x is also regarded as a Lebesgue measure w on a singleton $\{x\}$, since the linear model is equivalent to a Lebesgue integral $z = (L) \int_X f dw$. The Choquet integral model breaks the restriction that the combined contribution of $\{x_i, x_j\}$ toward the objective z is the weighted sum of their respective contributions. Instead, it uses a non-additive measure μ, which is defined over the powerset of X, and a Choquet integral, $z = (C) \int_X f d\mu$. It is clearly more powerful than the Lebesgue integral model since the non-additive measure μ considers the interaction among attibutes toward the objective. In such a setting, $\mu(\{x_i, x_j\})$ may not be a linear sum of $\mu(\{x_i\})$ and $\mu(\{x_j\})$; Lebesgue integral model thus becomes a special case of the Choquet integral model where the linear sum equation holds.

The non-additive model based on Choquet integral (the *Choquet model* in short) has been shown useful in many practical applications, such as classification [11], multicriteria decision making [6], [17], image and pattern recognition [7], [13], [22], data modeling [12], [21], and so on, due to its distinguished feature that the interaction among predictive attributes toward the objective attribute can be properly reflected through a set of non-additive measures. However, even though the theory and applications of the Choquet model have been well studied at length in the last half century, the practical uses of the Choquet model are quite limited within a few research groups in the community of fuzzy set theory.

The main problem of applying the Choquet model is how to determine the non-additive measure μ defined over the powerset of attributes X. The basic idea to solve the Choquet model is to reduce the non-linear regression model to the traditional linear multiregression model, so that the Choquet model can be easily solved by using the least-square method in a quadratic running time [9]. The idea was orginally proposed in [16], and was successfully applied on classification [5]. However, there is a problem that "bad" solutions ("extreme" values near 0) are often generated [9], [12] because raw attribute data are often not good enough to identify all 2^n coefficients. In practial applications, the obtained optimal solutions are sometimes too unreasonable to support the decision making. To get around such a bad-solution problem, a suboptimal algorithm, called HLMS (Heuristic Least Mean Squares) [7], was proposed. The heuristic strategy is based on a gradient algorithm and the idea of equilibrium point [7]. However, the obtained suboptimal solution may be quite different from the optimal one. Another popular strategy is to use genetic algorithm (GA) to determine the non-additive measures [21], [12], where each measure μ is coded in a gene. Due to randomness of chromosome generation and the huge 2^n dimensional search space, the obtained solutions at different running times are not unique, and actually quite different in general. Similar to the previous alogithms, "bad" solutions are often generated due to the trap of local optimum.

In this paper, we present a new approach on the Choquet model toward its efficient yet precise determination of non-additive measures. The main contributions of the paper can be summerized as follows.

- The main objective of the Choquet model is to identify all the 2^n non-additive measure coefficients, given a set of n attributes. However, raw attribute data are often not good enough to identify all 2^n non-additive measure coefficients, due to the reason that some measure coefficients are not directly applied during the data modeling. For this reason, we introduce a new data normalization algorithm, baseda on *median alignment*, so that all the non-additive measure coefficients defined over the preprocessed data

978–1–4244–1674–5/08/$25.00

are used in the model, and can hence be identified properly.

- Non-additive measures are traditionally determined in heuristic strategies or genetic algorithms; the least-squared algorithm for solving linear equations have to be applied repeatedly until no more or under certain very little improvement could be made, and the solution for the Choquet model is not unique in general. It often occurs that different runs on the same dataset give quite inconsistent coefficient results. With the help of our new data normalization algorithm, the non-additive measure can then be identified deterministically by applying the standard least squared algorithm once.
- We illustrate an application of the Choquet model on interdependency measure and significance analysis for the wireless cross-layer network design under uncertainties. The proposed approach captures the interdependency among the MAC protocol parameters and identify which subset of system parameters has the most significant effect on the performance metrics of interest under the current system and network conditions, so that we can fine tune those system parameters to improve the network throughput performance if necessary.

It is worthy to be mentioned that this paper is targetted to those application problems with the small number of attributes ($n \leq 10$), otherwise, the fuzzy measure containing around 2^n coefficients is too large to determine. Strategies how to solve Choquet models with large number of attributes can be found in [20], [15], [18], [8], [10]. The proposed approach in this paper is also useful to handle similar issues for solving large Choquet models.

The rest of the paper is organized as follows. Section II gives a brief introduction of the Choquet integral model, and how the coeefficients of the model is traditionally determined. Section III presents a new approach, based on the median-alignment data normalization, to determine the non-additive measures of the Choquet model. Section IV shows a case study and the simulation results. Finally, Section V gives a conclusion.

II. The Choquet Integral Model

In this section, we give a brief introduction on non-additive measure, the Choquet integral model, and how to determine the non-additive measure of the model.

A. Non-additive measure

Let $X = \{x_1, x_2, \cdots, x_n\}$ be a set of attributes. A *fuzzy measure* on X is a set function $\mu : \mathcal{P}(X) \to \mathcal{R}$ with a constraint $\mu(\emptyset) = 0$, where n is the number of attributes, $\mathcal{P}(X)$ is the powerset of X, and $\mathcal{R}$ is the real domain. We relax the following two traditional restrictions on fuzzy measures: (i) the co-domain of the set function μ is $\mathcal{R}$ instead of $\mathcal{R}^+$; (ii) the monotonicity, $A \subset B \subseteq X$ implies $\mu(A) \leq \mu(B)$, is not necessary.

The main characteristic of fuzzy measures is that they can express interactions among attributes being aggregated in a more flexible and precise manner. A fuzzy measure is said to be *additive* if $\mu(A \cup B) = \mu(A) + \mu(B)$ whenever $A \cap B = \emptyset$, otherwise, it is *non-additive*. In the applications of multicreteria decision making [10], the case of $\mu(A \cup B) > \mu(A) + \mu(B)$ is often called *postive interaction* or *positive synergy* between criteria A and B; whereas the case of $\mu(A \cup B) < \mu(A) + \mu(B)$ is called *negative interaction* or *negative synergy*, where the union of criteria (or attributes) does not bring anything more to the objective.

B. The model

The main feature of the Choquet integral model is that the interaction among attributes toward the objective can be properly measured through a non-additive fuzzy measure. Assume the data consists of l observations of the attributes $x_1, x_2, \cdots, x_n$ and the objective z, in the following form:

x_1	x_2	$\cdots$	x_n	z
f_{11}	f_{12}	$\cdots$	f_{1n}	z_1
f_{21}	f_{22}	$\cdots$	f_{2n}	z_2
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
f_{l1}	f_{l2}	$\cdots$	f_{ln}	z_l

where each row is an observation of attributes $x_1, x_2, \cdots, x_n$ and z. The observation of $x_1, x_2, \cdots, x_n$ can be regarded as a function $f : X \to \mathcal{R}$; hence the j-th observation of $x_1, x_2, \cdots, x_n$ is denoted by f_j, and we write $f_{ji} = f_j(x_i)$ where $1 \leq i \leq n$ and $1 \leq j \leq l$.

The interaction among predictive attributes X toward the objective z is described by a set function μ defined on the power set of X satisfying the condition of vanishing at the empty set, i.e., $\mu : \mathcal{P}(X) \to \mathcal{R}$ with $\mu(\emptyset) = 0$. The new non-additive multi-regression model is expressed as

$$z = e + \int_{(c)} f d\mu + N(0, \delta^2), \tag{1}$$

where e is a regression constant, $\int_{(c)}$ is the Choquet integral, f is an observeration of $x_1, x_2, \cdots, x_n$, μ is a fuzzy measure, and $N(0, \delta^2)$ is a normally distributed random perturbation with expection 0 and vairance δ^2. The Choquet integral $\int_{(c)}$ of the data observation f, w.r.t. a fuzzy measure μ, is defined as:

$$\int_{(c)} f d\mu = \int_{-\infty}^{0} [\mu(F_\alpha) - \mu(X)] d\alpha + \int_{0}^{+\infty} \mu(F_\alpha) d\alpha \tag{2}$$

where $F_\alpha = \{x \mid f(x) \geq \alpha\}$ for any $\alpha \in (-\infty, +\infty)$, and is called the α-cut set of f.

C. To determine the non-additive measure

The basic idea to solve the Choquet model is a two-step procedure. The first step is to reduce the non-linear multiregression model to the traditional linear multiregression model by converting each n-dimensional vector attribute datum to a 2^n-dimensional vector datum, which is defined over the powerset of attributes; and thus, the second step is to solve the linear model by using the standard least-square method.

x_1	x_2	x_3
5	10	8
7	0	5
-2	3	5
-6	-9	-2
$\vdots$	$\vdots$	$\vdots$

(a) Raw Data

$\{x_1\}$	$\{x_2\}$	$\{x_1, x_2\}$	$\{x_3\}$	$\{x_1, x_3\}$	$\{x_2, x_3\}$	$\{x_1, x_2, x_3\}$
0	2	0	0	0	3	5
2	0	0	0	5	0	0
0	0	0	2	0	5	-2
0	0	0	4	3	0	-9
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$

(b) Data after Transformation

Fig. 1. Data Transformation: $X \Rightarrow \mathcal{P}(X)$

Since the non-additive measure in the Choquet model is defined over the powerset $\mathcal{P}(X)$, the reduction step basically aggregates the observed data of individual attributes to the observation on sets. Consider a small set of sample data with the attribute set $\{x_1, x_2, x_3\}$ as shown in Figure 1. The first vector $(5, 10, 8)$ is aggregated to an equivalent subset-based representation, so that the contribution to the objective can be calculated using the non-additive measure in a linear way:

$$2 * \mu(\{x_2\}) + 3 * \mu(\{x_2, x_3\}) + 5 * \mu(\{x_1, x_2, x_3\}),$$

which is consistent to $\int_0^{+\infty} \mu(F_\alpha) d\alpha$ in the equation 2. On the other hand, if the observed data contains negative numbers, their contribution to the objective is cacluated via $\int_{-\infty}^{0} [\mu(F_\alpha) - \mu(X)] d\alpha$. That is, a negative unit of a set S ($S \in \mathcal{P}(X)$) contributes $\mu(X - S) - \mu(X)$ to the objective instead of $-\mu(S)$. Therefore, the contribution of the data vector $(-6, -9, -2)$ is equivalent to:

$$4 * \mu(\{x_3\}) + 3 * \mu(\{x_1, x_3\}) - 9 * \mu(X).$$

Once the observed data on X are converted to the new ones on $\mathcal{P}(X)$, then we have a linear multi-regression problem, in which the non-additive measure is basically the set of regression coefficients.

The method to determine the non-additive measure was orginally proposed in [16]. Direct application of this method often generates "bad" solutions [9], [12] because raw attribute datum are often not good enough to identify all 2^n coefficients. For example, consider a model of two attributes x_1 and x_2, and the datum $f(x_1)$ is consistently greater than $f(x_2)$ in each observation. Then the data under $\{x_2\}$ after conversion are all 0's, since all data of x_2 would be aggregated to the set $\{x_1, x_2\}$ due to $f(x_2) < f(x_1)$ in each observation. As a result, the measure $\mu(\{x_2\})$ cannot be determined by using the least-squared method.

Heuristic strategies, such as using gradient algorithms [7] or genetic algorithms [21], [12], were applied to get around such a bad-solution problem. However, there are a few problems: (1) the heuristic algorithms are not efficient since they are basically repetitive optimization procedures based on generating and comparing mutiple non-additive measures based on their regression residual errors; (2) the solutions are not unique in general, and possibly quite different from the expected ones due to the trap of local optimum.

```
procedure norm(vectors f, f', vector M)
/* j-th observed: f_j : (f_j1, ..., f_jn) */
/* j-th normalized: f'_j : (f'_j1, ..., f'_jn) */
/* alignment coefficient M : (M_1, ..., M_n) */
{
    for(k = 1; k <= n; k++)                    (1)
        M_k <- quickMedian(k);                 (2)

    for(k = 1; k <= l; k++)                    (3)
        normalize(f_k, f'_k, M);               (4)
}

procedure normalize(vector f_j, f'_j, M)
{
    for(k = 1; k <= n; k++)                    (5)
        f'_jk <- f_jk * M_1/M_k;               (6)
}
```

Fig. 3. A pseudo-code implementation of data normalization

III. New Approach using Median Alignment

A. Overview of new approach

Let $X = \{x_1, \cdots, x_n\}$ be a set of predicative attributes, and n be the number of attributes; l be the number of observation records. Figure 2 is a dataflow diagram illustrating how the observation data $f(X)$ are processed step by step toward the determinitation of its fuzzy mesaure μ defined over $\mathcal{P}(X)$. In the diagram, each rectangle represents a set of observed data, processed data, or computed results; each rounded rectangle represents a data processing procedure; and the directed edge shows the processing sequence and dataflow direction. Detailed explanation on procedures are shown in the following subsections.

B. Data Normalization

As mentioned earlier, direct reduction on raw data from the Choquet model to linear multiregression model may cause "bad" solutions, where non-additive measures on some subsets are often not able to be determined. In order to determine the non-additive measure on all subsets, it is important to ensure a necessary condition that the aggregated observation have *enough* non-zero data for each $S \in \mathcal{P}(X)$, so that the least-squared method is able to find its associated coeffecient.

We present a data normalization approach based on median alignment. Let $X' = \{x'_1, \cdots, x'_n\}$ be the set of new

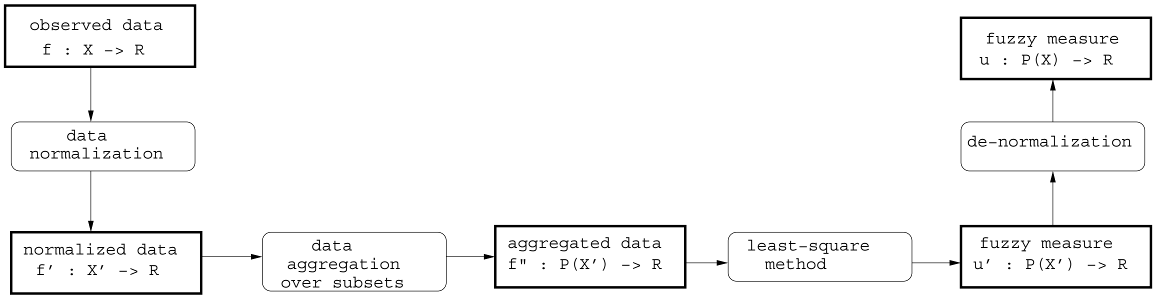

Fig. 2. A dataflow diagram

predicative attributes (different from the original attributes in scales) after normalization, and the function $f' : X' \leftarrow R$ be the normalized data observation. The main idea is to align the observed data of each predictive attribute with others' data along their medians, so that if the data samples are reasonably large, it is expected to have non-zero aggregated observation for each subset $S \in \mathcal{P}(X')$; and its non-additive measure can therefore be determined properly.

Figure 3 shows a simple pseudo-code implementation of the data normalization based on median alignment. The procedure `quickMedian(`k`)` is a standard divide-and-conquer, partition-based selection algorithm [1] using median-of-three partitioning to return the median among the observed data of x_k, which is a linear algorithm in time complexity $O(l)$. The procedure `normalize`$(f_j,\ f'_j\ M)$ is used to adjust each observed data of f_{jk} by multiplying a factor M_1/M_k, which is obviously an $O(n)$ algorithm (lines 5-6). Therefore, the total time complexity of the procedural `norm` is $O(nl)$, which is basically linear to the total number of individual data items.

C. Data Aggregation over Subsets

This subsection shows an efficient algorithm in Fig 4 on how the attribute data (n-dimensional vectors) after normalization are aggregated to a 2^n-dimensional vector data defined over the powerset of attributes. The Choquet model can then be reduced to a linear multiregression model after data aggregation.

As shown in Figure 4, given a vector of normalized data $f'_j = (f'_{j1}, \cdots, f'_{jn})$, we use a vector $f''_j = (f''_{j1}, ..., f''_{(j,2^n-1)})$ to represent aggregated data vector over the powerset of X, where $f''_{ji} = f''_{(j,i)}$ through the rest of paper. For each f''_{ji}, $1 \leq i \leq 2^n - 1$, the binary representation of i indicates whether an attribute variable is a member of the subset, that is, i is a subset indicator of $\{x'_k \mid \text{the } k\text{-th lowest bit of binary representation of } i \text{ is a } 1\}$.

The algorithm first applies an indirect insertion sorting procedure on the vector of normalized data. The purpose of indirect sorting is to sort the data items without really moving the data around, because each datum is associated to a data column under an attribute variable. The indirect sorting is implemented by using an auxiliary vector I of pointers to data and only rearranging the pointers. We simply use the insertion sorting due to the fact that it is assumed that the vector dimension n is less than or equal to 10. The data items in an increasing order are shown as $f'_{jI_1}, \cdots, f'_{jI_n}$.

procedure aggregate(vectors f', f")
/* j-th normalized: $f'_j : (f'_{j1}, \cdots, f'_{jn})$ */
/* j-th aggregated: $f''_j : (f''_{(j1)}, \cdots, f''_{(j,2^n-1)})$ */
{
(1) for($j \leftarrow 0;\ j < l;\ j$++) {
(2) indirectSort(f'_j, I);
(3) for($k \leftarrow 1;\ k < 2^n;\ k$++)
(4) $f''_{(jk)} \leftarrow 0$;
(5) d $\leftarrow 0.0$;
(6) s $\leftarrow 2^n - 1$;
(7) for($k \leftarrow 1; k \leq n; k$++) {
(8) $f''_{js} \leftarrow f'_{jI_k} - d$;
(9) $d \leftarrow f'_{jI_k}$;
(10) $s \leftarrow s - 2^{I_k}$;
(11) }
(12) }
}

procedure indirectSort(vector f'_j, vector I)
{
(13) for($k \leftarrow 1;\ k \leq n;\ k$++) {
(14) $q \leftarrow k$;
(15) for($t \leftarrow k;\ t > 1$ and $f'_{(j,q)} < f'_{(j,I_{(t-1)})}; t$--)
(16) $I_t \leftarrow I_{(t-1)}$;
(17) $I_t \leftarrow q$;
(18) }
}

Fig. 4. A pseudo-code implementation of data aggregation

We then use an efficient procedure, linear to n, to identify those subsets with non-zero aggregated values. We assume that the aggregated value f''_{jk} of each subset indictor k is initially set to zero (lines 3-4). As shown in Fig 4, we then introduce a subset indicator s, initially $2^n - 1$ representing the full set $\{x'_1, \cdots, x'_n\}$. Since the data vectors are sorted, the first datum, f'_{jI_1}, must be the value of $f''_{(j,2^n-1)}$, which represent the aggregated datum of the full set. The `for`-loop (k from 1 to n in lines 7-11) removes the variable x'_{I_k}, which holds the least data f'_{jI_k}among those variables in s, out of the set indicator s in each iteration. We use d denote the least datum among those variables in the previous (iteration) s. Therefore, f''_{js}, the aggregated observation datum on the subset s, should be assigned to $f'_{jI_k} - d$ (line 8).

The running time complexity of our data aggregation algorithm is $O(2^n)$, where n is supposed to be within a constant 10 or so. Actually, except that the initialization (lines 3-4) takes expenential running time due to the 2^n-dimension of f''_j, the indirect sorting takes $O(n^2)$ and the rest of the algorithm takes $O(n)$ running time. We use the insertion sort in the `indirectSort` procedure because insertion sort is supposed to be more efficient than quicksort in practice if n is small. This aggregation algorithm is better than the ones proposed in [21], [22], which are $O(n \times 2^n)$ time complexity.

D. De-normalization

Once the data in Choquet model have been transformed to 2^n-dimensional matrix linear model, the standard least-squared method is applied to determine the fuzzy measure u' over the powerset of X'. We omit the detail of this step since the least-squared method cannot be found from any textbook covering the multi-regression. The next question is how to derive the expected fuzzy measure u over the powerset of original attribute set X based on the fuzzy measure u'.

```
procedure deNorm(vector u', u)
/* measure over P(X'): u' : (u'_1, ..., u'_(2^n-1)) */
/* measure over P(X): u : (u_1, ..., u_(2^n-1)) */
{
(1)   for(s ← 1; s < 2^n; s++) {
(2)       for(k ← 1; k ≤ n; k++) {
(3)           x_k ← s mod 2;
(4)           s ← s/2;    % integer division
(5)       }
(6)       normalize(x, x', M);
(7)       transform(x', y);
(8)       u_s ← Σ_{k=1}^{2^n-1} y_k × u'_k;
      }
}
```

Fig. 5. A pseudo-code implementation of back reduction

A pseudo-code implementation of the de-normalization algirithm is shown in Fig 5. The variable s is a set indicator, from 1 to $2^n - 1$, which indicates an attribute set $S = \{x_k \mid \text{the } k\text{-th lowest bit of the binary representation of } s \text{ is } 1\}$. For each set indicator s, code lines $(2-5)$ construct a new vector of observed data by assigning all the $x_k \in S$ to 1 and the rest attribute varialbes to 0. After normalization to aligned vector x' and then transformation to a 2^n-dimensional y, we apply the Choquet integral with the fuzzy measure u' on the transformed y to find its objective contribution, which is equivalent to u_s. The time complexity of the procedure `deNorm` is $O(2^{2n})$, where n is a small constant within 10 or so. That is because each application of the Choquet integral model to an observation data takes $O(2^n)$; and the fuzzy measure u over the powerset $\mathcal{P}(X)$ contains $2^n - 1$ values, where each value, corresponding to a non-empty subset $S \in \mathcal{P}(X)$, requires an application of the Choquet integral model with the fuzzy measure u' to determine.

IV. Interdependency Measure and Significance Analysis on Cross-Layer Design

In this section, we show an application how to apply the Choquet model with the proposed approach for interdependency measure and significance analysis on cross-layer wireless network design [4]. We choose the IEEE 802.11 WLAN as our platform due to its popularity and complicated interdependency among different system parameters such as the number of users, the minimum contention window size, frame size, and data rate. So far, even though the IEEE 802.11 WLANs have gained a huge commercial success, the working dynamics of the MAC protocol of the IEEE 802.11 WLAN is still largely remaining unknown due to 1) the heuristic design of the protocol at its early development stage such as the exponential backoff scheme and 2) the backward compatibility requirement among different versions of the standard.

There has been a lot of research done on how to use cross-layer design to improve the throughput performance. In literature [19], [2], [14], throughput performance was improved by adapting frame size, contention window size, data rate, and so on. However, research also shows that none of the schemes for throughput enhancement work well in all cases due to the fact that the significance of each system parameter is also time-varying under different system and network situations. For example, increasing frame size is a good yet simple approach when the channel quality is good but not so when the channel quality is bad, where lowering the transmission data rate might be a good option. Things will be even more complicated when multidimensional optimization schemes are adopted in cross-layer design, where several parameters might be adjusted simultaneously for better throughput. Cautions have to be given in this case since different system parameters might react to the same system condition differently, resulting in losing gain of cross-layer design, some times, even degrading the system performance. Therefore, dynamically switching from one cross-layer scheme to another based on system dynamics is necessary to get all-time high system performance.

	lightly-loaded		heavily-loaded	
	good channel	bad channel	good channel	bad channel
$\mu(\{f\})$	18.655	25.444	23.07	-3.950
$\mu(\{c\})$	-19.129	-0.123	-3.301	-0.038
$\mu(\{f,c\})$	12.885	4.007	1.478	5.721
$\mu(\{d\})$	112.482	205.120	168.685	73.449
$\mu(\{f,d\})$	110.065	158.776	183.862	135.825
$\mu(\{c,d\})$	-31.615	-4.017	-16.223	6.615
$\mu(\{f,c,d\})$	16.312	43.448	5.595	33.335

TABLE I
Non-additive measures over different networks

We will use the proposed approach to quantitatively capture the interdependency among IEEE 802.11 MAC protocol parameters (frame size, minimum contention window size, and data rate, corresponding to `f`, `c` and `d` as shown in Table I) in cross-layer design under uncertainties and identify which set of system parameters has the most significant effect on the throughput performance under the network conditions. Table I shows our experimental results under four different network conditions considering both network status (lightly-loaded or heavily-loaded) and channel quality (good or bad). Each $\mu(s)$ in the table gives the quantified significance (positive or negative) from a set of system parameters s toward the throughput performance.

From experimental results, we may have the following observations, which are actually consistent to the literatures.

- When the network is lightly loaded, increasing data rate would improve the network throughput performance.
- When the network is heavily loaded, choosing larger frame size and higher data rate would lead to the most significant improvement on the throughput performance.
- For the case of lightly loaded network with good channel, there exists significant negative interaction between frame size and data rate, that is, $\mu(\{f, d\}) < \mu(\{f\}) + \mu(\{d\})$; and for the rest of the cases, there exists positive interaction between frame size and data rate.
- For the case of heavily loaded network with bad channel, $\mu(\{f, c\}) > 0$ even though both $\mu(\{f\})$ and $\mu(\{c\})$ have negative effects. It is true that increasing minimal contention window size may reduce the throughtput performance directly, but on the other hand, it reduces the packet loss rate caused by collision as well. Thus, if the frame size is increased as the same time, it is possible to increase the throughput performance instead.

With those observations, we can identify a small and major subset of system parameters to fine tune and improve the system performance effectively. To summarize, through this simple case study we demonstrated the basic idea of our proposed approach and validated the effectiveness of our model. The proposed approach is very helpful for us to understand the behaviors of the cross-layer design of a dynamic wireless network system.

V. Conclusion

We presented a new approach to identify the non-additive fuzzy measures in the Choquet model. The approach is based on a data normalization strategy to align the data among different predictive attibutes over their medians. Data alignment better ensures that once the observed data are aggregated over subsets, there will be *enough* non-zero values for each subset, so that each fuzzy measure can be identified. We gave a detailed description, including a dataflow diagram, pseudo-code algorithms for each major step and their time complexities, on how this normalization strategy is applied to solve the complete Choquet model. Our appraoch resolves a practical issue of applying the Choquet model that "bad" and non-deterministic solutions are often generated when heuristic strategies or genetic algorithms are adopted. Additionally, we presented an application how the Choquet model is applied to the significance analysis and interdependency measure for the cross-layer wireless network design.

References

[1] M. Blum, R.W. Floyd, V. Pratt, R. Rivest and R. Tarjan. Time bounds for selection. Journal of Computer and System Sciences, 7(1973):448-461.

[2] F. Cali, M. Conti, and E. Gregori: Dynamic Tuning of the IEEE 802.11 Protocol to Achiveve a Theoretical Throughput Limit. *IEEE/ACM Transactions of Networking*, 8(6): 785-799, 2000.

[3] G. Choquet. Thoery of capacities. *Annales de l'Institut Fourier*, 5:131-295, 1953.

[4] S. Ci, H.-F. Guo. *Characterizing Cross-Layer Behaviors Using Nonlinear and Incommensurable Observations*. The IEEE Global Communications Conference (GLOBECOM), 2007.

[5] M. Grabisch and J.M. Nicolas. Classification by fuzzy integral – performance and tests. *Fuzzy SEts & Systems, Special issues on Pattern Recognition*, 65:255-271,1994.

[6] M. Grabisch. Fuzzy integral in multicriteria decision making. *Fuzzy Sets & Systems*, 69 (1995), 279-298.

[7] M. Grabisch. *A new algorithm for identifying fuzzy measures and its application to pattern recognition*. Int. Joint Conf. of the 4th IEEE Int. Conf. on Fuzzy Systems and the 2nd Int. Fuzzy Engineering Symposium, march 1995, Yokohama, Japan, 145-150.

[8] M. Grabisch. K-order additive discrete fuzzy measures and their representation. *Fuzzy Sets and Systems*, 92:167–189, 1997.

[9] P. Miranda and M. Grabisch. Optimization issues for fuzzy measures. *International Journal of Uncertainty, Fuzziness, and Knowledge-Based Systems*, 7(6):545-560, 1999.

[10] M. Grabisch and M. Roubens. Application of the Choquet integral in multicriteria decision making. *Fuzzy Measures and Integrals - Theory and Applications*, pp. 348-274, Physica Verlag, 2000.

[11] M. Grabisch. Fuzzy integral for classification and feature extraction. In *Fuzzy Measures and Integrals — Theory and Applications*, Physica Verlag, 2000, pp. 348–374.

[12] M. Grabisch. Modelling data by the Choquet integral. In *Information fusion in data mining*, V. Torra (ed.), Physica Verlag, 2003, 135-148.

[13] J.M. Keller, P.D. Gader, and A.K. Hocaoglu. Fuzzy integrals in image processing and recognition. *Fuzzy Measures and Integrals – Theory and Applications*, pp. 435–466, Physica Verlag, 2000.

[14] Q. Liu, S. Zhou, and G. B. Giannakis: Cross-layer Combining of Adaptive Modulation and Coding with Truncated ARQ over Wireless Links. *IEEE Trans. Wireless Communications*, 3(5): 17461755, 2004.

[15] R. Mesiar and D. Vivona. Two-step integral with respect to fuzzy measure. *Tatra Mt. Math. Publ.*, 16(1999): 359-368.

[16] T. Mori and T. Murofushi. *An analysis of evaluation model using fuzzy measure and the Choquet integral*. In 5th Fuzzy System Symposium, pp. 207-212, Japan, 1989. In Japaneses.

[17] Y. Narukawa and T. Murofushi. *Decision Modelling Using the Choquet Integral*. First International Conference on Modeling Decisions for Artificial Intelligence, pp. 183–193, 2004.

[18] Y. Narukawa and V. Torra. Choquet integral based models for general approximation. *Artificial Intelligence Research and Development*, 100(2003): 39–50, 2003

[19] D. Qiao, S. Choi, and K. Shin: Goodput Analysis and Link Adaptation for IEEE 802.11a Wireless LANs. *IEEE Trans. on Mobile Computing*, 1(4): 278-292, 2002.

[20] M. Sugeno and K. Fujimoto. *Hierarchical Decomposition Theorems for Choquet Integral Models*. IEEE International Conference on Fuzzy Systems, pp. 2245-2252, 1995.

[21] Z. Wang, K. S. Leung, and J. Wang. A genetic algorithm for determining nonadditive set functions in information fusion. *Fuzzy Sets and Systems*, 102 (1999), 463-469.

[22] K. Xu, Z. Wang, M. L. Wong, and K.-S. Leung. Discover dependency pattern among attributes by using a new type of nonlinear multiregression. *International Journal of Intelligent Systems*, 16(8): 949-962, 2001.