

ABOUT ONE APPROACH TO AHP/ANP STABILITY MEASUREMENT

Vitaliy Tsyganok
National Academy of Sciences of Ukraine
The Institute for Information Recording
Kyiv, Ukraine
E-mail: vitaliy.tsyganok@gmail.com

ABSTRACT

AHP/ANP stability measurement methods are described. In this paper we define the method's stability as the measure of its results dependence on the expert's errors, made during pair comparisons. Ranking Stability (order preservation in alternative ranking under natural expert's errors, made during expert estimation) and Estimating Stability (maintaining alternative weights within the specified maximal relative inaccuracy range) are considered. Targeted Genetic Algorithm search procedure is used for possible stability violation detection. Then division-in-half (dichotomy) method is applied to calculate stability metric of a given criteria hierarchy.

Keywords: AHP/ANP stability measurement, Ranking/Estimate stability, Genetic Algorithm

<http://dx.doi.org/10.13033/ijahp.v3i1.50>

1. Introduction

In the Decision Making Systems (DSS), particularly, while using AHP/ANP (Saaty, 1980; Saaty, 2008) for getting decision variants, the question of the decision's stability degree arises. In other words, the question is: what is the extent of dependence between expert's errors (made during criteria hierarchy/network model building and during alternatives estimation according to a chosen set of criteria) and decision's stability?

For example, in case of criterion hierarchy, used for choosing the best alternative from the set, it is undesirable for the object's place in the rating to depend upon a single slight difference in pair-wise comparisons of two sub-criteria influence coefficients (like difference between "Very, very strong" or "Extreme" domination).

It is obvious that decision stability depends on both criteria hierarchy structure and numerical values reflecting experts' priorities.

Having analyzed the essentials of AHP method, we can come to a conclusion that complexity of criteria hierarchy structure, and therefore, stability of decisions made on its basis, depends on lengths of ways from any alternative in the hierarchy to the main criterion/goal, because, the longer the ways are, the more ratios characterizing criteria influence degree "contribute" to alternative weights' calculation.

2. Development of decision stability estimation method

2.1 Problem classes

Let's consider two main classes of problems solved by AHP. These are: alternatives' ranking and their estimation. In the case of alternative ranking-associated decisions we shall define decision *stability* as the property of alternatives' rank order preservation under natural expert's errors, made during expert estimation. We suggest using the probability of rank order preservation as stability metric. In order to ensure decisions' stability in case of alternative weights' calculation, obtained estimates are required to lie within the range, limited by specified relative error. In this case, the probability of keeping the alternative weight within the range may be chosen as stability metric. It should be stressed that the solution procedure for decision stability-related problem is the same for both aforementioned cases.

So, let's formulate the problem statement.

2.2 Problem statement

What is given:

- 1) A ready-built criteria hierarchy;
- 2) Expert priorities, given in the form of pair-wise comparison matrices (PCM) in Fundamental scale.

We suppose that experts have already built PCM-s for calculating criterion and alternative weights (it is presumed that alternatives have been compared according to each of bottom-level criteria).

What is to be obtained: stability metric value under given conditions.

2.3 Stability metrics

Let's model expert's errors, made in pair-wise comparisons k_i , $i \in \{1..n\}$ by setting a specified *relative deviation* level of dominance values (in order to avoid complications with indices we define n as the total number of pair-wise comparison values in all the matrices built for the given hierarchy). In our model the deviation value is considered the same for all pair-wise comparisons of criterion weights and of all alternatives according to each criterion of the given hierarchy. Hereinafter, deviation is the boundary random deflection value of every pair-wise comparison k_i for the given criteria hierarchy measured as percentage of its current value. A hierarchy is considered subject to influence of deviation Δ when every pair comparison is varied according to the following law: $k_i^* = k_i + R_i \cdot k_i \cdot \Delta / 100$, where R_i is a random value, evenly distributed in the $[-1;1]$ range. It is significant, that, for modeling purposes, while evaluating stability metric the Fundamental scale will be considered continual.

We shall define the maximal deviation Δ , under which AHP result remains stable (in

other words, no rank reversals occur, or, in case of another problem class, obtained alternative weights remain within the given range) as *stability metric*.

2.4 Solution process

Let's obtain maximal alternative weight deviation for a given criteria hierarchy as relative value measured in percentage, through *division-in-half (dichotomy) method*. In order to do this let's specify a certain arbitrary deviation value within the search range between 0% and 100% and subject the hierarchy to the influence of the deviation. For the specified deviation value let's check whether AHP result keeps stable in comparison with the result under deviation 0%. If the result remains stable then the search for maximal deviation is continued in the range limited by the value used in previous calculation and upper bound of the range (i.e., 100%) (to be exact, the middle of the range is considered). Otherwise, i.e. when the result is not stable, the middle of another range (lying to the left of the value used in the previous calculation step) is considered while checking. The searching range converges with every such iteration. This iterative process ends when desired accuracy of deviation is achieved. It means that the process stops when the middle value of the search range is different from one of the boundary values of the range, limited by desired accuracy. For example, if the initial relative value deviation value is 10% and after all pair comparisons have been subjected to such deviation AHP results remained stable, the next deviation value will be 55% (the average of 10 and 100%). Otherwise (i.e. in case of stability breach), the next deviation value will constitute 5%, which is the average of 0 and 10%.

2.5 Minimizing the possibility of undetected stability breach

It should be noted that since the process of expert's errors modeling is stochastic, it is possible that stability breach won't be detected at some current deviation level. While applying the specified searching method it is very important to minimize this possibility, because unlike in sequential search when we are using division-in-half (dichotomy) method, though the searching process is faster, any omission of stability violation switches the search to the wrong range, which often leads to incorrect results and conclusions. For example, suppose it is known for sure, that rank reversal in some hierarchy occurs under the deviation of more than 3%, but at some dichotomy step, when the value of deviation amounted to 5%, rank reversal wasn't detected. Then, even in the case when during all the following tests of the rest of deviation values existing rank reversals are detected, the deviation value obtained by division-in-half (dichotomy) method will be greater than 5%.

One of the obvious ways to avoid non-detected stability violations under specified deviation is the complete enumeration of possible variants of paired comparisons k_i at this deviation level. Under the obvious assumption, that dependence of stability violation's probability from deviation value is a monotonously increasing function, we can conclude that, in the first place, stability violations begin to occur under extreme deflections of pair comparisons values (i.e., on the boundaries of deviation range). So, in order to reduce the search range significantly and to detect possible stability violations as fast as possible, let's specify only the outermost values of the deflection range during enumeration:
 $k_i^* = k_i \pm k_i \cdot \Delta / 100$.

In this case a number of pair comparison values' variants that should be considered during exhaustive search is 2^n , where n is a number of pair comparisons that were carried out for a given criteria hierarchy. This kind of enumeration makes sense only under small values of n , when n does not exceed 15-20. In order to detect AHP results' stability breaches in a real criteria hierarchy, we suggest applying Genetic Algorithm (Holland, 1994).

2.6 Specific problem

Let's consider the problem statement, which, in author's opinion, can be most easily solved by applying classic genetic algorithm. The problem could be formulated in the following way.

What is given:

- 1) A ready-built criteria hierarchy, based on n pair comparisons, conducted by the expert;
- 2) Deviation of the pair comparisons δ .

We should find: a set of pair comparisons' values obtained under influence of deviation δ , on which maximal stability violation of AHP results is observed.

2.7 Formulating genetic algorithm's utility function

When the last formulated condition is satisfied, detection of stability violation can be guaranteed (i.e., maximal stability breach will certainly be detected). Utility function, realized in genetic algorithm, will be based on this condition. This utility function should possess larger values while more drastic stability violations are observed. In the case under consideration the function will be specified algorithmically by means of *Eigenvector* method realization for each set of objects, whose pairs are compared, followed by AHP method application and calculation of maximal result values' deflection from stable values (i.e. values calculated under zero deviation $-\delta = 0$).

In order to calculate maximal deflection of AHP method's results from stable values, let's consider each of the two aforementioned problems' classes separately. While considering alternative ranking-related problems, it seems logical to use Kemeny's length or distance (Kemeny & Snell, 1972) as AHP results' deflection measure. This value is determined on two relations A and B , where A characterizes ranking α – in the absence of deviation, B characterizes ranking β – based on a specified set of pair comparisons' results under deviation δ :

$$D(A, B) = \sum_{i, j} |\alpha(i, j) - \beta(i, j)|. \quad (1)$$

But after further analysis it was found, that Kemeny's distance did not fully satisfy the requirements to utility functions, used in genetic algorithm, because it was not sensible to such changes of AHP results, which didn't cause rank reversals. That is, on various input data sets (various pair comparisons' sets) utility function value remains the same in case

there is no rank reversal. And in this case, criteria for targeted search, applied in Genetic Algorithm, are absent. Therefore, for this class of decision making problems, concerning alternatives' ranking, the following empiric function was suggested as utility function to be used in genetic algorithm:

$$f_1 = \sum_{i=1}^{k-1} \frac{b_i - b_{i+1}}{a_i - a_{i+1}},$$

(2)

where a_i is a cardinal estimate of alternative with i -th rank, obtained under zero deviation; b_i is a cardinal estimate of alternative with i -th rank, obtained on a given set of pair comparison results, that were subjected to deviational influence δ ; k is a number of alternatives in the hierarchy.

As for utility function for a problem class, where hierarchies are analyzed as to keeping the results of alternatives' estimations within the bounds of the maximal relative error, the following function for calculating that maximal relative error corresponds to the specificity of the genetic algorithm-based search:

$$f_2 = \max_i \left(\frac{|a_i - b_i|}{a_i} \right),$$

(3)

where a_i and b_i are cardinal estimates of the i -th alternative, obtained without and under deviation of pair comparison results, respectively.

2.8 Adapting genetic algorithm to the specified problem

“Population” concept, used in genetic algorithm, corresponds to a set of fixed potency, consisting of sets of pair comparison results in criteria hierarchy, i.e. decision variants (in genetic algorithm they are called “individuals” in population).

The difference between the specified procedure and classical implementation of genetic algorithm is that during the search process, stability violation checks are performed, and the process stops if such violation is detected. In case genetic algorithm has completed its work and stability violation is not detected during the search, it is deduced that decisions made under given deviation δ are stable.

This feature of genetic algorithm realization allows to discard the possible issue of function (2) domain, namely its behavior while $a_i \leq a_{i+1}$ (when zero or a negative value can appear in the fraction's denominator). Under such values decision stability violation (i.e. rank reversal) is detected, so genetic algorithm ends its work and, therefore, further calculation of utility function is unnecessary.

Thus, the specified genetic algorithm is a single step in the solution procedure of the more general problem of obtaining AHP stability metric, stated in section 2.2 of this paper.

2.9 Software used in the experiment

In order to analyze criteria hierarchies using the aforementioned technique, the software applications were developed and integrated into an existing DSS. A new mode called “Decisions Stability Analysis”, envisioning the input of data for calculations, was added to the DSS.

During calculations, the subsystem of DSS allows detection of two types of decision stability violations: alternative rank reversals and cases when calculated alternative estimate goes beyond the bounds of permissible deflection.

In order to verify genetic algorithm’s efficiency under specified parameters, namely “Power (cardinality, potency) of population”, “Mutation probability” and “Number of generations with the same results”, exhaustive search method was applied for smaller criteria hierarchies.

Screenshots, illustrating the software used in the experiment, are shown on Figures 1 and 2.

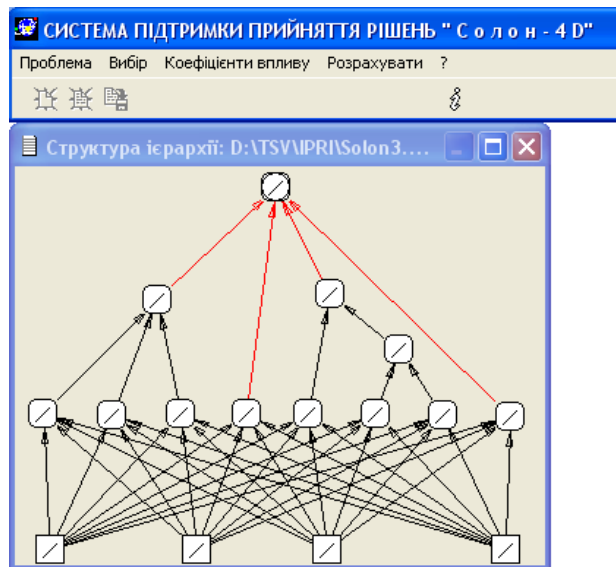


Figure 1. DSS interface with a given hierarchy structure.

девіація ЧКВ	5,00 %	тип пошуку	Макс. кількість перерахунків	500
<input checked="" type="checkbox"/> визначити макс. девіацію ЧКВ методом половинного ділення		<input type="radio"/> стохастичний	Кількість осіб в популяції	100
<input checked="" type="checkbox"/> Визначити ЧКВ по парних порівняннях		<input checked="" type="radio"/> генетичний алгоритм	Ймовірність мутації	0,25
при умові коли		<input type="radio"/> повний перебір	Кількість поколінь з однаковим результатом	500
<input type="radio"/> не виникає реверс рангів			область пошуку	<input type="radio"/> +/-0 <input checked="" type="radio"/> +/-
<input checked="" type="radio"/> макс. відносна похибка не перевищує	5,00 %			

Вперед

Figure 2. Search parameters input window.

3. Example

What is given:

- 1) Criterion hierarchy with main criterion (denoted as 0) and sub-criteria (denoted as 1-11);
- 2) Set of alternatives, whose weights are being calculated (denoted as 12-15). (See Figure 3).

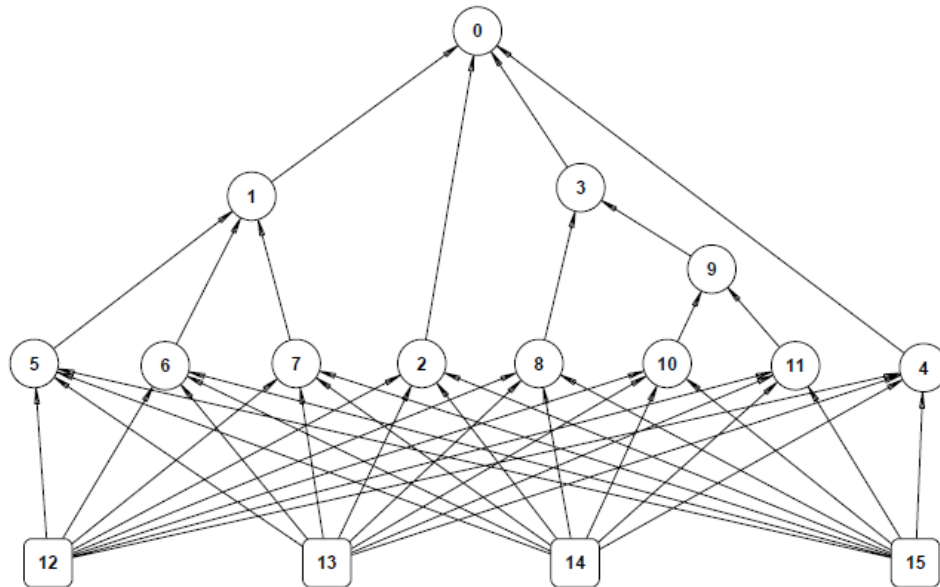


Figure 3. Graph of criterion hierarchy.

All experts' priorities are given in the matrices below.

PCM, where influences of sub-criteria 1, 2, 3 and 4 upon criterion 0 are compared in

Fundamental Scale:
$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ & 1 & 2 & 2 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}.$$

PCM, where influences of sub-criteria 5, 6 and 7 upon criterion 1 are

compared:
$$\begin{bmatrix} 1 & 2 & 3 \\ & 1 & 2 \\ & & 1 \end{bmatrix}.$$

PCM, where influences of sub-criteria 8 and 9 upon criterion 3 are compared:
$$\begin{bmatrix} 1 & 1/2 \\ & 1 \end{bmatrix}.$$

PCM, where influences of sub-criteria 10 and 11 upon criterion 9 are compared:
$$\begin{bmatrix} 1 & 6 \\ & 1 \end{bmatrix}.$$

PCM for comparing alternatives (nodes 12-15) according to criterion

$$\begin{array}{l}
 2: \begin{bmatrix} 1 & 1/2 & 1/3 & 1/4 \\ & 1 & 1/2 & 1/2 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}; \quad \text{according to criterion} \quad 4: \begin{bmatrix} 1 & 1/2 & 1 & 2 \\ & 1 & 1/2 & 2 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}; \\
 5: \begin{bmatrix} 1 & 2 & 3 & 4 \\ & 1 & 2 & 2 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}; \quad 6: \begin{bmatrix} 1 & 2 & 3 & 4 \\ & 1 & 2 & 2 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}; \quad 7: \begin{bmatrix} 1 & 2 & 2 & 1/2 \\ & 1 & 1 & 1/2 \\ & & 1 & 1/4 \\ & & & 1 \end{bmatrix}; \quad 8: \begin{bmatrix} 1 & 3 & 2 & 5 \\ & 1 & 2 & 2 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}; \\
 10: \begin{bmatrix} 1 & 3 & 1/2 & 1/2 \\ & 1 & 1/6 & 1/5 \\ & & 1 & 1 \\ & & & 1 \end{bmatrix}; \quad 11: \begin{bmatrix} 1 & 5 & 2 & 2 \\ & 1 & 1/3 & 1/4 \\ & & 1 & 2 \\ & & & 1 \end{bmatrix}.
 \end{array}$$

It should be noted here that given matrices are consistent enough to apply eigenvector method for alternative weights' calculation.

What is to be obtained?

Maximal relative deflection of pair comparisons from their given values, under condition that relative deflection of calculated weights does not exceed 5%.

Using AHP, we have calculated weights of alternatives 12-15 without deviation. They are: {0.2845873027, 0.20527370971, 0.2767528361, 0.23338615149}; let's call them *reference weights*.

During the targeted search we calculate alternative weights, under condition, that some expert priorities have changed (increased or decreased) by some specific value measured in percentage. For a start, let's define this value as, say, 10%. It will be the starting point of dichotomy search.

We should note that during targeted genetic algorithm-based search, the search range includes all possible variants of expert priorities' (i.e. PCM elements) change. In the given example all PCM elements (except diagonal ones) were changed by $\pm 10\%$. The total number of such non-diagonal PCM elements in our example is 59. Let's define this set as $K = \{k_i\}$, $i = \{1..59\}$. Thus, $K = \{2, 3, 4, 2, 2, 2, 2, 3, 2, 1/2, 6, 1/2, 1/3, 1/4, 1/2, 1/2, 2, 1/2, 1, 2, 1/2, 2, 2, 2, 3, 4, 2, 2, 2, 2, 3, 4, 2, 2, 2, 2, 2, 1/2, 1, 1/2, 1/4, 3, 2, 5, 2, 2, 2, 3, 1/2, 1/2, 1/6, 1/5, 1, 5, 2, 2, 1/3, 1/4, 2\}$

So, at present stage, genetic algorithm-based search is targeted at defining whether the weight of any alternative can change by more than 5% as a result of changing any of 59 PCM elements by $\pm 10\%$. In the given example while $K = \{1.8, 2.7, 4.4, 1.8, 2.2, 1.8, 2.2, 3.3, 1.8, 0.45, 5.4, 0.55, 0.3, 0.225, 0.55, 0.55, 2.2, 0.45, 0.9, 1.8, 0.55, 1.8, 1.8, 1.8, 3.3, 4.4, 2.2, 2.2, 1.8, 2.2, 3.3, 3.6, 2.2, 2.2, 1.8, 2.2, 1.8, 0.55, 1.1, 0.55, 0.275, 3.3, 2.2, 4.5, 1.8, 2.2, 2.2, 2.7, 0.45, 0.55, 0.18(3), 0.18, 0.9, 5.5, 1.8, 2.2, 0.367, 0.275, 1.8\}$, alternative weights are {0.31069680293, 0.2068284784, 0.26215412441,

0.22032059426}, so the 1st alternative's weight exceeds the reference value by 8.40353%, which is more than 5%. So, at this point the further search under given deviation value (10%) stops. To continue the search the new deviation value of $(0\%+10\%)/2=5\%$ is set. Under this deviation value (5%) the procedure finds another variant of K , on which calculated alternative weights differ from reference values by more than 5%.

Under deviation value of $(0\%+5\%)/2=2.5\%$ genetic algorithm doesn't find any variants of K , on which alternative weights differ from reference values by more than 5%, so the next deviation value is $(2.5\%+5\%)/2=3.75\%$. Under this deviation value there exists a variant of K , on which stability violation is witnessed (hereafter we shall mark such variants with "+", and negative search results with "-"). The search process continues with the following results: under deviation $(2.5\%+3.75\%)/2=3.13\%$ (value 3.125 is approximated to the second decimal digit) – "+", under deviation $(2.5\%+3.13\%)/2=2.82\%$ – "+", 2.66% – "+", 2.58% – "+", 2.54% – "-", 2.56% – "+", 2.55% – "+". So, we have found the maximal relative deviation of pair comparison values. It is 2.54%, and under such deviation the absolute relative difference between reference alternative weights and their calculated values does not exceed 5%.

4. Conclusions

A technique for calculating decisions' stability measure is suggested. We consider AHP-based decisions, which are made using expert data, input into an existing DSS. The developed decision stability estimation methodology can also be applied to obtaining stability measure in ANP method, and integrated into other systems, where expert information is used.

The described stability metric calculation technique can be used for stability breach detections, resulting from deviations of criterion weights (influence coefficients) and alternative estimates.

In cases when stability metric values do not satisfy posed requirements it is possible to develop methods for improving decision stability using the suggested technique.

REFERENCES

Holland, J.H. (1994). *Adaptation in natural and artificial systems. An introductory analysis with application to biology, control, and artificial intelligence*. London: Bradford book edition.

Kemeny, J.G., & Snell, J.L. (1972). *Mathematical Models in the Social Sciences*. Cambridge, MA: MIT Press.

Saaty, T.L. (1980). *The Analytic Hierarchy Process*. New York, NY: McGraw-Hill.

Saaty, T.L. (2008). Relative Measurement and Its Generalization in Decision Making. Why Pairwise Comparisons are Central in Mathematics for the Measurement of Intangible Factors. The Analytic Hierarchy/Network Process. *Statistics and Operations Research*, 102 (2), 251–318.