EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

ELSEVIER European Journal of Operational Research 110 (1998) 342–367

Theory and Methodology

A weight-assessing method with habitual domains

Gwo-Hshiung Tzeng *, Ting-Yu Chen, Jih-Chang Wang

Energy and Environmental Research Group and Institute of Traffic and Transportation, College of Management, National Chiao Tung University, 114, 4F, Sec. 1, Chung Hsiao W. Rd., Taipei 100, Taiwan, ROC

Received 7 September 1996; accepted 6 June 1997

Abstract

Weights of criteria are the important factors in decision-making. However, from a behavioral perspective, traditional weighting methods account for too few factors to deal with decisions properly. Based on the behavior mechanism and the theory of habitual domains, this study is undertaken to develop a new weight-assessing model that treats decision-making as a dynamically adjusting process proceeding from the ideal to actual states. The new model is built upon the dynamic analysis for the connectivities between criteria instead of the static analysis of traditional models. Finally, we have studied Taipei City motorcycle users, mode-choice behavior through questionnaire in order to show the applicability of the new model. From the empirical results, it is found that our weight-assessing method has significant application potential in practice.

Keywords: Weight; Criteria; Decision-making; Behavior mechanism; Habitual domains; Connectivity

1. Introduction

In the process of decision-making, trade-offs between the criteria that influence the final decision must be made. These trade-offs can be computed in terms of a relative ratio of their importance, which can be presented in a "weight" form. From the viewpoint of behavior, the influential factors determining the weight of a decision-making criterion include: the difference between the ideal and actual values of the criterion (i.e., level of charge), the diversification and intensification of other ideas which can activate the criterion (i.e., connectivity), the duration (or tenure) of the criterion belonging to the core of habitual domains (i.e., the frequency of input stimuli), the decision-maker's personality and social–economic attributes, the intrinsic value of the weight, and the interaction among other criteria. Traditional weighting methods (Hwang and Yoon, 1981), such as the eigen-vector method, weighted least-square method, entropy method, utility function method, consider only the "interaction among other criteria" and forego consideration of the other factors.

* Corresponding author. E-mail: ghtzeng@ccsun5.nctu.edu.tw.

0377-2217/98/$19.00
PII S0377-2217(97)00246-4

In general, most traditional weighting methods are based on static analysis, and their results usually only reflect the intuition or perception of decision-makers at the time of analysis. In fact, weights of criteria must be variable in different situations including input information, time, learning process, environment, etc. Thus, it is not easy to clearly delineate the absolute weightings for decision criteria. On the other hand, our model treats decision-making as a dynamically adjusting process from the ideal state to the actual state, allowing us to realize the dynamic change of weights depending on different situations.

This research considers most of these factors simultaneously using the concept of habitual domains. The habitual domains concept was first introduced by Yu in 1980. It claims that a human being's decision-making process is gradually fixed by habit. The main idea of habitual domains is that the set of ideas and concepts that are encoded and stored in the brain tend to progressively stabilize in the absence of an extraordinary destabilizing event. Thus, thinking processes will reach some steady state or may even become fixed. Once the habitual domain is extended, it could greatly enhance the quality of decision-making.

In order to express the changes in weights throughout the decision-making process, we use connectivities between criteria to set up a fuzzy directed graph that is a collection of crisp criterion sets. According to the connectivity network and neighborhoods identified by the connectivities, this research develops the weight-assessing method with habitual domains as an alternative to traditional weighting methods. Our model can assess weights based on behavior mechanisms and overcome many of the disadvantages of traditional weighting methods.

This paper is organized as follows: in Section 2, we review traditional weighting models. In Section 3, we introduce human behavior mechanisms and describe four hypotheses that capture the basic workings of the brain. We then introduce the concept of habitual domains. Section 4 presents a model of activation propensity and connectivity. We have started with two very practical examples before doing the mathematical work, including the mode choice behavior and the house-purchasing decision-making. Then we use the connectivities between criteria to establish the network structure in Section 5. Section 6 introduces an algorithm for weight assessing and provides an empirical study of motorcycle travelers. We end the paper with concluding remarks in Section 7.

2. Review of weight-assessing method

There are five primary components in any decision-making process, including: (i) decision alternatives; (ii) decision objectives or criteria; (iii) decision outcomes; (iv) preference structure; (v) information inputs (Yu, 1990). Weight is the most general form of preference structure. There are certain benefits in defining a preference structure by a set of weights, such as: (i) it allows the importance of each objective to be represented as a set of numbers; (ii) the ratio of two objectives is equal to their "relative importance"; (iii) the sum of all weights is equal to 1 (Saaty, 1980; Hwang and Yoon, 1981).

There is an abundance of research on weighting characteristics which might be applied to the decision-making process including: rating method (Eckenrode, 1965); utility function method (Keeney and Raiffa, 1976; Keeney and Nair, 1977); entropy method (Zeleny, 1974; Nijkamp, 1977); extreme weight approach; random weight approach (Voogd, 1983); LINMAP (Srinivasan and Shocker, 1973); analytic hierarchy process (AHP) (Saaty, 1977, 1980); least-square method; logarithmic least-square method; geometric mean method (Krovak, 1987; Cook and Kress, 1988). From a structural viewpoint, there are two types of weighting criteria: subjective and objective.

The objective weight can be computed from the outcomes without asking the perceptions of the decision-makers. For example, the extreme weight approach, random weight approach, and entropy method are all objective weight-assessing methods. The entropy method is one of the best objective weight-assessing methods. Entropy is a physical measurement of the second law of thermal-dynamics and has become an important concept in the social sciences as well as in the physical sciences.

In the information theory of Shannon, the entropy is used to measure the expected information content of certain messages. Entropy in information theory is a criterion for the amount of "uncertainty" represented by a discrete probability distribution. The entropy method assumes that those criteria with less uncertainty are more important. Moreover, the method computes the anti-uncertainty (amount of information) of each criterion based on its possible outcomes, and normalizes them to a set of weights whose summation is equal to 1.

AHP is the well-known type of subjective weight-assessing method. AHP was introduced by Saaty in the 1970s. AHP organizes all objectives into a hierarchical structure. In AHP, the objectives are independent of each other in the parallel level, and the summations of weights in the same level are equal to their direct higher objective. Saaty suggests two techniques for obtaining the information on preference: pairwise comparison and eigen-vector computing. In fact, we can get these values by a direct-rating process or compute them through the least-square method, without affecting the validity of the AHP model.

These objective methods show that the weight of a criterion is relative to its clearness. Thus, in the entropy method, the clear criteria are more important than the fuzzy criteria. Subjective methods show us to obtain the information on preference by asking the decision-maker. Some of these methods provide ratings or pairwise-comparison techniques, while some suggest that we organize the objectives as a hierarchical structure. There is a common assumption in all of these weight-assessing schemes that the preference structure exists; the problem is how to obtain it. Thus, the question arises: if the preference structure is not stable, will a weight-assessing approach be useful? Is the importance of food to a hungry person similar to that for a normal person? Surely it is not. From the viewpoint of habitual domains, the weight comes from the charge structure, and the difference between the perceived actual state and the ideal state is the primary consideration within the charge structure. In other words, the distance between "where you want to go" and "where you are" decides the weight.

3. Habitual domains

In Section 3.1 we introduce some basic human behavior mechanisms. In addition, we shall describe four hypotheses that capture the basic workings of the brain: circuit pattern, unlimited capacity, efficient restructuring, and the analogy and association hypothesis. Then, we introduce the concept of habitual domains in Section 3.2. In Section 3.3, we shall introduce four classes of decision problems that are to be selected for particular situations based on the regularity and availability of skill sets.

3.1. Behavioral bases for decision-making

The main idea of the model of Yu (1985) of human decision behavior is that each human being has an endowed internal information-processing and problem-solving capacity that is consciously allocated as needed to various activities and events over time to adapt to, and achieve in, the multi-dimensional human environment. The brain is the human internal information processing center. Moreover, it is recognized that when external stimuli are cognitively attended to a human being, a special sequence of circuit patterns of activated neurons, containing the cognitive function, appears in the brain. This sequence represents one of the many possible cognitive functions that has been engendered by the stimuli. Some of the major categories of cognitive brain function include: encoding, storing, retrieving, and interpretation.

Yu (1990) summarized memory and thought processes according to four basic hypotheses: (i) the circuit pattern hypothesis (that is, human memory and thought can be represented by electrochemical patterns in the brain cells); (ii) the unlimited capacity hypothesis (that is, the capacity of memory that our brain can encode (not retrieve) is practically unlimited); (iii) efficient restructuring hypothesis (that is, our memory can be restructured in an efficient way so as to effectively process the encoded information); (iv) analogy and association hypothesis (that is, our brain interprets incoming information using analogy and association with the existing memory).

Among these four hypotheses, the analogy and association hypothesis is one of the most pervasive and important observations concerning human cognitive processes. Most people conjure up an impression in response to receiving an abstract symbol and their cognitive ability then enables them to retrieve information and perform analysis is conserving the symbol. According to Yu's definition, the analogy and association hypothesis can be stated as follows:

> The perception of new events, subjects, or ideas can be learned primarily by analogy and/or association with what is already known. When faced with a new event, subject, or idea, the brain first investigates its features and attributes in order to establish a relationship with what is already known by analogy and/or association. Once the right relationship has been established, the whole of the past knowledge (preexisting memory structure) is automatically brought to bear on the interpretation and understanding of the new event, subject or idea.

The main points to be emphasized are: (i) there is a preexisting code or memory structure which can potentially alter or aid in the interpretation of an arriving symbol; (ii) a relationship between the arriving symbol and the preexisting code must be established before the preexisting code can play its role in interpreting the arriving symbol.

3.2. Habitual domains

The concept of habitual domains (HD) was first formulated by Yu (1980). It claims that a human being's decision-making process is gradually fixed by habit. The main idea of habitual domains is that the set of ideas and concepts that are encoded and stored in the brain tend to progressively stabilize with time and in the absence of an extraordinary destabilizing event will approach a steady state (Yu, 1990).

There are two kinds of thoughts stored in human memory: (i) the ideas that can be activated in thinking processes; and, (ii) the operators that transform the activated ideas into other ideas. The operators are related to thinking processes or judging methods. Generally speaking, operators are also ideas. However, because of their ability to transform or generate (new) ideas, they are called operators.

Habitual domains have four primary elements (Yu, 1991): (i) potential domain, PD_t, which is the collection of ideas or operators that can be potentially activated at time t (or stage t); (ii) actual domain, AD_t, which is the set of ideas or operators that are actually activated at time t (or stage t); (iii) activated probability, AP_t, which is defined for each subset of PD_t and is the probability that a subset of PD_t is actually activated or is in AD_t; and, (iv) reachable domain, $R(I_t, O_t)$, which is the set of ideas or operators that can potentially be reached from the initial set of ideas, I_t, and the initial set of operators, O_t.

Given a decision-making problem E that catches a decision-maker's attention at time t (or stage t), the propensity for an idea v_i to be activated is denoted by $P_t(v_i, E)$. As with probability functions, $P_t(v_i, E)$ may be estimated by determining its relative frequency, as well as statistical probability (Yu and Zhang, 1989). The *α-core* K_t of the HD for problem E at time t (or stage t) is the collection of the ideas or concepts that can be activated with a propensity larger than or equal to α. That is,

$$K_t(\alpha, E) = \{v_i \in \mathrm{HD} \mid P_t(v_i, E) \geqslant \alpha\}. \tag{1}$$

3.3. Classification of decision problems

Depending on the perception of the availability of skill sets and the α-cores of HD, we can classify decision problems into four categories: routine problems, mixed routine problems, fuzzy problems and challenging problems (Shi and Yu, 1987; Yu and Zhang, 1992a, b). We note that what is unknown to the decision-maker may be known to another. Therefore, the classifications below depend on each decision-maker's HD.

(i) *Routine problems.* For a routine problem, the idea set that is needed to successfully solve problems is well known to the decision-maker, and the decision-maker has acquired the set. With proper training and practice, most decision-makers can develop the capability to solve routine problems.
(ii) *Mixed routine problems.* A decision problem is called a mixed routine problem if it consists of a number of routine subproblems.
(iii) *Fuzzy problems.* The truly needed idea set is only fuzzily known to the decision-maker. Therefore, the decision-maker has not yet mastered the skills, concepts and operations necessary for successfully solving these problems.
(iv) *Challenging problems.* The truly needed idea set is unknown or only partially known to the decision-maker. These problems cannot be successfully solved by the α-core of HD, no matter how small α is.

4. Activation propensity and connectivity

4.1. Measurable space

In order to specify measurable space in our study, we first introduce the concept of goal setting. For a decision problem, denoted by E, there exists a set of goal functions to be achieved for its satisfactory solution. The goal functions in the internal information-processing center are used to measure the many dimensional aspects of the decision problem. For each goal function there is an ideal state or equilibrium point to reach and maintain. This process is called goal setting. Goal functions can be measured by a collection of elementary criteria, $\{v_1, v_2, v_3, \ldots, v_n\}$ (n is the number of criteria), which is finite. Each goal function is a subset of the total collection of all elementary criteria.We say that two goal functions are related or associated if their two corresponding criterion subsets have a non-empty intersection. We can consider the elementary criteria $\{v_1, v_2, v_3, \ldots, v_n\}$ to be the discussion universe, HD, for the problem E.

For example, when drivers of private vehicles begin a trip, the first problem is how to choose their route. The objectives which drivers would consider during the route choice process include the minimal travel cost, the fastest driving speed, optimal safety and comfort, the fewest risks, and the most familiar route. It is not necessary to optimize each objective for users, but to acquire only the satisfactory level of each objective. Therefore, for the route choice problem E, its discussion universe HD = {travel time, delay, driving speed, degree of safety, degree of comfort, degree of risk, and familiarity to the route}. Another example: when we desire to purchase a house, the criteria to be taken into consideration would be of price, size of the house, age of the house, distance of the house from the office, and convenience to shopping. Additionally, these criteria would have made up as the discussion universe for the decision-making of house-purchasing.

Let HD (habitual domains) be the discussion universe, i.e. a set consisting of all vertices (i.e. criteria) for the discussion of problem E. The family of sets consisting of all the subsets of HD is referred to as the power set of HD and is indicated by P (HD). $\sigma(\mathrm{HD}) \subset \mathrm{P}(\mathrm{HD})$ is a family of subsets of HD, so that:

(i) $\phi \in \sigma(\mathrm{HD})$ and $\mathrm{HD} \in \sigma(\mathrm{HD})$;
(ii) if $\Lambda \in \sigma(\mathrm{HD})$, then $\overline{\Lambda} \in \sigma(\mathrm{HD})$;
(iii) $\sigma(\mathrm{HD})$ is closed under the operation of set union; that is, if $\Lambda_1, \Lambda_2, \Lambda_3, \ldots \in \sigma(\mathrm{HD})$, then $\bigcup_{i=1}^{\infty} \lambda_i \in \sigma(\mathrm{HD})$

The set $\sigma(\mathrm{HD})$ is usually called the Borel field or σ-field. We usually take $\sigma(\mathrm{HD})$ to be the smallest σ-field containing as numbers all of the sets of particular interest, and $\sigma(\mathrm{HD})$ is called a σ-algebra. We treat HD as a measurable space,

$$\mathrm{HD} = (\mathrm{HD}, \sigma(\mathrm{HD}), \mu), \tag{2}$$

where $\sigma(\mathrm{HD})$ is a σ-algebra generated by HD, and μ is a meaning measure. Assume the HD is finite and that $\sigma(\mathrm{HD}) = \mathrm{P}(\mathrm{HD})$, i.e., every subset of HD is measurable.

4.2. Activation propensity and connectivity

Given an input stimulus S_t at time t (or stage t), the propensity for criterion v_i to be activated is denoted by $P_t(v_i S_t)$. For $\alpha \in [0, 1]$, the α-core of HD at time t (or stage t), denoted by $K_t(\alpha, \mathrm{HD})$, or simply by $K_t^{\alpha}(\mathrm{HD})$ when no confusion can occur, is defined as the collection of criteria that can be activated with a propensity larger than or equal to α. That is,

$$K_t(\alpha, \mathrm{HD}) = \{v_i \in \mathrm{HD} \mid P_t(v_i, S_t) \geqslant \alpha\}. \tag{3}$$

Take route choice as an example: when road users have received the message of a traffic jam on the road ahead, the difference between the ideal and actual values of travel time will be enlarged. Thus, the activation propensity of travel time will increase and such criterion will become the core of decision-making as travelers make their en-route switching. Should road users receive the message of rock slide on the road ahead, their decision core will be transferred to safety. Similarly, the price of the house is most often not the core of habitual domain for rich people during their house-purchasing decision process, and price is, therefore, not a major factor to be considered.

Note that the α-core $K_t(\alpha, \mathrm{HD})$ is just the closed α-cut of a fuzzy subset of HD with membership function P_t (Yu and Zhang, 1990). Furthermore, the activation propensity function can be generalized into a function defined on $\mathrm{HD} \times \mathrm{HD}$:

Definition 4.1. For any v_i and v_j in HD, $P_t(v_j, v_i)$ denotes the propensity that criterion v_j will be activated when the criterion v_i is presented to the input stimuli at time t (or stage t).

It is trivially known that the higher the activation propensity, the more strongly the criteria are connected and vice versa. Therefore, we can reasonably treat the activation propensity function as an approximation of the connectivity function and define the connectivity function following the definition of (Yu and Zhang, 1990) as:

Definition 4.2. A function $C_t(v_i, v_j)$ defined on $\mathrm{HD} \times \mathrm{HD}$ at time t (or stage t) is called a connectivity function on HD if it satisfies the following axioms:
(i) $C_t(v_i, v_j) \in [0, 1]$;
(ii) $C_t(v_i, v_i) = 1, \forall v_i \in \mathrm{HD}$.

Because the number of stimuli perceived by the senses of an individual at any point in time is enormous, selective perception is required. Thus, arriving symbols are processed in an unrefined way for the sake of convenience and efficiency. According to the analogy and association hypothesis, when faced with an external stimulus, the brain first investigates its features and attributes in order to establish a relationship with what is already known by analogy and association. The relationship can be treated as the connectivity function. Once the right connectivity relationship has been established, the whole of the past knowledge (preexisting memory structure) is automatically brought to bear on the interpretation and understanding of the stimuli, resulting in a decision being made.

According to the definition, if C_t is a connectivity function, then $C_t(v_i, v_j)$ is called the connectivity from v_i to v_j at time t (or stage t). If $C_t(v_i, v_j) > C_t(v_i, v_k)$, then we say that v_i is more strongly connected to v_j than to v_k. In terms of the route choice problem, should the driving speed of the alternative route be slower, the road user would then have to spend more time to reach his destination. Thus, driving speed can be connected with travel time by analogy and association. Likewise, when the driving speed is slower, the feeling of comfort that the route brings to users would be less; as a result, it can be seen that driving speed is connected with degree of comfort. Generally speaking, driving speed is more likely to be associated with travel time; in a word, the connection between driving speed and travel time is stronger than that of driving speed with comfort.

It should be noted that learning processes are usually directed, i.e., the connectivity from v_i to v_j is not equal to the connectivity from v_j to v_i. Thus, C_t is not necessarily a symmetric connectivity function; in other words, the activation propensity is non-symmetric. We can summarize the above discussion with Proposition 4.1.

Proposition 4.1. *Assume that $P_t(v_j, v_i)$ is such that $P_t(v_j, v_i) = 1$ if $v_t = v_j$. Let $C_t(v_i, v_j) = P_t(v_j, v_i)$ for all v_i and v_j in the* HD *at time t (or stage t). Then, $C_t(v_i, v_j)$ is a connectivity function that satisfies the conditions of Definition 4.2.*

It is noted that people usually utilize an approximation of the necessary criteria in solving a particular decision-making problem E based on the limits of personal charge structure, external information, attention allocation, self-suggestion or physiological monitoring, etc. Decision-makers are also apt to be affected by external considerations such as time and cost. Only in the absence of palpable external stimulation, goal setting, charge structure, and information input will the solution process maintain steady. This indicates that the decision-makers' perceived connection between criteria is only fuzzily known. In other words, the connectivity between the criteria for a decision-making problem E is roughly, but not clearly, known. Consequently, the connectivity function on HD will continue to change as long as the situation, viewpoint, or physiological state varies. This being the case, we define $C_t(v_i, v_j)$ as a function of time t (or stage t).

5. Connectivity network

5.1. Network structure

We assume that HD is discrete and finite; that is, $\text{HD} = \{v_1, v_2, v_3, \ldots, v_n\}$ is finite. We treat HD as a vertex (or node) set. Let A denote a subset of HD × HD, i.e., $A \subseteq \text{HD} \times \text{HD}$ and A is called an arc (or link) set. An element in A is called an arc. For any v_i and v_j in HD, $(v_i, v_j) \in A$ is called an arc that joins v_i and v_j (starting from v_i and arriving at v_j). Now we set up a connectivity network G.

A fuzzy directed graph (i.e., digraph) $G(\text{HD}, R_t)$ consists of a finite set $\text{HD} = \{v_1, v_2, v_3, \ldots, v_n\}$ and a fuzzy relation R_t on HD at time t (or stage t), where the relation R_t satisfies:

(i) $R_t(v_i, v_j) \in [0, 1]$;

(ii) $R_t(v_i, v_i) = 1, \forall v_i \in \text{HD}$ (reflexivity).

Let the fuzzy relation R_t on HD be the connectivity function $C_t(v_i, v_j)$ on HD at time t (or stage t). Therefore, the digraph is represented by $G(\text{HD}, C_t)$, i.e., the connectivity network. Consider a connectivity function C_t defined on a finite set HD. The connectivity C_t is interpreted in terms of a connectivity network. That is, for $v_i, v_j \in \text{HD}, C_t(v_i, v_j)$ is the grade of adjacency from v_i to v_j at time t (or stage t). Furthermore, the adjacency matrix of $G(\text{HD}, C_t)$ is defined as $[C_t(v_i, v_j)]$, where $v_i, v_j \in \text{HD}$.

An α-core of $G = (\text{HD}, C_t)$ is defined to be a crisp digraph $K_t^\alpha(G)$, when

$$K_t^\alpha(G) = \{v_i, v_j \in G \mid C_t(v_i, v_j) \geqslant \alpha\}. \tag{4}$$

Clearly the α-core $K_t^\alpha([C_t(v_i, v_j)])$ is the crisp adjacency matrix of the crisp digraph $K_t^\alpha(G)$.

The concept of a connectivity network G can be considered to be a collection of its α-cores. In particular, $K_t^0(G)$ is always a complete digraph; that is, every pair of criteria in $K_t^0(G)$ is adjacent.

5.2. Assessment of connectivity

In this section we discuss a method for assessing connectivity functions based on max–min operators. Before further discussion, we know that the arc (v_i, v_j) is active at the level α if $C_t(v_i, v_j) \geqslant \alpha$; that is, (v_i, v_j) is present in $K_t^\alpha(G)$ at time t (or stage t). Thus, $C_t(v_i, v_j)$ is equal to the maximum level at which the arc (v_i, v_j) is active.

A walk in a connectivity network $G(\mathrm{HD}, \mathrm{C_t})$ from v_{i1} to v_{ik} is a sequence of criteria $v_{i1}, v_{i2}, \ldots, v_{ik}$ that are connected by the arcs $(v_{i1}, v_{i2}), \ldots, (v_{ik-1}, v_{ik})$ (these arcs are usually considered to be part of the walk). A walk is called simple if each criterion appears in it only once at most, and a simple walk is called a path. Hence a path in G is a sequence of distinct criteria such that for all (v_i, v_j), $C_t(v_i, v_j) \geqslant \alpha$; in other words, a path in G is also called active at the level α if the path is present in $K_t^\alpha(G)$. The strength of the path is min $\{C_t(v_i, v_j)\}$ for all criteria contained in the path. The length of a path is the number of criteria contained in the path. The directed path also represents a learning sequence. It should be noted that some learning processes perhaps express a directed walk, not a path. However, as we know, if the walk is active at some level, the path contained in the walk is also active at the same level. Such being the case, in our study we consider only those situations in which learning processes are represented by the paths for simplicity.

A criterion v_j of G is called α-reachable from another criterion v_i if there is a path from v_i to v_j in the crisp digraph $K_t^\alpha(G)$. Furthermore, any criterion v_i is α-reachable from v_i itself, for any $\alpha \in [0, 1]$. G is α-connected if and only if all pairs of criteria of G are α-reachable; that is, G is called α-connected if $K_t^\alpha(G)$ is connected (Miyamoto, 1990).

Given a connectivity network $G(\mathrm{HD}, C_t)$, a connectivity function C_t^2 is defined by $C_t^2(v_i, v_j) = \alpha$ if and only if there is a path of length 2 from criterion v_i to criterion v_j in $K_t^\alpha(G)$, and for any $\varepsilon > 0$ there is no path of length 2 from v_i to v_j in $K_t^{\alpha+\varepsilon}(G)$ at time t (or stage t). Therefore, $C_t^2(v_i, v_j)$ is the maximum level of the α-core so that there is a path of length 2 starting from v_i and arriving at v_j.

The discussion concerning C_t^k, $k > 2$, can be carried out in a similar manner. First, $C_t^k(v_i, v_j) = \alpha$ if and only if there is a path of length k from criterion v_i to criterion v_j in $K_t^\alpha(G)$ at time t (or stage t). Moreover, if $C_t^k(v_i, v_j) = \alpha$, then there is no path of length k from v_i to v_j in $K_t^{\alpha+\varepsilon}(G)$ for any $\varepsilon > 0$ at time t (or stage t). That is, $C_t^k(v_i, v_j)$ is the maximum level of the α-core so that there is a path of length k starting from v_i and arriving at v_j. Consequently, the connectivity $C_t^k, k \geqslant 2$ is calculated by the max–min composition rule:

Proposition 5.1. *Given a connectivity network $G(\mathrm{HD}, C_t)$, then the connectivity C_t^k at time t (or stage t) is calculated by the following iterative formula:*

$$C_t^2(v_i, v_j) = \max_{v \in \mathrm{HD}} \min[C_t(v_i, v), C_t(v, v_j)], \tag{5}$$

$$C_t^k(v_i, v_j) = \max_{v \in \mathrm{HD}} \min[C_t^{k-1}(v_i, v), C_t(v, v_j)] \quad (3 \leqslant, k \leqslant, n), \tag{6}$$

where n is the number of criteria in G.

Proof. Find out v' such that v' is the element belonging to HD that achieves

$$\max_{v \in \mathrm{HD}} \min[C_t(v_i, v), C_t(v, v_j)] = \min[C_t(v_i, v'), C_t(v', v_j)].$$

Let $\alpha' = \min[C_t(v_i, v'), C_t(v', v_j)]$. Then both arc (v_i, v') and arc (v', v_j) are active at the level α'. This means that there is a path of length 2 at the level α' (i.e., in $K_t^{\alpha'}(G)$) starting from v_i and arriving at v_j. Therefore, we have the following inequality:

$$C_t^2(v_i, v_j) \geqslant \min[C_t(v_i, v'), C_t(v', v_j)] \\ = \max_{v \in \mathrm{HD}} \min[C_t(v_i, v), C_t(v, v_j)].$$

On the other hand, for any $\varepsilon > 0$ there is no path of length 2 from v_i to v_j in $K_t^{\alpha'+\varepsilon}(G)$ at time t (or stage t). That is, for any $\varepsilon > 0, C_t(v_i, v) < \alpha' + \varepsilon, C_t(v, v_j) < \alpha' + \varepsilon$, for all $v \in$ HD. Furthermore, from the definition of $C_t^2(v_i, v_j)$, we have $C_t^2(v_i, v_j) < \alpha' + \varepsilon$.
This means that

$$C_t^2(v_i, v_j) > \max_{v \in \mathrm{HD}} \min[C_t(v_i, v), C_t(v, v_j)]$$

does not hold for any $v_i, v_j \in$ HD.

Hence, we have proved the first equation of the proposition. Moreover, the second equation can be proved in the same way except for the "path of length k" condition. For simplicity, we omit the details. The connectivity function C_t^k can be calculated by the iterative formula described in Proposition 5.1. □

Let $C_t^*(v_i, v_j) = \alpha$ if and only if the criterion v_j is reachable from v_i in $K_t^{\alpha}(G)$, and v_j is not reachable from v_i in $K_t^{\alpha+\varepsilon}(G)$ for any $\varepsilon > 0$ at time t (or stage t). Then, $C_t^*(v_i, v_j)$ is the maximum level such that v_j is reachable from v_i; that is, v_j is α-reachable from v_i if and only if $C_t^*(v_i, v_j) \geqslant \alpha$. The connectivity C_t^* on HD described above is called a connectivitiy index. The connectivity index is calculated from $C_t^k, , k = 1, 2, 3, \ldots, n$ (n is the number of criteria), as described in Proposition 5.2.

Proposition 5.2. *Given a connectivity network* $G(\mathrm{HD}, C_t)$, *then the connectivity index* C_t^* *at time t* (or stage t) *is calculated from* C_t^k, *where* $k = 1, 2, 3, \ldots, n$:

$$C_t^*(v_i, v_j) = \max\{C_t(v_i, v_j), C_t^2(v_i, v_j), \ldots, C_t^k(v_i, v_j), \ldots, C_t^n(v_i, v_j)\}. \tag{7}$$

Furthermore, the connectivity index C_t^* *is also given by:*

$$C_t^*(v_i, v_j) = \max\{\min\{C_t(x, y) | (x, y) \in Path(v_i, v_j)\} \,|\, Path(v_i, v_j)\}, \tag{8}$$

where $Path(v_i, v_j)$ *is a path in G from* v_i *to* v_j.

Proof. From the definition it is obvious that the connectivity index $C_t^*(v_i, v_j)$ is equal to the maximum level of all the paths of arbitrary length (it is noted that the length of the path is always less than or equal to the number of criteria in HD) from v_i to v_j, which in turn is equal to the maximum of $C_t^k(v_i, v_j)$ the maximum level of paths of length k, for all k. This implies the first equation. As for the latter equation, that is simply another expression of the first equation. □

Remark 5.1. Given a connectivity network $G(\mathrm{HD}, C_t)$, $C_t^*(v_i, v_j)$ is the connectivity index from v_i to v_j at time t (or stage t) which satisfies the following axioms:
(i) $C_t^*(v_i, v_j) \in [0.1]$;
(ii) $C_t^*(v_i.v_t) = 1, \ \forall v_i \in$ HD.

Consequently, we can assess the connectivity index C_t^* by the max-min extension (according to the max–min composition rule) discussed above. Moreover, the max–min extension of the connectivity function C_t can be called the "conservative extension" (Yu and Zhang, 1990).

We have already shown that learning processes are usually directed, i.e., the connectivity from v_i to v_j may not be equal to the connectivity from v_j to v_i. Hence, C_t is not necessarily a symmetric connectivity function. Furthermore, the connectivity function may be non-transitive. C_t is (max–min) transitive if and only if

$$C_t(v_i, v_k) \geqslant \max_{v_j \in \mathrm{HD}} \min[C_t(v_i, v_j), C_t(v_j, v_k)] \tag{9}$$

is satisfied for each pair $(v_i, v_k) \in \mathrm{HD} \times \mathrm{HD}$. A C_t failing to satisfy the above inequality for some criteria of HD is called non-transitive; that is,

$$C_t(v_i, v_k) < \max_{v_j \in \mathrm{HD}} \min[C_t(v_i, v_j), C_t(v_j, v_k)] \tag{10}$$

for some $(v_i, v_k) \in \mathrm{HD} \times \mathrm{HD}$. For example, if there is no arc connecting v_k directly from v_i, this implies that it is practically impossible to learn v_k from v_i directly. Therefore, the connectivity function C_t does not require symmetry and transitivity.

5.3. Generalized current domain

As mentioned before, there exists a set of goal functions that must be achieved for the satisfactory solution of a decision problem E. Goal functions can be measured by finite elementary criteria, $\{v_1, v_2, v_3, \ldots, v_n\}$ (n is the number of criteria). The ideal values of criteria are denoted by $\boldsymbol{q}^* = \{q_1^*, q_2^*, q_3^*, \ldots, q_n^*\}$. In parallel with goal setting, goal state evaluation is constantly being performed in the brain. For the external stimuli, we continuously investigate, measure and attempt to detect any and all current deviations from ideal goal states. This process is called state evaluation. The actual values of criteria are denoted by $\boldsymbol{q} = \{q_1, q_2, q_3, \ldots, q_n\}$. Goal setting and state evaluation are dynamic, interactive processes that are affected by physiological forces, self-suggestion, external information forces, current memory and information-processing capacity (Yu, 1990).

Each stimulus is related to a set of goal functions. When there is an unfavorable deviation of the perceived value from the ideal, each goal function will produce a corresponding level of charge. Take the dynamic route choice problem as example: when users receive information of the real-time traffic jam on the road ahead, their expected travel time and delay will increase if they proceed straight ahead. Furthermore, their driving speed will also be slowed. Thus, the charge of travel time, delay, and driving speed directly related to the real-time information will, as well, increase. The totality of the charge by all goal functions is called the charge structure. The charge structure can change dynamically since, at any point in time, our attention will be drawn to the event that has the greatest influence on charge structure. The difference between the ideal and actual values of each criterion is calculated by $\|\boldsymbol{q}^* - \boldsymbol{q}\|$, which also measures the level of charge. We can summarize the above discussion with Definition 5.1.

Definition 5.1. For any criterion v_i in HD, q_i^* and q_i respectively denote the ideal and actual values of criteria v_i. Furthermore, the level of charge, denoted by Q_i, of v_i is measured by the difference between ideal and actual values

$$Q_i = \|q_i^* - q_i\| \quad \forall i \in [1, n], \tag{11}$$

where $\| \ \|$ is a meaning norm such that $0 \leqslant Q_i \leqslant 1$ for all $v_i \in \mathrm{HD}$.

If $S_t \setminus \mathrm{HD} \neq \phi$, it indicates that the stimulus S_t is unknown or only partially known to our existing HD. This situation implies that S_t contains some elementary criteria outside of the existing discussion universe (HD) for the decision problem E. It also implies that S_t cannot be completely contained in any α-core, no matter how small α is. For example, if road users do not have enough experience about ramp control of the highway, they would not be sure whether everything goes well as they enter the highway. Additionally, if users have no information about the traffic situation of other substitute path, it is very difficult for them to conduct route choice decision-making. Furthermore, even if the decision has been made, the result, such as total travel time, would be not quite desirable. In this case, the decision problem E will be challenging to the decision-maker. However, we notice that what is unknown to one person may possibly be known to someone else; that is, what is a challenging problem to one person may be a fuzzy or routine problem to another. Even so, E is still a challenging problem to all decision-makers. Generally speaking, challenging problems are very difficult to solve unless decision-makers can expand and/or restructure their HD. For simplicity, we will not discuss the case of $S_t \setminus \mathrm{HD} \neq \phi$.

On the other hand, even though what is not in HD may be more important than what is in it, people usually ignore those criteria belonging to $S_t \setminus \mathrm{HD}$. This observation can be attributed to the property of habitual domains that is an indication of the way that people process arriving stimuli. In this context of arriving stimuli, absorption is defined as the possibility that information input will be accepted. A suggestion is more easily accepted if it strikes a consonance in the receiver's memory; therefore, the degree of absorption will depend on a decision-maker's memory, goal setting, state evaluation, charge structure, and charge release. For example, sales advertisements about houses would often publicize such public facilities as swimming pool, exercise facilities, etc. However, this kind of information would be ignored by those who cannot swim or hate sports. Generally speaking, people will actively or progressively learn to accept those ideas, concepts and experiences which can help them reach their goals; that is, they accept the stimuli which are related to individual goal functions. If the external information is not related to a decision-maker's goal setting, charge structure and charge release, it is likely to be rejected. Therefore, it is not necessary to consider criteria outside of the existing discussion universe (HD) of decision problem E.

Given an external stimulus S_t at time t (or stage t) of the problem E, we assume that the arriving stimulus can be broken down into several elementary criteria belonging to HD; that is, $S_t \subseteq \mathrm{HD}$. The corresponding level of charge for each criterion v_i is denoted by Q_i. For $\alpha \in [0, 1]$, the α-core of S_t at time t (or stage t), denoted by S_t^α, is defined as the collection of criteria that can be activated with a level of charge larger than or equal to α. That is,

$$S_t^\alpha = \{v_i \in S_t \cap \mathrm{HD} \mid Q_i \geqslant \alpha\}. \tag{12}$$

S_t^α is the actual domain in a narrow sense at the time t (or stage t) concerned with an external stimulus S_t. Before illustrating the actual domain in a broad sense and the reachable domain, we define the connectivity from the existing domain to a particular criterion. Let us first give the following axiomatic definition:

Definition 5.2. The mapping $\mathscr{C}_t : \sigma(\mathrm{HD}) \times \mathrm{HD} \to C_{t^1}$ is called a connectivity function of criteria with subsets of *HD* at time t (or stage t) if it satisfies the following axioms:
(i) $\mathscr{C} \geqslant 0$ (non-negativity);
(ii) $\mathscr{C}_t(\Lambda_j, v_i) = 1, \forall v_i \in \Lambda_j$;
(iii) $\mathscr{C}_t(\Lambda_j, v_i) \leqslant \mathscr{C}_t(\Lambda_k, v_i), \forall v_i \in \mathrm{HD}$; if $\Lambda_j \subseteq \Lambda_k$ (monotonicity), where $\sigma(\mathrm{HD})$ is a σ-algebra generated by HD.

As mentioned before, S_t^α is the existing domain which can represent a criterion set activated by an external stimulus at time t (or stage t) for the problem E; that is, S_t^α is the collection of criteria that can be activated with a level of charge larger than or equal to α. Now, since $\mathscr{C}_t$ is a connectivity function, we can denote $\mathscr{C}_t(S_t^\alpha, v_i)$ as the connectivity of a criterion v_i with the existing domain (the actual domain in a narrow sense) S_t^α. When there is no confusion, we treat a connectivity function of criteria with the actual domain as a connectivity function. That is, a function $C_t(S_t^\alpha, v_i)$ defined on $\sigma(\mathrm{HD}) \times \mathrm{HD}$ at time t (or stage t) is called a connectivity function. The proof of Proposition 5.3 demonstrates that $\mathscr{C}_t(S_t^\alpha, v_i) \in [0, 1]$.

Proposition 5.3. *$\mathscr{C}_t(S_t^\alpha, v_i)$ is bounded by 1.*

Proof. For any v_i in HD and any S_t^α in σ(HD), we can obtain the following inequality according to the monotonicity axiom of $\mathscr{C}_t$ and the condition that $S_t^\alpha \subseteq S_t \subseteq$ HD,

$$\mathscr{C}_t(S_t^\alpha, v_t) \leqslant \mathscr{C}_t(S_t, v_i) \leqslant \mathscr{C}_t(\mathrm{HD}, v_i) = 1. \qquad \square$$

As for the relationship between the connectivity function of a criterion to a criterion and that of a domain to a criterion, the latter can be considered to be an extension of the former. To make this concept clearer, let us give Proposition 5.4.

Proposition 5.4. *(1) Let* HD *be finite and C_t^* be the connectivity index with $C_t^*(v_i, v_j)$ the strongest connectivity from criteria v_i* to *v_j. Given an external stimulus S_t at time t (or stage t), the* α-core *of S_t is denoted by S_t^α, where $S_t^\alpha \in \sigma$*(HD). *For any criterion v_j in* HD, *if we define the connectivity of v_j with the existing domain S_t^α as*

$$\mathscr{C}_t(S_t^\alpha, v_j) = \max\{C_t^*(v_i, v_j) \mid v_i \in S_t^\alpha\}, \tag{13}$$

then $\mathscr{C}_t$ satisfies the axioms as defined in Definition 5.2.

(2) The power set of HD *is indicated by* P(HD). *Assume that* σ(HD) = P(HD). *Given that $\mathscr{C}_t$ is a connectivity function of criteria with the existing domain as defined in Definition 5.2 then*

$$C_t^*(v_i, v_j) = \mathscr{C}_t(\{v_i\}, v_j) \tag{14}$$

is the connectivity index starting from v_i arriving at v_j as described in Remark 5.1.

Proof. (1) According to Proposition 5.2, we know that the connectivity index C_t^* is given by

$$C_t^*(v_i, v_j) + \max\{\min\{C_t(x,y) \| \ (x,y) \in Path(v_i, v_j)\} \mid Path(v_i, v_j)\},$$

where $C_t(v_i, v_j)$ is the connectivity from v_i to v_j.

(i) For any $v_i \in S_t^\alpha$ and any $v_j \in$ HD, the non-negativity of $\mathscr{C}_t$ is implied by the relationship

$$\mathscr{C}_t(S_t^\alpha, v_j) \geqslant C_t^*(v_i, v_j) \geqslant C_t(v_i, v_j) \geqslant 0.$$

(ii) If $v_j \in S_t^\alpha$, then

$$\mathscr{C}_t(S_t^\alpha, v_j) \geqslant C_t^*(v_j, v_j) = 1.$$

By Proposition 5.1, we know that $\mathscr{C}_t(S_t^\alpha, v_j) \leqslant 1$. By combining this relationship with the above condition, we can see that $\mathscr{C}_t(\{S_t^\alpha, v_j) = 1$.

(iii) Let $S_t^\alpha \in \sigma$(HD); if $S_t^\alpha \subseteq S_t^\beta$, then $\forall v_j \in$ HD we have the following inequalities:

$$\max\{C_t^*(v_i, v_j) | v_i \in S_t^\alpha\} \leqslant \max\left\{C_t^*(v_i, v_j) | v_i \in S_t^\beta\right\}, \mathscr{C}_t(S_t^\alpha, v_j) \leqslant \mathscr{C}_t, (S_t^\beta, v_j).$$

It is obvious that there exists monotonicity of $\mathscr{C}_t$. Thus (1) is proved.

(2) (i) By the non-negativity of C_t and Proposition 5.3. we see that $0 \leqslant \mathscr{C}_t(\{v_i\}, v_j) \leqslant 1$ corresponding $0 \leqslant C_t^*(v)_i, v_j) \leqslant 1$.

(ii) If $v_i = v_j$, then $v_j \in \{v_i\}$ and $C_t^*(v_i, v_j) = \mathscr{C}_t(\{v_i\}, v_j) = 1$. Consequently, by (i) and (ii), $C_t^*(v_i, v_j)$ satisfies the following axioms: (i) $C_t^*(v_i, v_j) \in [0, 1]$; (ii) $C_t^*(v_i, v_j) = 1$ if $v_i = v_j$. This completes the proof. □

Remark 5.2. For any $v_i \in S_t^\alpha$ and any $v_j \in$ HD, $Path^*(v_i, v_j)$ is the optimal path from v_i to v_j such that $\mathscr{C}_t(S_t^\alpha, v_i)$ is a maximum. In other words, $Path^*(v_i, v_j)$ is the strongest path because its connectivity index is the maximal of all possible paths.

According to the analogy and association hypothesis, we can conclude that new things can be more easily learned if they are similar to some things that are already known. Because people associate arriving stimuli with preexisting memory, arriving stimuli tend to be initially treated as either positive or negative. If the arriving stimulus is perceived to be relevant to a decision-makers goal function, then it is more carefully examined and composed to fit preexisting codes. Also, we may say that if a newly arriving stimulus is very similar to a preexisting code, it will be quickly processed in a positive or favorable way. Conversely, a stimulus that is perceived as irrelevant or quite dissimilar to preexisting codes will be filtered or ignored. It should be noted that frequently repeated events will have a stronger influence on analogy and association. However, those events preexisting in weak codes, which are stored in remote areas of the brain, will have little impact on the analogy and association process. Therefore, we must first specify the influential domain from the preexisting memory through external stimuli. We can restrict the neighborhood of the actual domain in the narrow sense to be the reachable domain. In order to figure out the neighborhood of the actual domain in this case, we facilitate our discussion by using the connectivity of criteria with the existing domain.

Definition 5.3. Given a connectivity network $G(\mathrm{HD}, C_t)$ and an external stimulus S_t at time t (or stage t), the α-core of S_t is denoted by S_t^α and $\mathscr{C}_t$ is a connectivity function of the criteria with the existing domain. Let 2^{HD} denote the collection of all non-empty compact subsets of HD. The ε-neighborhood of S_t^α for $S_t^\alpha \in 2^{\mathrm{HD}}$ is defined by $N_t(S_t^\alpha, \varepsilon)$

$$N_t(S_t^\alpha, \varepsilon) = \{v_j \in \mathrm{HD} \setminus S_t^\alpha \mid \exists v_i \in S_t^\alpha, \ni \mathscr{C}_t(S_t^\alpha, v_j) \geqslant \varepsilon\}, \quad \forall S_t^\alpha \in 2^{\mathrm{HD}}. \tag{15}$$

S_t^α can be considered as the actual domain that contains the set of criteria that are actually activated. Moreover, $N_t(S_t^\alpha, \varepsilon)$ represents the reachable domain that contains the collection of criteria that are reachable from the existing domain through external stimuli. Therefore, $S_t^\alpha \cup N_t, (S_t^\alpha, \varepsilon)$ is the actual domain in a broad sense, and we call it the *generalized current domain.*

When external stimuli are repeated, the corresponding circuit patterns will be reinforced and strengthened. Furthermore, the stronger the circuit patterns become, the more easily the corresponding circuit patterns are retrieved in the learning processes. Therefore, it seems reasonable to assume that the connection between a pair of criteria in the actual domain and its ε-neighborhood will be reinforced as the learning process progresses. In other words, the connectivity between pairs of criteria in the generalized current domain will increase. Hence, the connectivity function C_t must be updated after each learning iteration.

The uncertainty that arises from human thought processes and the randomness associated with experiments is often confused by social scientists (Kaufmann and Gupta, 1991). Some of the data obtained in this manner are hybrid; that is, their components are not homogeneous, but a blend of precise and fuzzy information. Furthermore, a fuzzy relation, such as the connectivity function, is not a measurement. That is, the connectivity function is a subjective valuation assigned by one or more human operators. To simplify matters, we suppose that $C_t(v_i, v_j)$ is a continuous random variable, uniformly distributed within the interval

$$\left[\left(\bar{C}(v_i, v_j)\right)^{1/\beta}, \left(\overline{C}(v_i, v_j)\right)^{\beta}\right],$$

where $\overline{C}(v_i, v_j)$ is the mean of $C_t(v_i, v_j)$. β is called the determinate index and its value is in the unit interval [0,1]. β characterizes the degree of certainty since the higher the β value is the less change is performed by the connectivity function. That is, when β approaches 1, the connectivity is rather stable. On the other hand, when β approaches 0, we do not have sufficient evidence to point out the exact value of the connectivity function. Let $U(0,1)$ represent a continuous random variable that is uniformly distributed over the interval [0,1]. From simple proportionality, we can write

$$C_t(v_i, v_j) = \left(\overline{C}(v_i, v_j)\right)^{1/\beta} + \left[\left(\overline{C}(v_i, v_j)\right)^{\beta} - \left(\overline{C}(v_i, v_j)\right)^{1/\beta}\right] U(0, 1). \tag{16}$$

Thus, it is very simple to generate $C_t(v_i, v_j)$ from a given $U(0,1)$ provided the lower bound and upper bound are known. In order to reflect the fact that the connectivity between each pair of criteria is enhanced through the learning processes, we define an index parameter $I_t(v_i, v_j)$ for each pair (v_i, v_j) belonging to HD at time t (or stage t) and a concentration parameter δ as follows. The initial values of I for pairs of criteria are set to zero. If v_i is activated when v_j is presented to the input stimuli, the value of I increases by 1. The concentration parameter, δ, represents the change in size of the definition domain, and $0 < \delta < 1$. Consequently, the connectivity function is calculated within the adjustment interval

$$\left[\left(\overline{C}(v_i, v_j) + I_t(v_i, v_j)\delta\right)^{1/\beta}, \left(\overline{C}(v_i, v_j) + I_t(v_i, v_j)\delta\right)^{\beta}\right]$$

and now $C_t(v_i, v_j)$ is given by

$$C_t(v_i, v_j) = \min\Big\{1, \left(\overline{C}(v_i, v_j) + I_t(v_i, v_j)\delta\right)^{1/\beta} + \Big[\left(\overline{C}(v_i, v_j) + I_t(v_i, v_j)\delta\right)^{\beta} - \left(\overline{C}(v_i, v_j) + I_t(v_i, v_j)\delta\right)^{1/\beta}\Big] U(0,1)\Big\}. \quad (17)$$

To avoid the condition where the connectivity exceeds 1, we utilize a "min" operator. That is, we use the index parameter, I, and the concentration parameter, δ, to indicate the reinforcement change of circuit patterns.

6. Weight-assessing method with habitual domains

6.1. Architecture

The architecture of the weight-assessing method with habitual domains is shown in Fig. 1. All elements within the network are fully interconnected.

The weight-assessing method with habitual domains is based on competitive learning. As we described above, the elementary criteria are the basic elements of our discussion universe. Each criterion has routes that connect it to the neighboring ideas. In the presence of external stimuli, some criteria can be "fired" and "lit up" sequentially through the learning processes, while others will remain "dark". The external stimuli are registered and processed using the circuit patterns or sequences of circuit patterns in our discussion universe. In other words, external stimuli are encoded as digraphs, routes or paths of lit ideas, and activation occurs when attention is paid. This in turn triggers the appropriate interconnecting arcs to fire or the appropriate criteria to discharge. Therefore, the output ideas of the digraph compete with each other and only some of them are activated or fired at any one time. Such being the case, our weight-assessing method is based on the concept of competitive learning.

In the most general case, competitive learning belongs to unsupervised learning (Fausett, 1994). In unsupervised learning, the network is presented with a set of training patterns, but is not given a target answer for each input pattern. Thus, we can modify the weights of the network without specifying the desired output for any of the input patterns.

In competitive learning, the output ideas of the network compete among themselves to be active (or fired). Competitive learning begins with a random arrangement of weights and gives all output ideas a chance to compete. It also limits the strength of the weights. Each criterion has a temporally fixed amount of weight. The weights are limited to values between 0 and 1; that is,

$$w_t(v_i) \geqslant 0 \quad \text{for each } v_i, \sum_{i=1}^{n} w_t(v_i) = 1, \quad (18)$$

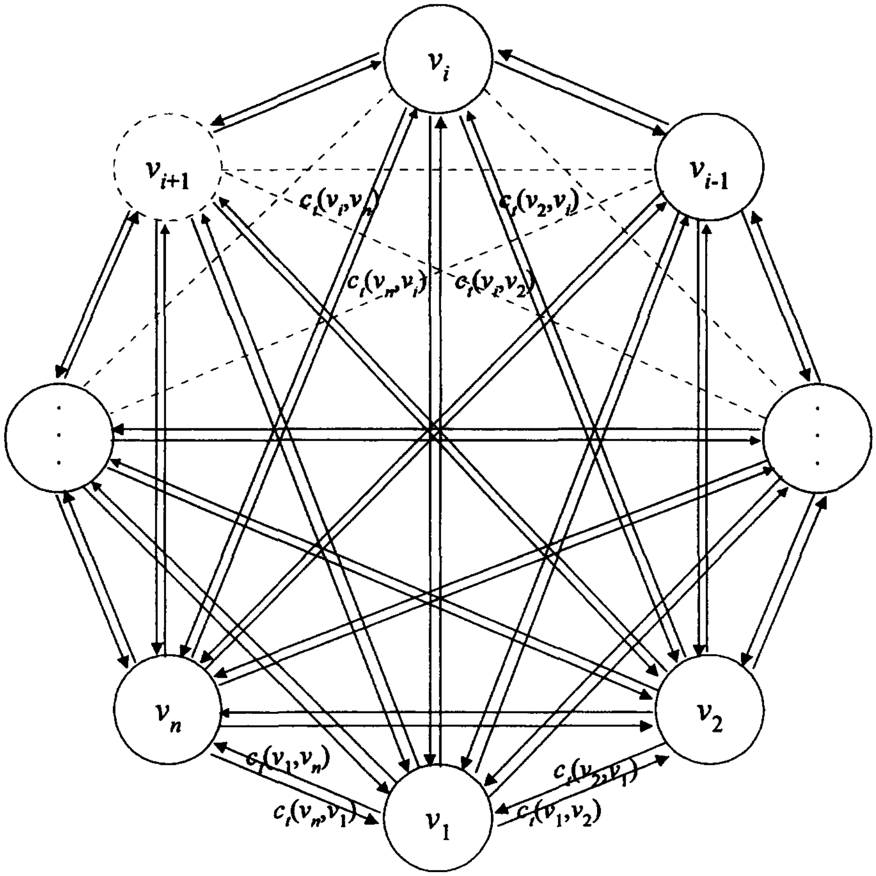

Fig. 1. Architecture of weight-assessing method with habitual domains.

where $w_t(v_i)$ is the weight of the criterion v_i (competitive layer) from the input stimulus S_t.

An internal mechanism creates a competition among the ideas for the right to respond to a given subset of input stimuli, so that only one output idea, or only one idea per group, is active (i.e., "on") at a time. The idea that wins the competition is called the "winner-take-all" idea (Haykin, 1994). In this study, the criteria in the generalized current domain are chosen as the winners during the learning process. This is different from the traditional concept of winner-take-all because there is not necessarily only one criterion in the winner group when the competition is completed. This form of competition among a group of criteria is called *generalized winner-take-all*, and our weight-assessing method performs a generalized winner-take-all competition.

Let S_t^α denote the actual domain of all winning criteria, and its ε-neighborhood (i.e., the reachable domain) be $N_t(S_t^\alpha, \varepsilon)$. The output signals of the generalized current domain are set equal to one; the output signals of all the criteria that lose the competition are set equal to zero. The output signal is also called the index parameter I_t. We use the winning set and its neighborhood to update the weights of the network. Then, we can form a new weight vector that is a linear combination of the old weight vector and the current input vector. Weight corrections are accumulated over an entire epoch of training patterns (i.e., batch updating). This updating procedure has a smoothing effect on the correction terms. The learning rule of the weight correction is thus

$$w_{t+1}(v_i) = w_t(v_i) + \Delta w_t(v_t). \tag{19}$$

We can define the adjustment factor η_i for each criterion v_i through the level of charge and the connectivity function. For each $v_i \in N_t(S_t^\alpha, \varepsilon)$, v_k is the precedent criterion of v_i so that

$$C_t^*(v_k, v_i) = \max\{C_t^*(v_j, v_i) \mid v_j \in S_t^\alpha\}.$$

The level of charge of v_k is denoted by Q_k. The adjustment factor η_i is computed by:

$$\eta_i = \begin{cases} Q_i & \text{if } v_i \in S_t^\alpha, \\ Q_k \mathscr{C}_t(S_t^\alpha, v_i) & \text{if } v_i \in N_t(S_t^\alpha, \varepsilon), \\ 0 & \text{otherwise.} \end{cases} \tag{20}$$

Assume that the input stimulus contains a set of p vectors. Let $\overline{\eta_1}$ be the average of all η_i's during the particular learning iteration; then $\overline{\eta_i}$ is

$$\overline{\eta_i} = \frac{\sum_p \eta_i}{p}. \tag{21}$$

Using the average adjustment factor, we can obtain the change $\Delta w_t(v_i)$

$$\Delta w_t(v_i) = \xi_t \times \left[\left(\frac{\overline{\eta_i}}{\sum_j \overline{\eta_j}}\right) - w_t(v_i)\right], \tag{22}$$

where ξ_t is the learning-rate parameter and its value is chosen by users. Note that the values of ξ_t' s must be between 0 and 1.

Remark 6.1. The learning rule of the weight correction, as described by (19)–(22), guarantees that the sum of the weights for all criteria in our discussion universe is always equal to one.

Proof. For all criteria in our discussion universe HD at time t (or stage t), the sum of weight changes is equal to zero; that is,

$$\begin{aligned} \sum_i \Delta w_t(v_i) &= \xi_t \times \left[\sum_i \left(\frac{\overline{\eta_i}}{\sum_j \overline{\eta_j}}\right) - \sum_i w_t(v_i)\right], \\ &= \xi_t \times [1 - 1] = 0. \end{aligned} \tag{22}$$

Thus, the sum of the weights for all criteria in HD at time $t + 1$ (or stage $t + 1$) is

$$\sum_i w_{t+1}(v_i) = \sum_i w_t(v_i) + \sum_i \Delta w_t(v_i) = 1 + 0 = 1. \quad \square$$

We now want to discuss equilibrium states of the routine and mixed routine systems. From Proposition 6.1, the weights of criteria clearly are only related to the level of charge and the connectivities with the existing domain in the equilibrium state.

Proposition 6.1. *For routine and mixed routine problems, the weight of criterion v_i is only related to the level of charge and the connectivity of v_i with the existing domain in the equilibrium state. Moreover, ceteris paribus, the weight of v_i is also fixed and stable.*

Proof. Let h_k be the probability that stimulus S_k is presented on any trail and $g_k(v_i)$ be the probability that criterion v_i wins (i.e., is in the generalized current domain) when stimulus S_k is presented. We consider the case in which

$$\sum_k \Delta w_t(v_i) h_k g_k(v_i) = 0,$$

that is, the case in which the average change in the weights is zero. We refer to such a state as the equilibrium state (Rumelhart and McClelland, 1986). Thus, using the learning rule and averaging over all of the stimulus patterns, we have

$$\xi_t \sum_k \left[\frac{\overline{\eta_i}}{\sum_j \eta_j} h_k g_k(v_i) \right] - \xi_t \sum_k [w_t(v_i) h_k g_k(v_i)] = 0,$$

and thus

$$w_t(v_i) = \frac{\frac{\overline{\eta_i}}{\sum_j \overline{\eta_j}} \sum_k h_k g_k(v_i)}{\sum_k h_k g_k(v_i)} = \frac{\overline{\eta_i}}{\sum_j \overline{\eta_j}}.$$

The average adjustment factor for criterion v_i is calculated using the charge and the connectivity function, so the weight of v_i is only related to the level of charge and the connectivity of v_i in the existing domain. Let us consider the situation in which the level of charge and the connectivity corresponding to the input stimuli remain stable. In this case, the average adjustment factor does not make significant changes; thus, the weight is a constant. Whenever the system is in the state in which, on average, the weights are not changing, we say that the system has reached an equilibrium state. When this happens, the system always responds the same way to a particular stimulus pattern. Note that the equilibrium state only holds for routine and mixed routine problems.

According to Yu (1990), a routine problem means that the needed idea set (i.e., the truly needed competence set) is well known, and the decision-maker has mastered the set. Because the needed idea set is well-known, ceteris paribus, the difference between the ideal and actual values of each decision criterion does not change. Furthermore, the decision-maker has acquired and mastered the truly needed set, so his connectivity network is quite stable. Since a mixed routine problem consists of a number of routine subproblems, the discussion concerned with mixed routine problem is similar. Therefore, the weights of criteria remain unchanged in response to any single stimulus in the equilibrium state in routine and mixed routine problem. □

From Proposition 6.1, we know that for routine and mixed routine problems, the weights of criteria are stable in the equilibrium state. However, it is possible that weights will be pushed out of equilibrium by an unfortunate sequence of stimuli. In this case, the system can move toward a new equilibrium state (or possibly back to a previous one).

For a fuzzy problem, the ideas, concepts and skills needed to successfully solve the decision problem are roughly, but not clearly, known. This implies that the decision-maker has not mastered the ideas and skills necessary for solving these problems (Yu, 1990). Such being the case, the connectivity network of the decision-maker does not reach a stable state. However, we can obtain a *temporally* fixed amount of weight through the algorithm introduced in Section 6.2 even though the weight does not remain stable throughout the decision-making process.

The architecture and algorithm introduced in the next section for the connectivity network can be used in routine, mixed routine, and fuzzy problems.

6.2. Algorithm

The algorithm given here is suitable for routine, mixed routine, or fuzzy problems. Note that weights for routine or mixed routine problems are rather stable, but weights for fuzzy problems are in a stable state only when there has been a convergence that satisfied the stopping rule. When the stopping condition is false, we cannot obtain the approximate amounts of weights.

Step 0. Initialize weights $w_t(v_i)$ for each v_i belonging to HD so that $\sum_{i=1}^{n} w_t(v_i) = 1$.
Initialize the continuous random variate, $U(0,1)$, that is uniformly distributed over the interval [0,1].
Initialize the index parameters $I_t(v_i, v_j) = 0$ for each pair (v_i, v_j) belonging to digraph G.
Initialize the ε-neighborhoods $N_t(S_t^\alpha, \varepsilon) = \phi$ for all $S_t^\alpha \in 2^{\mathrm{HD}}$.
Obtain the ideal values of criteria $\boldsymbol{q}^*$, the actual values of criteria $\boldsymbol{q}$, and the initial mean connectivity matrix $\overline{C}_t = [\overline{C}_t(v_i, v_j)]$ for each pair (v_i, v_j) belonging to the digraph G through the questionnaire survey.
Set the concentration parameter $\delta, 0 < \delta < 1$.
Set the determinate index $\beta, 0 \leqslant \beta \leqslant 1$.
Set the learning rate parameters ξ_t. Set the threshold parameter $\alpha, 0 \leqslant \alpha \leqslant 1$.

Step 1. Compute the initial connectivity for each pair (v_i, v_j) belonging to digraph G.

$$C_t(v_i, v_j) = \left(\overline{C}(v_i, v_j)\right)^{1/\beta} + \left[\left(\overline{C}(v_i, v_j)\right)^{\beta} - \left(\overline{C}(v_i, v_j)\right)^{1/\beta}\right] U(0,1).$$

Step 2. When the stopping condition is false, do steps 3–13.

Step 3. For each stimulus vector S_t, do steps 4–9.

Step 4. Specify the actual domain in the narrow sense, S_t^α:

$$S_t^\alpha = \{v_i \in S_t \cap \mathrm{HD} \mid Q_i \geqslant \alpha\},$$

where $Q_i = \|q_i^* - q_i\|, \quad \forall i \in [1, n]$.

Step 5. For each criterion $v_j \in \mathrm{HD} \setminus S_t^\alpha$, compute the connectivity of v_j with the existing domain S_t^α:

$$\mathscr{C}_t(S_t^\alpha, v_j) = \max\{C_t^*(v_i, v_j) \mid v_i \in S_t^\alpha\},$$

where $C_t^*(v_i, v_j) = \max\{\min\{C_t(x,y) \mid (x,y) \in Path(v_i, v_j)\} \mid Path(v_i, v_j)\}$.

Step 6. Find the ε-neighborhood of S_t^α:

$$N_t(S_t^\alpha, \varepsilon) = \{v_j \in \mathrm{HD} \setminus S_t^\alpha \mid \exists v_i \in S_t^\alpha, \ni \mathscr{C}_t(S_t^\alpha, v_j) \geqslant \varepsilon\}, \qquad \forall S_t^\alpha \in 2^{\mathrm{HD}}.$$

Step 7. For each $v_i \in S_t^\alpha, v_j \in N_t(S_t^\alpha, \varepsilon)$, $Path^*$ (v_i, v_j) is the optimal path from v_i to v_j such that $\mathscr{C}_t(S_t^\alpha, v_i)$ is a maximum.
Update the index parameters for all criteria x, y within $Path^*(v_i, v_j)$:

$$I_t^{(\mathrm{new})}(x,y) = I_t^{(\mathrm{old})}(x,y) + 1.$$

Step 8. For each v_i, compute the adjustment factor η_i:

$$\eta_i = \begin{cases} Q_i & \text{if } v_i \in S_t^\alpha, \\ Q_k \mathscr{C}_t(S_t^\alpha, v_i) & \text{if } v_i \in N_t(S_t^\alpha, \varepsilon), \\ 0 & \text{otherwise}. \end{cases}$$

where v_k is the precedent criterion of v_i.

Step 9. For all $v_i, v_j \in S_t^\alpha \cup N_t(S_t^\alpha, \varepsilon)$, update the connectivity from v_i to v_j:

$$C_t^{(\mathrm{new})}(v_i, v_j) = \min\left\{1, \left(C_t^{(\mathrm{old})}(v_i, v_j) + I_t^{(\mathrm{new})}(v_i, v_j)\delta\right)^{1/\beta} + \left[\left(C_t^{(\mathrm{old})}(v_i, v_j) + I_t^{(\mathrm{new})}(v_i, v_j)\delta\right)^{\beta} - \left(C_t^{(\mathrm{old})}(v_i, v_j) + I_t^{(\mathrm{new})}(v_i, v_j)\delta\right)^{1/\beta}\right] U(0,1)\right\}.$$

Step 10. Let $\overline{\eta_i}$ be the average of all η_i's.

$$\overline{\eta_i} = \frac{\sum_p \eta_i}{p}.$$

Step 11. Compute the weights $w_{t+1}(v_i)$ for each criterion v_i within HD at stage $t+1$:

$$w_{t+1}(v_i) = w_t(v_i) + \Delta w_t(v_i),$$

where $\Delta w_t(v_i) = \xi_t \times \left[\left(\frac{\overline{\eta_i}}{\sum_j \overline{\eta_j}}\right) - w_t(v_i)\right].$

Step 12. Update the learning rate.

Step 13. Test for the stopping condition:

Let $\partial w_t(v_i) = \frac{|w_{t+1}(v_i) - w_t(v_i)|}{w_t(v_i)}.$

If $\max_i \partial w_t(v_i)$ is smaller than a specified tolerance $\forall i \in [1, n]$, then stop; otherwise, continue.

6.3. Empirical study

In this section, we will discuss an application that employs the weight-assessing method with habitual domains. In the following, we will introduce the background of the problem and the questionnaire content of our empirical study.

In Taiwan, a great number of people use motorcycles as their primary commuting mode because of low car-ownership and low car-operating costs, comparatively small size, high mobility, and ease of operation. Take Taipei City, for instance: registered motorcycles amounted to 54% of all motor vehicles at the end of 1995, and the percentage of growth has risen 10% in the last two years. Furthermore, such statistics do not account for the number of motorcycles coming from other cities, which has not yet been taken into account. The massive amount of mixed traffic flow created by motorcycles has had severe effects in urban areas. And it has also posed a serious threat to safety and driving in urban traffic, not to mention air and noise pollution as well as the high risk of damages resulting from the inferior stability and fragile structure of motorcycles. Therefore, how to promote public transportation in motorcycle users habitual domains becomes more and more important for traffic authorities.

This empirical study is undertaken to test the applicability of our weight-assessing method through questionnaires to motorcycle travelers in Taipei. We investigated the decision attributes as well as the grade of importance considered by motorcycle travelers during their mode choice process. It is possible to encourage motorcycle users to shift to public transit if the latter possesses those attributes. Such being the case, the traffic authorities can realize how to attract motorcycle travelers to use transit under the specific attribute stimulus. Therefore, the empirical results for motorcycle travelers will help to set up transportation marketing strategies and corresponding policies favoring public transportation usage in the future.

A two-stage approach was used for the empirical research of motorcycle users in Taipei. We first obtained the influential factors of mode choice for motorcycle users through the first stage questionnaire with open questions. Nine criteria of high membership are selected, which includes walking time (v_1), waiting time (v_2), in-vehicle time (v_3), transfer time (v_4), accessibility (v_5), travel cost (v_6), punctual arrival (v_7), difficulty of parking (v_8), and degree of traffic jam (v_9). Then the second stage questionnaire will be subsequently conducted. The population is the people living in the administrative zones comprising Taipei City. We employ the stratified sampling, and the criterion of proportion allocation is according to the subpopulation of each zone. A total of 160 questionnaires were sent out and 155 valid copies retrieved.

The main content of investigation includes the connectivity between any two criteria. The connectivity means the degree of ease from one criterion to another by analogy and association. The mean value of the connectivity $C(v_i, v_j)$ for all interviewees will be denoted as $\overline{C}(v_i, v_j)$, then we can construct the connectivity network. Because it is difficult to specify the ideal values of all criteria, we intend to use the tolerance value of each criterion instead of the ideal value. Furthermore, we use the difference between the tolerance and actual values of each criterion to specify the level of charge, and [0,1] is the range of the norm for the level of charge. When the actual value of a criterion is equal to or has exceeded the tolerance value, there is an unfavorable deviation of the perceived value from the goal; the corresponding level of charge then is 1. When the actual value of a criterion is smaller than tolerance value, the corresponding level of charge approximates to 1 if the actual value is close to the tolerance value. The collection of criteria whose levels of charge have exceeded a threshold value is called the significant stimulus set.

In view of the nine criteria obtained from the first stage investigation, we asked interviewees the tolerance values of above criteria according to their commuting experiences. It is noted that accessibility is measured by the distance between the destination and the place that their mode can reach. Furthermore, punctual arrival is expressed by the delay time, and difficulty of parking means the searching or waiting time for parking space. Then the tolerance values obtained will be compared to the actual values of each criterion so as to find out what criteria might create change to the charge structure of motorcycle users, and users attentions will be directed to those criteria.

This study assumes that motorcycle travelers choice set includes all traffic modes. The mean values of criteria for each mode according to the past investigation data would be taken to be the actual values for the model (Chen et al., 1997). The in-vehicle time is found by taking the walking distance out of the average origin-to-destination distance, then dividing it by the speed of the mode. Additionally, the delay time in Taipei amounts to 41.2% of the travel time as found from the report of optimal control on the entire transportation system within Taipei City; thus, the delay time should be resolved by having such percentage multiplied by the in-vehicle time. And the travel cost of the train can be found based on the pricing structure stipulated by the Taiwan Railway Administration.

Now we use nine criteria and six input stimuli to present the following illustration. The input data include the connectivity matrix of all criteria ($\overline{C}$), the tolerable values of criteria ($\boldsymbol{q}^*$), and the actual values of criteria for each input stimulus ($\boldsymbol{q}$). Suppose the learning rate (geometric decrease) is

$$\xi_0 = 0.6, \quad \xi_{t+1} = 0.5\xi_i,$$

Step 0. We obtain the values of, $\boldsymbol{q}*$, and $\boldsymbol{q}$ through questionnaire survey.

II I	v_1 Walking time	v_2 Waiting time	v_3 In-vehicle time	v_4 Transfer time	v_5 Accessibility	v_6 Travel cost	v_7 punctual arrival	v_8 Difficulty of parking	v_9 Degree of traffic jam
v_1	1.000	0.419	0.573	0.406	0.672	0.209	0.590	0.545	0.303
v_2	0.406	1.000	0.457	0.542	0.320	0.307	0.788	0.325	0.745
v_3	0.494	0.461	1.000	0.437	0.525	0.465	0.818	0.424	0.805
v_4	0.518	0.546	0.412	1.000	0.417	0.291	0.664	0.275	0.566
$\overline{C} = v_5$	0.708	0.406	0.527	0.432	1.000	0.401	0.420	0.301	0.265
v_6	0.234	0.385	0.510	0.293	0.417	1.000	0.252	0.302	0.394
v_7	0.721	0.750	0.786	0.609	0.287	0.318	1.000	0.581	0.789
v_8	0.491	0.203	0.446	0.225	0.274	0.255	0.610	1.000	0.558
v_9	0.259	0.553	0.680	0.462	0.297	0.332	0.725	0.544	1.000

The above table indicates the connectivity between criteria of total samples, and the connectivity stands for the degree of difficulty being associated from criterion I to criterion II. The connectivity from the in-vehicle time to punctual arrival ranks the highest (0.818), followed by from the in-vehicle time to traffic jam (0.805), from punctual arrival to traffic jam (0.789), and from the waiting time to punctual arrival (0.788). On the other hand, those of lower connectivity are from the association of the walking time to travel cost (0.209), and from difficulty of parking to the transfer time (0.225).

	v_1 (min)	v_2 (min)	v_3 (min)	v_4 (min)	v_5 (m)	v_6 (NT$)	v_7 (min)	v_8 (min)	v_9 (km/hr)
$q^* =$	12.60	14.89	34.78	16.17	803.04	45.23	14.23	4.33	24.87

In terms of the tolerance values of each criterion, the tolerable walking time is 12.60 min, waiting time is 14.89 min, in-vehicle time is 34.78 min, transfer time is 16.17 min, the distance between the destination and the place which the mode can reach would be 803.04 m, the tolerable travel cost is no more than 45.23 NT dollars, the delay time is no more than 14.23 min, and the speed of the mode is no less than 24.87 km/h. The following table shows the actual values of each criterion for each mode, including motorcycle, private vehicle, bus, train, taxi, and commuting bus.

	v_1(min)	v_2(min)	v_3(min)	v_4(min)	v_5(m)	v_6(NT$)	v_7(min)	v_8(min)	v_9(km/hr)
Motorcycle	2.46	–	34.25	–	113. 53	0.67	14.11	1.30	14.10
Private vehicle	3.24	–	43.90	–	149.53	2.32	18.09	7.90	11.00
$q =$ bus	11.38	10.48	37.73	8.50	525.23	1.70	15.54	–	12.80
Train	11.30	6.46	14.95	16.10	521.50	2.04	3.76	–	32.30
Taxi	0.69	3.69	43.90	–	31.85	14.21	18.09	–	11.00
Commutingbus	7.02	8.66	43.90	–	324.00	1.26	18.09	–	11.00

Suppose the intrinsic preference structure of criteria is unknown; initialize the weight $w_t(v_i)$ for each criterion v_i to be 0.1111 and $\sum_{i=1}^{9} w_t(v_i) = 1$.

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$W =$	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111

Set the concentration parameter $\delta = 0.0001$.
Set the determinate index $\beta = 0.5$.
Set the threshold parameter $\alpha = 0.5$.
Initialize the learning rate:

$$\xi_0 = 0.6.$$

Suppose the threshold value for an ε-neighborhood of S_t^α is

$$\varepsilon = 0.7.$$

Step 1. Compute the initial connectivity for each pair (v_i, v_j).

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
v_1	1.000	0.539	0.440	0.489	0.456	0.304	0.416	0.518	0.108
v_2	0.519	1.000	0.491	0.636	0.389	0.528	0.753	0.415	0.685
v_3	0.424	0.344	1.000	0.353	0.467	0.366	0.812	0.397	0.686
v_4	0.399	0.514	0.375	1.000	0.202	0.357	0.716	0.300	0.512
$C_t = v_5$	0.767	0.455	0.395	0.282	1.000	0.407	0.232	0.234	0.256
v_6	0.154	0.230	0.704	0.378	0.335	1.000	0.205	0.487	0.350
v_7	0.826	0.644	0.707	0.547	0.475	0.342	1.000	0.692	0.826
v_8	0.326	0.218	0.584	0.200	0.494	0.327	0.426	1.000	0.414
v_9	0.111	0.335	0.633	0.479	0.307	0.539	0.605	0.359	1.000

Step 2. Because the stopping condition is false, we do steps 3–13. Let $t = 1$ represent the first iteration.
Step 3. For the first stimulus vector, do steps 4–9.

First	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$S_t =$	2.46	–	34.25	–	113.53	0.67	14.11	1.30	14.10

Step 4. Calculate the level of charge Q_i of each v_i.

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$Q_i =$	0.80	0.00	0.02	0.00	0.86	0.99	0.01	0.70	0.43

Specify the actual domain in the narrow sense, S_t^α:

$$S_t^\alpha = \{1, 5, 6, 8\}$$

Step 5. Compute the connectivity of v_j with the existing domain S_t^α:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
v_1	1.000	0.539	0.539	0.539	0.494	0.539	0.539	0.539	0.539
v_2	0.753	1.000	0.707	0.636	0.494	0.539	0.753	0.692	0.753
v_3	0.812	0.644	1.000	0.636	0.494	0.539	0.812	0.692	0.812
v_4	0.716	0.644	0.707	1.000	0.494	0.539	0.716	0.692	0.716
$C_t^* = v_5$	0.767	0.539	0.539	0.539	1.000	0.539	0.539	0.539	0.539
v_6	0.704	0.644	0.704	0.636	0.494	1.000	0.704	0.692	0.704
v_7	0.826	0.644	0.707	0.636	0.494	0.539	1.000	0.692	0.826
v_8	0.584	0.584	0.584	0.584	0.494	0.539	0.584	1.000	0.584
v_9	0.633	0.633	0.633	0.633	0.494	0.539	0.633	0.633	1.000

$C_t^*(v_i, v_j)$	HD\ S_t^α								
	–	0.539	0.539	0.539	–	–	0.539	–	0.539
	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–
	–	–	–	–	–	–	–	–	–

S_t^α	–	0.539	0.539	0.539	–	–	0.539	–	0.539
	–	0.644	0.704	0.636	–	–	0.704	–	0.704
	–	–	–	–	–	–	–	–	–
	–	0.584	0.584	0.584	–	–	0.584	–	0.584
	–	–	–	–	–	–	–	–	–
$\mathscr{C}_t(S_t^\alpha v_i)$	–	0.644	0.704	0.636	–	–	0.704	–	0.704

Step 6. Find the ε-neighborhood of S_t^α:

$$N_t(S_t^\alpha, \varepsilon) = \{3.7.9\}.$$

Step 7. Update the index parameters for all criteria x, y within $Path^*(v_i, v_j)$:

		v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
	v_1	–	–	–	–	–	–	–	–	–
	v_2	–	–	–	–	–	–	–	–	–
	v_3	–	–	–	–	–	–	2	–	–
	v_4	–	–	–	–	–	–	–	–	–
$I_t^{(new)}$	v_5	–	–	–	–	–	–	–	–	–
	v_6	–	–	3	–	–	–	–	–	–
	v_7	–	–	1	–	–	–	–	–	1
	v_8	–	–	–	–	–	–	–	–	–
	v_9	–	–	–	–	–	–	–	–	–

Step 8. Compute the adjustment factor η_i:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$\eta =$	0.805	0	0.693	0	0.859	0.985	0.693	0.7	0.693

Step 9. For all $v_i, v_j \in S_t^\alpha \cup N_t(S_t^\alpha, \varepsilon)$, update the connectivity from v_i to v_j: (Continue until one epoch of training is completed).

		v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
	v_1	1.000	0.663	0.211	0.274	0.419	0.305	0.305	0.504	0.193
	v_2	0.618	1.000	0.539	0.755	0.243	0.600	0.687	0.502	0.663
	v_3	0.187	0.219	1.000	0.408	0.645	0.335	0.713	0.516	0.502
	v_4	0.563	0.607	0.234	1.000	0.277	0.535	0.813	0.428	0.314
$C_t^{(new)}$	v_5	0.682	0.623	0.570	0.498	1.000	0.605	0.386	0.294	0.068
	v_6	0.097	0.225	0.584	0.235	0.145	1.000	0.221	0.499	0.394
	v_7	0.903	0.723	0.832	0.599	0.452	0.338	1.000	0.786	0.710
	v_8	0.131	0.099	0.460	0.304	0.604	0.363	0.586	1.000	0.563
	v_9	0.069	0.427	0.415	0.661	0.204	0.733	0.530	0.271	1.000

Step 10. Calculate the average $\overline{\eta_i}$ of all η_i's.

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
η_1	0.80	0.00	0.69	0.00	0.86	0.99	0.69	0.70	0.69
η_2	0.74	0.00	0.00	0.00	0.81	0.95	0.00	0.82	0.56
η_3	0.69	0.69	0.71	0.69	0.69	0.96	0.69	0.69	0.69
η_4	0.68	0.57	0.57	0.50	0.00	0.95	0.74	0.65	0.42
η_5	0.95	0.75	0.57	0.64	0.96	0.69	0.57	0.57	0.56
η_6	0.72	0.00	0.72	0.00	0.60	0.97	0.72	0.46	0.56
$\overline{\eta_i}$	0.72	0.00	0.72	0.00	0.60	0.97	0.72	0.46	0.56

Step 11. Compute the weights $w_{t+1}(v_i)$ for each criterion v_i at stage $t + 1$:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$w_t(v_i)$	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111
$\overline{\eta_i}/\sum_i \overline{\eta_i}$	0.1435	0.0630	0.1023	0.0576	0.1229	0.1727	0.1068	0.1220	0.1092
$\Delta w_t(v_i)$	0.0194	−0.0288	−0.0053	−0.0321	0.0071	0.0370	−0.0026	0.0066	−0.0012
$w_{t+1}(v_i)$	0.1305	0.0823	0.1058	0.0790	0.1182	0.1481	0.1085	0.1177	0.1099

Step 12. Update the learning rate.

$$\xi_{t+1} = \xi_i \times 0.5 = 0.1.$$

Step 13. Test for the stopping condition:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$w_t(v_i)$	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111	0.1111
$w_{t+1}(v_i)$	0.1305	0.0823	0.1058	0.0790	0.1182	0.1481	0.1085	0.1177	0.1099
$\partial w_t(v_i)$	0.1746	0.2592	0.0477	0.2889	0.0639	0.3330	0.0234	0.0594	0.0108

if $\max \partial w_t(v_i) = 0.3330 > 0.01$, then continue.
Modifying the adjustment procedure for the learning rate over seven iterations (epochs) gives the following results:

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$w_t(v_i)$	0.1305	0.0797	0.1105	0.0753	0.1196	0.1706	0.1124	0.1015	0.0998
$w_{t+1}(v_i)$	0.1306	0.0794	0.1104	0.0750	0.1195	0.1707	0.1129	0.1014	0.1000
$\partial w_t(v_t)$	0.0008	0.0038	0.0009	0.0040	0.0008	0.0006	0.0044	0.0010	0.0020

if $\max_i \partial w_t(v_i) = 0.0044 < 0.01$, then stop.
This gives the weight vector (Iteration 10):

	v_1	v_2	v_3	v_4	v_5	v_6	v_7	v_8	v_9
$\mathbf{w} =$	0.1306	0.0794	0.1104	0.0750	0.1195	0.1707	0.1129	0.1014	0.1000

Among the nine evaluation criteria, the most important criterion in the motorcycle drivers mind is the travel cost (v_6) and its weight is 0.1707; the second is walking time (v_1) with weight 0.1306. The criterion of least importance is the transfer time (v_4) and its weight is 0.0750; the next is the waiting time (v_2) with weight 0.0794.

In actual practice, the purpose of the transportation authorities is to attract motorcycle travelers to use public transit. After finding out the major criteria considered by motorcycle travelers, we can promote the criteria in which public transportation performs better than motorcycle. On the other hand, transportation authorities should improve public transit in those criteria with worse performance than motorcycle or employ relevant policy to place constraints on motorcycle. In a word, if motorcycle users cantake into account those criteria with better performance in transit than motorcycles, it is possible to persuade them to switch their mode to transit.

As for the values of parameters α, β, δ, ξ, and ε, they are obtained by the subjective judgment of researchers and, as well, according to the needs of the authorities. For those who employ motorcycle as their commuter mode, habitual decision exists in their mode choice behavior (Chen et al., 1997). Due to the fact that motorcycle users mode choice is a routine or mixed routine problem, the weights of criteria are rather stable. Such being the case, we will devise the specification principles of parameter values. Also, the measurement of the ups and downs for parameter designation depends primarily on the effect of convergence on the network. For routine and mixed routine problems, the determination of parameter should avoid the oscillation of weight values or divergence of the network resulting from over-revision during the learning process of the network. As a result, reasonable parameter determination should perform desirable convergence within the value range and acceptable sensitivity so as to enhance the efficiency of network. It is suggested that we use simulation data or partial empirical data to put into the network as the test illustration. According to the experiment experience, we can determine more favorable ranges of parameter values and obtain desirable convergence within such range for the network, which can be offered as reference for future empirical application.

Before the application of weight-assessing method in this study, we conducted the sensitivity analysis of the initial weight with the simulation data. Regardless of initial values of criterion weights, it is found that there would not be much difference after several learning iterations. Since weight assessing is low sensitive to the initial weights, we have assumed that the initial weight value of each criterion is the same during the empirical study. It is further suggested that subsequent study can investigate the intrinsic preference structure of each criterion so as to learn about the initial values of weights.

7. Conclusion

Weights of criteria are the deciding factors in decision-making, but from a behavioral perspective, traditional weighting methods take into account too few factors to deal with them effectively. Based on a model of behavior mechanism and the theory of habitual domains, we use connectivities between criteria to establish a network structure. We define a fuzzy directed graph using the concept of connectivity, and call it the connectivity network. We then identify the neighborhood of the fired criteria during the stimulus-response process, and we specify the actual domain in a broad sense as the generalized current domain. Our model treats decision-making as a dynamically adjusting process from the ideal state to the actual state. We develop the new weight-assessing method using the theory of habitual domains, which is based on competitive learning. Our assessing method performs a generalized winner-take-all competition. We prove that the weight is only related to the level of charge and the connectivity of the criterion with the existing domain in the equilibrium state for routine and mixed routine problems. Moreover, ceteris paribus, the weight also has a stable fixed value. However, for fuzzy problems, the weight does not remain stable during the whole decision-making process. The algorithm presented is suitable for routine, mixed routine, or fuzzy problems.

Finally, we have studied motorcycle users' mode-choice behavior in Taipei City through questionnaires in order to show the applicability of the weight-assessing method with habitual domains. From the analysis results of the empirical study, it is found that the most important criteria for motorcycle travelers are the travel cost and the walking time, while the criteria with less importance are the transfer time and the waiting time. The weight-assessing method being put forward in this study, aside from being applied to decision-making problem, can help describe the consumption and choice behaviors. Furthermore, it can be also used in new product design, marketing management, pricing and market segmentation. To sum up, our weight-assessing method, in comparison with other weight methods, can help researchers to approximate human thinking process more accurately. Furthermore, its implementation procedures are feasible in practice.

References

Chen, T.-Y., Chang, H.-L., Huang, S.-F., 1997. A study of attributes interaction in decision making for mode choice using habitual domains – a case of motorcycle users in Taipei. Transportation Planning Journal 26 (1), 1–36 (in Chinese).

Cook, W.D., Kress, M., 1988. Deriving weights from pairwise comparison ratio matrices: An axiomatic approach. European Journal of Operational Research 37 (3), 355–362.

Eckenrode, R.T., 1965. Weighting multiple criteria. Management Science 12 (3), 180–192.

Fausett, L., 1994. Fundamentals of Neural Networks: Architectures, Algorithm, and Applications. Prentice-Hall, Englewood Cliffs, NJ.

Haykin, S., 1994. Neural Networks – A Comprehensive Foundation. Macmillan, New York.

Hwang, C.L., Yoon, K., 1981. Multiple Attribute Decision Making – Methods and Applications: A State-of-the-art Survey. Springer, Berlin.

Kaufmann, A., Gupta, M.M., 1991. Introduction to Fuzzy Arithmetic – Theory and Application. Van Nostrand Reinhold, New York.

Keeny, R.L., Nair, K., 1977. Nuclear siting using decision analysis. Energy Policy 8 (2), 223–231.

Keeny, R.L., Raiffa, H., 1976. Decisions with Multiple Objectives: Preferences and Value Trade-offs. Wiley, New York.

Krovak, J., 1987. Ranking alternatives: Comparison of different methods based on binary comparison matrices. European Journal of Operational Research 24 (1), 86–95.

Miyamoto, S., 1990. Fuzzy Sets in Information Retrieval and Cluster Analysis. Kluwer Academic, Dordrecht.

Nijkamp, P., 1977. Stochastic quantitative and qualitative multicriteria analysis for environmental design. Papers of the Regional Science Association 39, 175–199.

Rumelhart, D.E., McClelland, J.L. and the PDP Research Group, 1986. Parallel Distributed Processing, Explorations in the Microstructure of Cognition, vol. 1. Foundations. MIT Press, Cambridge, MA.

Saaty, T.L., 1977. A scaling method for priorities in hierarchical structures. Journal of Mathematical Psychology 15 (3), 234–281.

Saaty, T.L., 1980. The Analytic Hierarchy Process. McGraw-Hill, New York.

Shi, Y., Yu, P.L., 1987. Habitual domain analysis for effective decision making. Asia-Pacific Journal of Operational Research 4 (2), 131–150.

Srinivasan, V., Shocker, A.D., 1973. Linear programming techniques for multidimensional analysis of preference. Psychometrika 38 (3), 337–369.

Voogd, H., 1983. Multicriteria Evaluation for Urban and Regional Planning. Pion, London.

Yu, P.L., 1980. Behavior bases and habitual domains of human decision/behavior – concepts and applications. In: Fandel, G., Gal, T. (Eds), Multiple Criteria Decision Making Theory and Application. Springer, New York, pp. 511–539.

Yu, P.L., 1985. Multiple Criteria Decision Making: Concepts, Techniques, and Extensions. Plenum Press, New York.

Yu, P.L., 1990. Forming Winning Strategies: An Integrated Theory of Habitual Domains. Springer, New York.

Yu, P.L., 1991. Habitual domains. Operations Research 39 (6), 869–876.

Yu, P.L., Zhang, D., 1989. Competence set analysis for effective decision making. Control Theory and Advanced Technology 5 (4), 523–547.

Yu, P.L., Zhang, D., 1990. A foundation for competence set analysis. Mathematical Social Sciences 20 (3), 251–299.

Yu, P.L., Zhang, D., 1992a. Competence set analysis for fuzzy problems. IEEE, pp. 323–330.

Yu, P.L., Zhang, D., 1992b. Effective expansion of a practically known competence set. In: Tzeng, G.H. et al. (Eds.), Multiple Criteria Decision Making. Springer, New York, pp. 189–198.

Zeleny, M., 1974. Linear Multiobjective Programming. Springer, Berlin.