ISAHP 2007, Viña Del Mar, Chile, August-2-6, 2007

# SELECTING TEST LOCATIONS: AN AHP APPLICATION TO BRAZILIAN DIGITAL INCLUSION PROJECTS

***Graziella C. Bonadia; José Carlos Lima Pinto; Giovanni M. de Holanda***
CPqD – Research and Development Center in Telecommunications
{bonadia, jcarlos, giovanni}@cpqd.com.br

**Abstract:** *This paper describes a process that allows choosing the test locations of two digital inclusion services to be widely deployed in Brazil. The process is based on a multi-criteria analysis, specifically the analytic hierarchy process (AHP), since the services have different characteristics and should be jointly tested in telecenters. The cities should be selected according to the following general requirement :the "easy to find potential users" criterion should include the illiteracy density, demand for health services, balance between urban and rural populations and a fair number of elderly people over 60 years old. The higher the figures, the more appropriate the city. Based on the results obtained from such an analysis, it was possible to evaluate how useful the method is. Moreover, it may be also applied to support other decision-making processes, mainly when several elements should be considered and the problem to be solved involves complex and multi-faceted issues.*

**Keywords:** *Multi-criteria analysis, Analytic Hierarchy Process, digital divide, ICT*

## 1 Introduction

The problem of digital divide has been thoroughly discussed around the world, and it is more and more the focus of research and development aiming at reducing that gap. In the process of developing services based on ICTs (Information and Communication Technologies), it should be taken into account that a considerable part of the Brazilian public now excluded is comprised by people with low levels of literacy. For those people, the mere supply of access to ICT will not help to reduce their digital divide. Therefore they must be able to understand and use ICT-based services so that they can really benefit from those technologies. Thus, within the scope of the project named Telecommunication Solutions for Digital Inclusion – STID[1], two services for the promotion of citizenship are being developed: a service to support retirement and another for the health service. Although those services can be used by any user profile, its interface also takes into account the audience with low or no literacy levels.

Due to the diversity of the target audience regarding the services, it is important that their interfaces are validated in a real day-to-day environment. For that purpose, the choice of test sites should take into account the existence of a certain potential demand for the use of services, besides a relevant number of people with the profiles most targeted by the services. The main focus of this article is to present the method used to identify a group of Brazilian municipal districts that fall within the requirements of the experiment, in a real environment, presenting all the conditions suitable for the inclusion solutions developed by this project. Besides a description of the process to select the municipal districts, the article also intends to show the applicability of the employed method, in terms of robustness and discriminant power.

The process described herein is based on the definition of analysis criteria, and indicators associated to those criteria, considering the characteristics required by the inclusion services to be developed. Thus, Analytic Hierarchy Process (AHP) is used to calculate the importance of each indicator for the most appropriate choice of municipal districts.

[1] An overview on the general line of this project is presented in (Holanda & Dall'Antonia, 2006).

## 2 Criteria and indicators

For an effective validation of those services, particularly in terms of meeting the needs of the target audience regarding digital and social inclusion, the analysis and selection of municipal districts where the field tests will be conducted should consider places in which: (i) the users that belong to the target audience are strongly present, and (ii) the users' habits are compatible with the services in development, such as demand for health services and a balanced distribution of the target audience among rural and urban areas. Besides, implementation and operation cost aspects of potential test environments should be considered.

Thus, the general goal of the selection method is to obtain a group of municipal districts where it is more simple to recruit the users specified in (i) and (ii). The criteria that translate the easiness of user selection were established according to the quantification and relativization of some indicators, which are: illiteracy density, elderly people rate (over 60 years old), demand for health services, and balance between urban and rural populations. Then, the indicators related to those criteria were analyzed, as detailed in (Pinto *et al*., 2007) and summarized next.

### 2.1 High rate of illiteracy

Great differences exist among complete and functional illiterates, which can generate development requirements for specific inclusion solutions for each profile. Thus, the possibility of having different solutions to meet the different profiles leads to the need of finding both complete and functional illiterates to perform the tests. However, the indicators that measure the density of complete illiteracy and the density of functional illiteracy are strongly correlated (0.94). Thus, to prevent the high correlation from interfering in the calculation of the variable, only one of the two illiteracy indicators was used. Then, it was established to adopt the indicator "density of complete illiteracy per municipal district", built from the ratio between the number of complete illiterates and the total area of each municipal district, according to IBGE [2] data (Census 2000).

### 2.2 High rate of elderly people

This criterion translates the need of a senior audience to test some applications and functionalities of the retirement service. The indicator used to check whether this criterion is met is "rate of elderly per municipal district" which allows the selection of municipal districts with greater percentile of elderly compared to the total population. The "rate of elderly per municipal district" was obtained by the ratio between the number of individuals aged 60 or above, and the total population of each municipal district, based on IBGE data (Census 2000).

### 2.3 High demand, and lack of basic health care services

The health care service proposed in the STID project will act in synergy with basic health care programs. The most appropriate places to conduct the tests are those where the demand for health services is higher. An indicator that translates both the demand and the lack of basic health service is obtained by a combination of the "annual average of medical consultations per inhabitant in the basic specialties" and the "monthly average of home visits per family", generating a single indicator called "demand for health services"[3].

### 2.4 Balance between urban and rural populations

To minimize the difficulties of users' locomotion and access network infrastructure, the service offers that are being developed under STID should occur via a telecenter structure with distributed access

---

[2] Source: IBGE – Instituto Brasileiro de Geografia e Estatística (Brasilian Institute of Geography and statistics).

[3] Source: DATASUS – Ministério da Saúde (Health Ministry).

(contemplating both urban and rural areas). To assure the presence of the appropriate target audience for the tests, the criterion “balance between urban and rural populations” was established. The indicator that measures this situation is calculated considering the percentage of rural and urban illiterates compared to the total population of complete illiterates. The criterion seeks to classify the municipal district with the broadest balance between those two percentages, i.e., the same amount of illiterates living in rural and urban areas.

### 2.5 Cost of implementation and operation of tests

Another factor taken into account for the choice of municipal districts where the field tests will be performed was the cost of implementation and operation. Generally, the main regional differences related to costs depend on the distance between the research center coordinating the development of services, and the municipal districts where the tests will be conducted. The relative difference between the costs of six months of operation on the test fields, classified by average distance is listed in Table 1.

**Table 1** Relative difference between the implementation and operation costs regarding the location of the initiative

| Distance between the research center and the test field | Relative difference between costs - 6 months |
|---|---|
| Up to 200 km | 1,000 |
| Up to 400 km | 1,028 |
| Southeastern Region, beyond 400 km | 1,032 |
| Southern Region | 1,042 |
| Midwestern Region | 1,049 |
| Northern Region | 1,073 |
| Northeastern Region | 1,079 |

## 3 Application of the method

Once the indicators for all Brazilian municipal districts are gathered, the score can be obtained for each municipal district in relation to their values. Although many indicators have been identified for application in the process of choosing a test site, they are considered with different degrees of importance. Thus, it is advised to use a calculation procedure of variables with the purpose of describing the levels of importance of each indicator in order to choose a site.

The choice of weights to establish a score standard is not an easy issue. Due to the fact that the indicators used are applied to services with different characteristics, they present different rates when applied to the context of each of the two services. The main purpose, however, is a single parameter applied to a rule for test site selection, i.e., a single formulation where each indicator used has only one associated weight. Such circumstance of choice of weights is a distinctive situation of decision with different objectives, also called a multiple criteria decision situation (Ehrlich, 2004). There are many methods that address this condition, and some of them hierarchically breakdown the objectives in successive levels. The method selected for this analysis was the AHP developed by Saaty (1980). The hierarchical breakdown is applied with the purpose of obtaining the weights to be used in the scoring standard regarding the process of choosing municipal districts. Such process is illustrated in Figure 1, with the indicators associated to the two services.

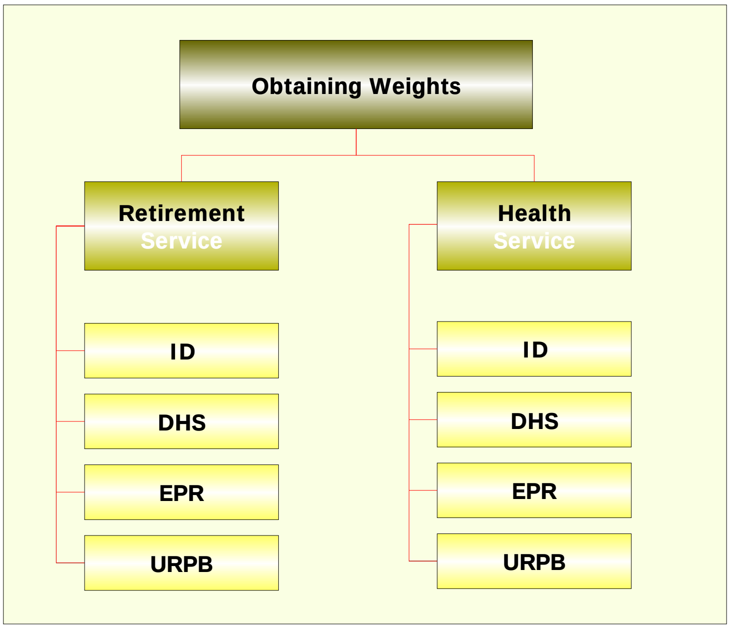


**Figure 1** Hierarchical breakdown of the problem of obtaining weights

In the representation of Figure 1, ID is the "illiteracy density", DHS is the "demand for health services", EPR is the "elderly people rate", and URPB is the "balance between rural and urban populations".

The relative importance and the inconsistency relationships are presented in Table 2.

**Table 2** Comparisons among pairs of attributes (services and indicators) used to define the weights ((a), (b) and (c))

(a) comparison of the relevance among services

| **Comparison of the relevance among services** | |
|---|---|
| **Comparison** | **Level of relevance** |
| Health service vs. Retirement service | 1 |

(b) comparison of relevance among indicators according to the health service viewpoint

| **Comparison of relevance among health service indicators** | |
|---|---|
| **Comparison** | **Level of relevance** |
| ID vs. EPR | 7 |
| ID vs. URPB | 3 |
| DHS vs. ID | 3 |
| DHS vs. EPR | 9 |
| DHS vs. URPB | 7 |
| URPB vs. EPR | 5 |

(c) comparison of relevance among indicators according to the retirement service viewpoint

| **Comparison of relevance among retirement service indicators** | |
|---|---|
| **Comparison** | **Level of relevance** |
| ID vs. DHS | 9 |
| ID vs. EPR | 5 |
| ID vs. URPB | 3 |
| EPR vs. DHS | 3 |
| URPB vs. DHS | 7 |
| URPB vs. EPR | 3 |

By applying the AHP method, the weights listed in Table 3 will be obtained.

**Table 3** Value of ($P_x$) weights obtained by the AHP method

| **Weight** | **Value** |
|---|---|
| $P_{ID}$ | 0,4133 |
| $P_{DHS}$ | 0,3197 |
| $P_{EPR}$ | 0,0746 |
| $P_{URPB}$ | 0,1924 |

In Table 3, $P_{ID}$ is “the weight of illiteracy density indicator”, $P_{DHS}$ is “the weight of demand for health services indicator”, $P_{EPR}$ is “the weight of elderly people rate indicator” and $P_{URPB}$ is “weight of balance between rural and urban population indicator.”

With the indicator weights, the score formula for site selection can be established. The application of a weighed average composed by parts that simply consist in the indicator value multiplied by the indicator weight can generate results that jeopardize an appropriate choice of sites. This is because the score is obtained based only on the sum of weighed indicators, not considering the balance regarding the value of those weighed indicators. Thus, if a municipal district obtains a high value in one indicator and near zero values in others, it is still possible to have a good score. The desired situation would be that, although the site score is high, there would also be a balance among the weighed indicators, to make the site the most favorable for tests, i.e., to have the largest number of desirable characteristics. Seeking to reach a better balance among the weighted variables, a formula was used (Equation (1)) based on the Euclidian distance between the analyzed municipal district and a point that describes a hypothetical municipal district with the best possible characteristics (indicators). Thus, it is possible to check how close the municipal districts are to that point. The smaller the distance, the closest it is to the ideal value, and therefore its position on the list of test sites for selection will be better.

$$Dist_i = (((max(ID) - IDi) * PID)^2 + ((max(DHS) - DHSi) * PDHS)^2 + \\ + ((max(NEP) - NEPi) * PNEP)^2 + )((max(URPB) - URPBi) * PURPB)^2 )^{½} \qquad (1)$$

where $Dist_i$ is the distance of the i-nth municipal district to the point with ideal characteristics, max*(X)* is the greatest possible value for the *X* indicator, $X_i$ is the value of the X indicator for the i-nth evaluated municipal district, and P*X* is the weight of indicator X. The weights used in this equation, as described above, were obtained by the AHP method (Table 3).

With all the procedures described above, it was possible to obtain a rank of municipal districts related to the benefits offered as base for the tests. The next step was to include the costs in the analysis. For this, it was possible to observe the cost-benefit ratio via Equation (2).

$$Rat_i = Dist_i * Cost_i \tag{2}$$

In this equation, $Rat_i$ is the cost-benefit ratio of the i-nth municipal district appraised, $Dist_i$ is the distance from the i-nth appraised municipal district to the point with ideal characteristics, and $Cost_i$ is the cost of the i-nth evaluated municipal district. Thus, the smaller the $Rat_i$ value, the better the municipal district will be classified on the test site selection list.

## 4 Sensitivity analysis and discriminant power

The pair comparisons demanded by the process are subjective, and could cause variations in those entry data. Even knowing which attribute is more relevant than the other, the level of relevance cannot be very clear. Because of that, a desirable property for the analysis is the robustness of the system regarding small variations in values, attributed to the entry comparisons, i.e., it is desirable that small changes in entry data cause no significant variations in the final results.

Given the foregoing, an analysis of the sensitivity of the main comparisons was performed among the indicators with more influence on the obtaining of the final result. Those indicators would be the ones with the greatest weight in the score standard, i.e., the “illiteracy density” (ID) and the “demand for health services” (DHS). Then, it was possible to observe the positioning (score) of the municipal districts when actual entry data were used (data used as ideal for the problem). Then, changes were made to the entry data under analysis, and the new positioning of those municipal districts was observed. It should be emphasized that the entry data for this analysis are the comparisons among the indicators mentioned in AHP, however the output is not simply the result of AHP, i.e., the weights generated by this process for the indicators (ID), (DHS), (EPR) and (URPB) – the output is the positioning of the municipal districts, using the score standard based on the weights generated by AHP.

As the application of this process seeks to select ten municipal districts for an in-depth analysis of their characteristics, as described below, the sensitivity analysis made was intended to observe the behavior of the top ten municipal districts listed (municipal districts chosen using data deemed as ideal) due to the variations in entry data. This variation (related to the relevance comparison between the indicators ID and DHS) is intended to cover intervals in which the consistence ratio assumes values considered acceptable (not superior to 0.1 according to Saaty (*apud* Ehrlic,2004)). The values for this analysis are shown in Table 4. In this table, the levels of relevance presented as “actual” are those that were in fact used to solve problem. During the analysis, both services (health and retirement) were considered as having the same relevance.

**Table 4** Comparisons among attribute pairs used in the sensitivity analysis

| Relevance comparison of the ID indicator vs. DHS | | | |
|---|---|---|---|
| **Health service** | | **Retirement service** | |
| **Relevance levels** | **Consistency ratio generated for the service** | **Relevance levels** | **Consistency ratio generated for the service** |
| 1/5 | 0,12 | 5 | 0,09 |
| 1/3 (actual) | 0,08 | 7 | 0,05 |
| 1 | 0,08 | 9 (actual) | 0,03 |

As a result, it is possible to see that none of the used values caused any change in the choice of the top ten municipal districts. In each of the verification, the same municipal districts were listed as selected cities, still maintaining the same rank. Figure 2 shows a complementary view of this analysis. In this illustration, the top 20 municipal districts selected are shown. Then, it is possible to establish from which position some of those municipal districts are excluded from this group of 20 cities.

In Figure 2, each bar of the chart represents a selected municipal district. The chart that shows the ID/DHS list as being the "actual" is the one that contains the list of municipal districts initially chosen. Changes are made to the value of the relevance ratio between indicators ID and DHS for the health service (Figure 2 (a)) or for the retirement service (Figure 2 (b)). Then, it is possible to observe the displacement of those 20 municipal districts initially selected. It should be emphasized that this illustration is intended to show only the exclusion of the municipal districts initially selected among the top 20 positions. The new positions of those municipal districts among the initial placements are not analyzed. It can be seen that the worst case is the exclusion of two municipal districts from the list of selected municipal districts (Figure 2 (a); ID/DHS=1/5), and the list change only starts from the 19th position.

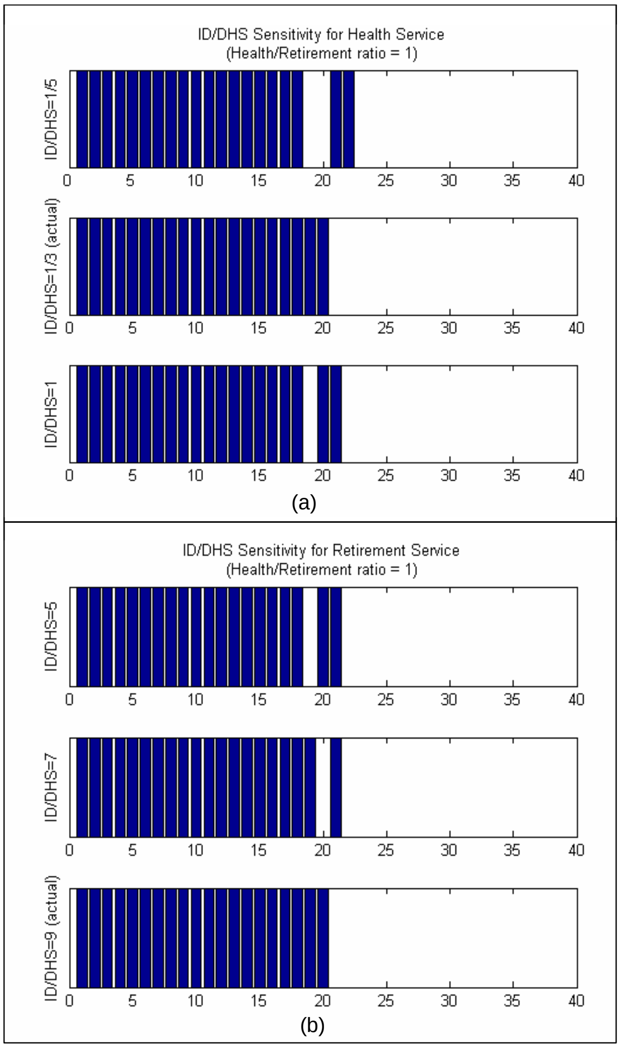


**Figure 2** Sensitivity analysis to the relevance variation between ID and DHS indicators: (a) for the health service; (b) for the retirement service

It is important to remember that the analysis by simultaneously varying the relationship between the indicators of both services is sought. The result obtained is shown in Figure 3, where it is possible to see that the behavior regarding the simultaneous variation causes no significant changes in the final result of the selection. Only the exclusion of one municipal district from the 19th position has occurred.

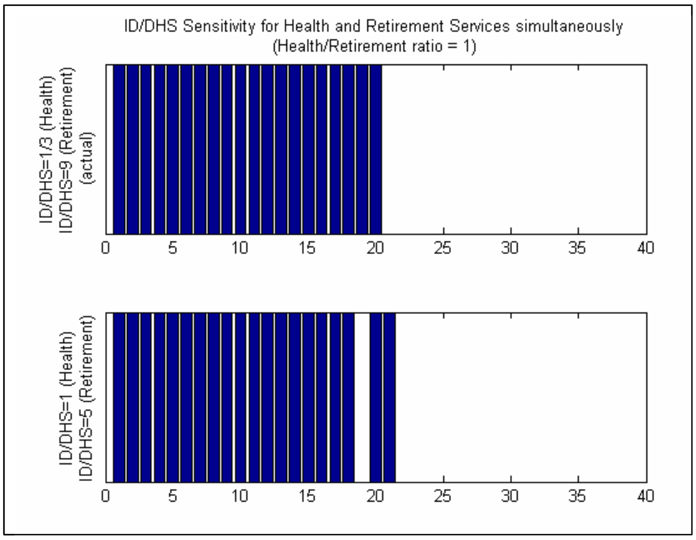


**Figure 3** Sensitivity analysis regarding the relevance variation between ID and DHS indicators for the health and retirement services simultaneously

From this analysis it is possible to find that the process used to establish the weights, based on the AHP procedure, although using subjective entries, presents no high-sensitivity variable. Although this low sensitivity is desirable, this can also mean a lack of discriminant power, thus compromising the reliability of the procedure used. To test this property (discriminant power), another test was performed in which the focus of municipal district selection was the indicator “balance between rural and urban populations” (URPB) instead of the “illiteracy density” (ID).

The values for comparing the relevance among attribute pairs (services and indicators) used in this new analysis are shown in Table 5. The result of the sensitivity analysis is presented in Figure 4.

**Table 5** Comparisons among pairs of attributes (services and indicators) used to define the weights (emphasis on the URPB indicator) ((a), (b) and (c))

(a) comparison of the relevance among services

| **Comparison of the relevance among services** | |
|---|---|
| **Comparison** | **Level of relevance** |
| Health service vs. retirement service | 1 |

(b) comparison of relevance among indicators according to the health service viewpoint

| **Comparison of relevance among health service indicators** | |
|---|---|
| **Comparison** | **Level of relevance** |
| ID vs. EPR | 5 |
| ID vs. URPB | 1/3 |

| | |
|---|---|
| DHS vs. ID | 7 |
| DHS vs. EPR | 9 |
| DHS vs. URPB | 3 |
| URPB vs. EPR | 7 |

(c) comparison of relevance among indicators according to the retirement service viewpoint

| **Comparison of relevance among retirement service indicators** | |
|---|---|
| **Comparison** | **Level of relevance** |
| ID vs. DHS | 7 |
| ID vs. EPR | 3 |
| ID vs. URPB | 1/3 |
| EPR vs. DHS | 3 |
| URPB vs. DHS | 9 |
| URPB vs. EPR | 5 |

Figure 4 illustrates the exclusion behavior of the municipal districts initially selected (emphasis on ID), among the first ten positions, when the application focus is changed (emphasis on URPB). It is assumed that, when the application focus is the "balance between rural and urban populations" (URPB), exclusions begin to happen in the initial list of selected municipal districts, from the second position on, and in this example the total exclusions is of 40%.

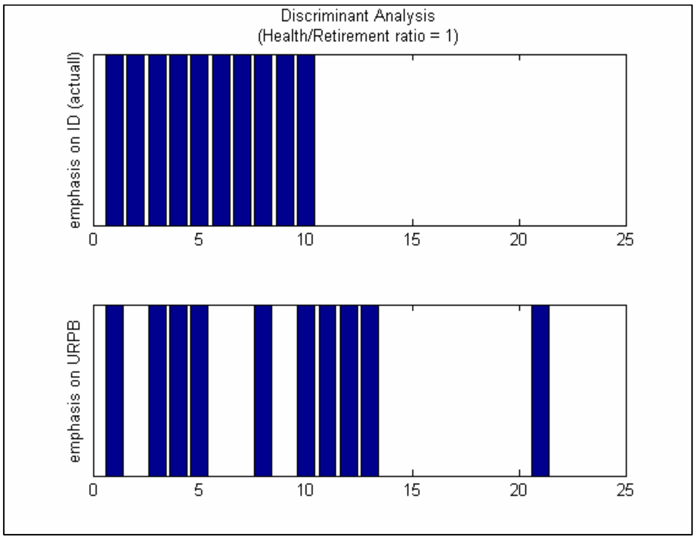


**Figure 4** Analysis of discriminant power for the method used in relation to the change in the indicator focus of the application

Additionally to Figure 4, observing the list of the top ten municipal districts as a relevant factor, only the first municipal district remained in its original position, i.e., there was a change in 90% of the positions of municipal districts associated to the change of focus for the application.

Based on the performed tests, changing the analysis focus (relationship between ID and DHS), it is possible to conclude that the method used had a significant discriminant power. This property is also joined to a high robustness.

## 5 Conclusions

The success of the service assessment in a real test environment will first depend on the number of persons – belonging strictly to the target audience – who will frequent the test locations and use the services developed. The high number of Brazilian cities and the discrepancies of the target audience profile for each service make the decision-making process more complex. While only a couple of cities must be chosen out of 5,560, the method has to consider cities where potential users for both services are easy to find. To meet this condition, a multi-criteria decision-making process was employed.

Taking into account the multifaceted nature of the experiment, it is essential to apply a method that is robust to slight variations in its inputs – due to the particularities of the subjective assessment for each criterion – and able to properly differentiate the locations according to the criteria and their weight. In this context, the AHP has proved to be effective in dealing with the problem posed in this paper, behaving either as robust or discriminant method.

In addition, the use of the AHP in association with a score standard based on the Euclidian distance (instead of an additive equation) has allowed setting the priority of the cities, balancing the values of the indicators. This procedure has prevented that a combination of indicators with extremely high values and indicators with extremely low values, for a given city, would result in a score higher than a score obtained from a balance among the indicators.

In short, this method has proved to be useful for a preliminary selection of cities with conditions to be test locations for the STID developments. A shortened amount of selected cities (ten) will make the next assessment stage easier, which will be based on *in loco* visits aiming at choosing two districts out of ten. The sensitivity analysis suggests that this method may be applied to select other test sites for other types of digital inclusion services, since taking into account the same priority criteria even though its relative weights are different. Nevertheless, if the new services require an analysis based on other selection criteria, new sensitivity analyses should be carried out to assess the robustness and the discriminant power.

## Ackowledgement

This work was supported by FUNTTEL – Fundo para o Desenvolvimento Tecnológico das Telecomunicações

## References

Ehrlich, P. J. (2004) “Procedimentos para Apoio às Decisões”. Available at: <www.fgvsp.br/academico/professores/Pierre_J_Ehrlich. 2004>. Accessed on Dec 12, 2006.

Saaty, T. L. (1980) *The Analytic Hierarchy Process – Planning, priority setting, resource allocation,* McGraw-Hill, Inc.

Holanda, G. M. and Dall'Antonia, J. C. (2006) “An approach for e-inclusion: bringing illiterates and disabled people into play”, *Journal of Technology Management & Innovation*, Vol 1, 3, 29-37.

Pinto J. C. L.*et al*. (2007) Identificação de Locais de Testes. Projeto Soluções de Telecomunicações para Inclusão Digital. Version AD. PD.30.12.36A.005A/RT-02AD. Campinas: CPqD-Funttel.