*Information Technology & Tourism,* Vol. 6 pp. 00–00
Printed in the USA. All rights reserved.

1098-3058/04 $20.00 + .00

www.cognizantcommunication.com

# INTELLIGENT ONE-STOP-SHOP TRAVEL RECOMMENDATIONS USING AN ADAPTIVE NEURAL NETWORK AND CLUSTERING OF HISTORY

MANOLIS WALLACE,* ILIAS MAGLOGIANNIS,† KOSTAS KARPOUZIS,*
GEORGE KORMENTZAS,† and STEFANOS KOLLIAS*

*Image, Video and Multimedia Systems Laboratory, Department of Electrical and Computer Engineering, National Technical University of Athens
†Department of Information and Communication Systems Engineering, University of the Aegean

The rapid growth of e-commerce during the last years has obliged a significant number of companies and professionals from diverse fields to turn to the Internet as a medium through which they aim to promote their products and services. A main issue for product and service providers is that, as this new market is characterized by the lack of personal contact, it is difficult to offer personalized services to end users; it is this type of service that end users look for and remain faithful to. Recommender systems belong to a new breed of software that aims to fill this gap; they rely on the analysis of past user actions to estimate the optimal way with which to interact with each user. In this article we explain why existing recommender systems are not adequate to provide for efficient personalization of interaction in the area of travel services, as they cannot support the user in all the phases of travel planning, and propose a new scheme to overcome the identified difficulties. Our approach considers the relation between different types of services in the usage history of the system. It is based on a hierarchical clustering of usage history to extract meaningful usage patterns, as well as an adaptive neural network structure that allows for online adaptation to the user, and enables the offering of intelligent recommendations.

Key words: Recommender systems, Travel planning, Collaborative user modeling, Hierarchical clustering, Neural network, Linear adaptation

## Introduction

As is the case with other aspects of economic and social life, tourism-related services are nowadays being made available through the Web. The majority of tourism/travel suppliers (particularly airlines, shipping lines, car rentals, and hotel chains) provide Web pages that incorporate useful supplier-based information. Many of these sites allow potential customers to directly access the supplier's reservation

Address correspondence to Ilias Maglogiannis, University of the Aegean, Department of Information and Communication Systems Engineering, 83200 Karlovasi, Greece. Tel: +30-2273-82239; Fax: +30-2273-82009; E-mail: imaglo@aegean.gr

system [e.g., British Airways (http://www.britishairways.com), Avis (http://www.avis.com), Marriott Hotels (see http://www.marriott.com), etc.].

A significant number of travel agencies have also obtained Web presence (e.g., see http://www.thomascook.com, http://www.lunnpoly.com, etc.), while a number of Web-based virtual travel agencies have also emerged (e.g., Expedia.com, ebookers.com, Travelocity.com, etc.); many travel agencies' sites drive to external reservation systems operating as intermediates to them (e.g., see http://www.opodo.com, http://www.orbitz.com, etc.). Well-known Internet portals (e.g., Yahoo, Altavista, Excite, etc.) provide travel/tourism content mainly through links in external sources (e.g., Web-based agencies and suppliers, etc.). Finally, several tourism/travel destinations have developed customized Web sites providing useful information (e.g., restaurants, museums, sights, etc.) for potential visitors (e.g., http://www.tiscover.com, http://www.holland.com, etc.).

As can be seen, the basic tourism/travel players have entered dynamically into the Web market, gaining a significant portion of its penetration to the end user (O'Connor & Frew, 2000). However, their share compared with traditional tourism/travel market still remains very low, less than 2% (Maglogiannis, Kormentzas, & Panagiotarakis, 2003). In order to focus on a greater market portion, online tourism/travel information systems need to incorporate some of the characteristics that lead customers to prefer taking their business to traditional travel agencies; personalization of services and expert recommendations are probably the most important characteristics that Web-based tourism/travel systems are currently missing. This would turn online systems into a direct extension of personalized services offered by small travel agencies, where the client's preferences are learned through personal interaction (Delgado & Davidson, 2002) and the travel agent applies his/her knowledge and experience to propose possible options and to offer advice to the customer.

To offer such services in an automated manner, *user modeling* technology needs to be applied; it has already been successfully used in a variety of domains related to information access: *information retrieval* (Brajnik & Tasso, 1994; Kay, 1995), *filtering* (Balabanovic & Shoham, 1997) and *extraction* (Benaki, Karkaletsis, & Spyropoulos, 1997), *adaptive user interfaces* (Chin, 1989; Langley, 1999), and *adaptive Web sites* (Ardissono & Goy, 2000). Lately, it has started to be integrated with e-commerce systems to provide the advantage of returning clientele (Shafer, Konstan, & Riedl, 2001).

In fields such as information or multimedia retrieval, where plenty of user interaction is available to the system, and user preferences may be modeled through their thematic categorization, it is relatively easy to generate *user profiles* by monitoring user actions and then use them to customize offered services (Chen & Kuo, 2000). Cases exist, though, where the accumulation of information about a specific user may not be possible. For example, user interaction may be sparse and limited, the user might not always identify him/herself to the system, or user monitoring might be considered as an invasion to the user's privacy. In such cases, user personalization is not a trivial task. Moreover, cases exist in which the modeling itself of user preferences is difficult.

Tourism-related information systems fall within these last categories. Users typically access them sparsely, often without any form of authentication/identification, and little feedback is offered to the system concerning the user's satisfaction from its performance. Thus, specialized personalization approaches need to be applied in such systems (Paliouras, Papatheodorou, Karkaletsis, & Spyropoulos, 2002). Moreover, even when adequate feedback has been provided by the user, the way to analyze it and transform it to user preferences is not obvious. These are the reasons that existing travel information systems have not yet integrated any recommendation services and act simply as information and service brokers (Baundisch & Terveen, 1999; Franke, 2003; Kautz, 1998; Resnick & Varian, 1997; Soboroff, Nicholas, & Pazzani, 1999), with very few exceptions (Ricci, Arslan, Mirzadeh, & Venturini, 2002).

In this work we apply a collaborative approach to user modeling to overcome the problem of sparse and limited user interaction. Starting from the system's usage history (i.e., the logs of all transactions made), we base the analysis on travel plans, rather than single travel services; this allows for a more intuitive modeling of users and user preferences. Applying a properly modified hierarchical clustering algorithm we identify typical patterns that appear often in travel plans and use them to drive a neural network. This network is able to provide efficient, quick recommendations. The issue of online adaptation to each user is also considered.

The structure of the article is as follows. In the second section we present the analysis of the usage history of the system using the modified agglomerative clustering algorithm for the extraction of typical travel plans. This information is utilized for the initialization of a resource allocating neural network structure. Continuing, in the third section we explain how this structure relates to the architecture of an online travel-related e-commerce system and discuss its online adaptation to each tourist. Finally, the fourth section presents a descriptive simulation of the proposed methodology and the fifth section concludes the article.

## User Modeling

Although systems that are aimed to assist users in retrieving and selecting information and services have existed for years, research in the field still remains open. The reason is that the identification of the user's wishes has proven to be a difficult thing to achieve. Initially, the need for adaptation of the system's operation to its users was handled with *stereotypes*; this approach assumes that differences among users rely solely on their different levels of expertise. Thus, asking users to answer a few simple questions could rapidly provide all the information required to classify them to distinct stereotypes, each one enjoying different information services.

Modern information systems are equipped with the ability to monitor and analyze users' actions, to determine how to best interact with them. Ideally, each user's actions are logged separately, thus forming the individual usage history. Continuing, this is analyzed to generate an individual user profile. The latter contains all the information about a user, extracted either by merely monitoring user actions or by considering the objects the user has evaluated (Burke, 2002), and is utilized to customize offered services. This user modeling approach is known as *content-based learning*. The main assumption behind it is that a user's behavior remains unchanged through time; therefore, the content of past user actions may be used to predict the desired content of future actions as well.

A different user modeling approach is that of *collaborative learning* (Varian, 1996). The main assumption behind this approach is that similar users will react in a similar manner, when faced with similar situations. This assumption permits the formation of user groups (Fig. 1); users of the same group may be assigned the same user profile and served in the same way. Although this approach is typically not ideal for customizing tasks such as Web retrieval, it is most suitable for cases in which limited feedback is available from the users. The fact that, irrelevant to the number of actual users, the number of user groups remains unchanged makes it possible to statically define the way in which users from each group should be served, or combine feedback from numerous members of a group in order to extract user profiles.

As most people use tourist recommendation and booking systems sparsely, information that may be

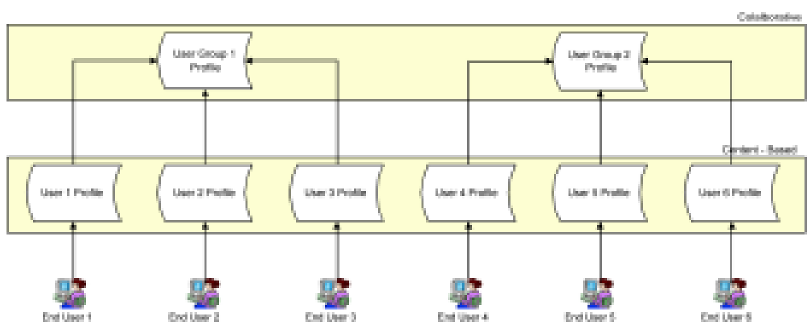


*Figure 1.* Content-based versus collaborative learning.

acquired by logging their actions is in most cases far from sufficient for the generation of individual user profiles. In addition, travel agencies book tickets on behalf of numerous and different clients, which makes it difficult to extract reliable personalization information using their booking records. Thus, it is not possible to adapt a content–based learning or a stereotype–based approach in an information system that aims to facilitate acquisition of travel/tourism information and booking of tickets/services; a collaborative approach needs to be applied. In order to do so, a number of questions have to be answered:

1. Which are the examined personalization elements? In other words, which elements in the usage history (in the logs of the system) contain useful information and may be utilized for the extraction of the user model.
2. How many are the distinct user groups that exist? To be more exact, the question to answer is how many different typical patterns exist; the same user may display more than one pattern in his/her behavior.
3. How are examined elements mapped to user groups? In other words, how are the typical patterns defined with respect to the characteristics of possible user actions?

Once these have been answered, the application of the collaborative learning approach to user modeling becomes similar to the application of a content–based approach (which, having received more attention by researchers in the past, is easier to solve.) We attempt to answer all these questions in the following.

Examined personalization elements are complete travel plans of the usage history. This is a major innovation of the proposed approach, as existing e-commerce recommender systems consider each part of the user interaction to be independent; existing approaches, due to the independent way with which they handle distinct selections of the end users, are only able to recommend single services. The proposed approach, relying on complete travel plans to extract usage patterns, is able to assist users in building their own complete travel plans. In other words, it is able to offer intelligent and justified (based on the usage history) recommendations for all travel/tourism-related services that may be available online, thus offering more of the services end users look for.

When it comes to the consideration of user groups, alterations to the classical approach are again needed. Typically, each user group is characterized by some behavior, and each user is mapped to a single user group and follows the corresponding behavior. Unfortunately, although easy to implement, this approach does not have the intuitive merits one might wish. We would prefer to have a system that is aware of all possible behaviors and dynamically relates each user with one or more of them, depending on the situation. Therefore, instead of detecting distinct user groups, in the proposed approach we detect distinct usage patterns. As has already been mentioned, each one of these patterns is related to the formulation of a complete travel plan, not just to the selection of a single travel/tourism service.

### *Clustering of Usage History*

As already mentioned, the usage history is the set of all logged user actions, and more specifically the set of all travel *services* purchased through the system. Our approach relies on analyzing logs of services purchased within the same Web session. Thus, we consider the *travel plan* (i.e., the set of all services purchased by a user in a single session) as the elementary article in the usage history. For example, a transport service from Athens international airport to the Athens Hilton Hotel is a service; this transport together with accommodation at the hotel for 3 nights and plane tickets from Heathrow airport to Athens airport form a complete travel plan. The aim is to group travel plans together based on some sort of similarity; groups of similar plans will describe a typical generic travel plan, which is information that can be utilized for the offering of intelligent recommendations.

As the count of distinct patterns that characterize the travel plans in the usage history is not known before hand, a hierarchical clustering needs to be applied on the usage history to extract the underlying patterns (Theodoridis & Koutroubas, 1998); for the application of such an algorithm one needs to first define/select a similarity (or dissimilarity) measure among clusters (groups) of the input elements. So, in our case, we need to define a distance measure among groups of travel plans. These are in turn based on the definition of distances among distinct travel plans. We elaborate on both in the following.

To define a distance metric among any couple of travel plans we start from the following, rather obvious, remark: two travel plans, possibly created by different users and at different times, are similar to the degree that the services they comprise are similar. Thus, to compare two travel plans we start by comparing their components.

Each one of them, depending on the type of service it represents, can be mapped to a feature vector of some finite dimension. Feature values may be related to type of service (e.g., guide service, transport service, lodging service), duration, relative price (cheapest available, cheaper 20%, average), etc. Each feature is normalized in the range, to facilitate the definition of similarity and distance metrics among services and plans.

Based on each one of these features, two services may be compared and assigned a distance value; depending on the case, one of these values, or a combination among them, may be the distance that best describes the degree to which the two services are related.

For two services $S_1$ and $S_2$, the following distance may be defined when considering feature *i*:

$$Dist_i(S_1,S_2)=\left|S_1^i-S_2^i\right|$$

where $S_1^i$, $S_2^i$ are the *i*th features of the services. For example, consider the services in Table 3. Each service is represented by a position in the three-dimensional space. The three axes correspond to relative price, time of year, and geographical location. All values range from 0 to 1. As far as relative price is concerned, 0 implies absence of the service, 0.5 the cheapest service (of this type) available, and 1 most expensive available. Time of year assumes value 0 for summer time and 1 for winter time. Depending on the exact data, intermediate values are also possible. Finally, geographical location takes values in {0,1}, each one corresponding to one of the two assumed destination islands (a small but simple variation of the distance metric is needed in the case that more than two destinations are assumed, to allow for the distance between any two distinct destinations to be the same, regardless of their mapping on the {0.1} interval). When comparing services 3 and 4 to each other, their distance is |0.78 – 0.88| = 0.1 if the feature considered is the location. If, on the other hand, we use the geographical location their distance is 0.

When it comes to two complete travel plans, $P_1$ and$P_2$, their distance may be described by the average distance among all of the services they comprise:

$$Dist_i(P_1,P_2)=\frac{\sum_{S_1\in P_1,S_2\in P_2}Dist_i(S_1,S_2)}{|P_1|\cdot|P_2|}$$

Finally, when comparing two groups (clusters) of plans, $C_1$ and $C_2$, we use the following metric:

$$Dist_i(C_1,C_2)=\frac{\sqrt{\sum_{P_1\in C_1,P_2\in C_2}\left(Dist_i(P_1,P_2)\right)^2}}{|C_1|\cdot|C_2|}$$

with respect to each feature *i*. The overall distance between two clusters $C_1$ and $C_2$ (i.e., their distance when considering all features) is defined as

$$Sim(C_1,C_2)=1-\frac{\sum_{i=1}^{F}\left(\frac{1}{F}+[t_i]^{\lambda}\right)Dist_i(C_1,C_2)}{1+\sum_{i=1}^{F}[t_i]^{\lambda}}$$

where vector *t* is a weighting of features, λ is a predefined constant, and *F* is the count of features. The $\frac{1}{F}$ component has been added to the weighting to assure that all features participate in the overall rating of the similarity. Thus, depending on the weighting *t* that is used, a different degree of similarity among services, plans, and clusters is assumed. The two clusters are considered similar if any *context* (i.e., any combination of features) exists via which the overall distance is small. Thus, the minimum value of Sim($C_1$,$C_2$) is an estimation of the similarity of the clusters.

When $\lambda = 1$ it is easy to see that the weighting that produces the best similarity is the one that promotes the feature(s) for which the distance among clusters is smallest. When $\lambda \neq 1$ and features exist for which the distance among the clusters is zero, then the optimal weighting is the one that promotes exactly those features. In the most typical case, though, such features will not exist. Typically, for every $i \in 1 \ldots F$, plans exist in the groups such that $\mathrm{Dist}_i(P_1,P_2) \neq 0$. For this case, it is proven in Wallace and Kollias (2003) by demanding that $\lambda \neq 1$ and that the first derivative is zero, that the optimal weighting $t$ is given by:

$$t_i = t_1 \left[ \frac{Dist_1(C_1,C_2)}{Dist_i(C_1,C_2)} \right]^{\frac{1}{\lambda-1}}, i \in 2 \ldots F$$

$$and\ t_1 = \frac{1}{\sum_{i=1}^{F} \left[ \frac{Dist_1(C_1,C_2)}{Dist_i(C_1,C_2)} \right]^{\frac{1}{\lambda-1}}}$$

Because the optimal weighting of the features (i.e., the optimal context) is calculated analytically, the estimation of the similarity between two services has the ideal complexity of $O(1)$. Thus, the process of comparison of all clusters to each other has a polynomial complexity of $O(n^2)$, as exactly $n(n-1)/2$ comparisons are needed, where $n$ is the count of all clusters.

The generic agglomerative clustering algorithm (Theodoridis & Koutroubas, 1998; Yager, 2000), which is the most important category of hierarchical clustering algorithms, initializes by considering each element as a cluster; in this case elements are travel plans in the usage history. Continuing, all clusters are compared to each other, as described above, to detect the two that are most similar. These two are merged, and the algorithm continues iterating, until a threshold on the similarity among clusters to be merged is met.

Small clusters (i.e., clusters that contain too few elements) do not correspond to usage behaviors that users typically demonstrate. In other words, travel plans that do not appear often in the logs of the system are not expected to be selected often by future tourists either. Thus, they are ignored as trivial. Each one of the remaining clusters, although containing various travel plans due to the way these plans were selected and brought together, corresponds to a typical travel plan (i.e., a typical tourist selection). It is this information that the intelligent recommender shall use to select the set services to propose to each user.

*Extraction of Patterns and Mapping Onto a Neural Network*

Each one of the clusters the agglomerative process produces contains numerous distinct travel plans. Each one of them comprises a set of services of different types, and quite probably with different feature values as well. For example, a group of users might prefer to use cheap transportation with luxuries accommodation.

To allow for such diversities to survive the process of pattern extraction, each type of service is analyzed independently when clusters are analyzed. Thus, if cluster $C$ contains services of different types, then the following center and standard deviation values can be used to fully characterize the pattern it describes:

$$\mu_{ij} = \frac{\sum_{S \in C \cap^j S} S^i}{|C|},\ \sigma_{ij} = \sqrt{\frac{\sum_{S \in C \cap^j S} \left(S^i - \mu_{ij}\right)^2}{|C|}},\ j = 1 \ldots q$$

where $q$ is the count of different types and $j_s$ is the set of all services of type $j$. Each selection a tourist makes in the future can be compared to these to detect which typical travel plans the user's choice resembles. Once these have been identified, other services contained in them can be utilized to form a reasonable recommendation to the user.

Two problems that may be identified in this process are the following:

- Serial comparison of the service(s) selected by the tourist to each one of the detected typical travel plans may lead to delays in a large-scale system.
- The system cannot adapt its operation to each user, as the centers and spreads of the clusters are shared among all users.

To tackle these problems, the comparison of services selected by a tourist and typical travel plans can be performed using a properly initialized *re-*

*source allocating neural network*, such as the one presented in Figure 2. This is a structure that allows for parallel implementation for rapid responses and online adaptation to each user via small perturbation of its parameters based on feedback acquired from the user.

Each detected cluster (i.e., each typical travel plan) is assigned one output layer node. This node will be activated every time the services selected by the user match this typical travel plan. For each type of services in the cluster, a different hidden layer node is initialized with position and spreads equal to the center and standard deviation that characterize the type of service in the cluster, as described above. Each hidden layer node is linked to the corresponding output layer node with a degree of one. Finally, out of the wide range of possible choices, we choose the sigmoid function for output layer nodes, as a nonlinear discriminant. It is worth noting that two of the detected typical travel plans may contain similar services. In this case, two similar but distinct hidden layer nodes will represent them, each one linked to exactly one output layer node. More on the structure, initialization, and learning formulas of this network can be found in Tsapatsoulis, Wallace, and Kasderidis (2003).

Together with the network parameters (positions, spreads, and activation values), a training set needs to be stored as well, to provide for efficient retraining of the network. As the network will be required to operate at online mode, the utilization of the complete usage history for retraining purposes is not possible; thus, a small set of representative samples is kept in the usage history.

## Intelligent Travel Recommendations

The proposed methodology is integrated in an online travel recommendation system as shown in Figure 3. Thus, all available history is analyzed using the methodology of the above section for the initialization of the network structure that forms the Intelligent Adaptive Recommender. This accepts as input the features that characterize an available service; the activation of the output nodes indicates to which travel behaviors this service is related. This way, service compatibility can be defined through the degree to which the network indicates that they are related to the same typical travel plans; this is the concept that drives the intelligent recommendation engine.

Specifically, when a user connects to the system, no information about him/her is available. Thus, no

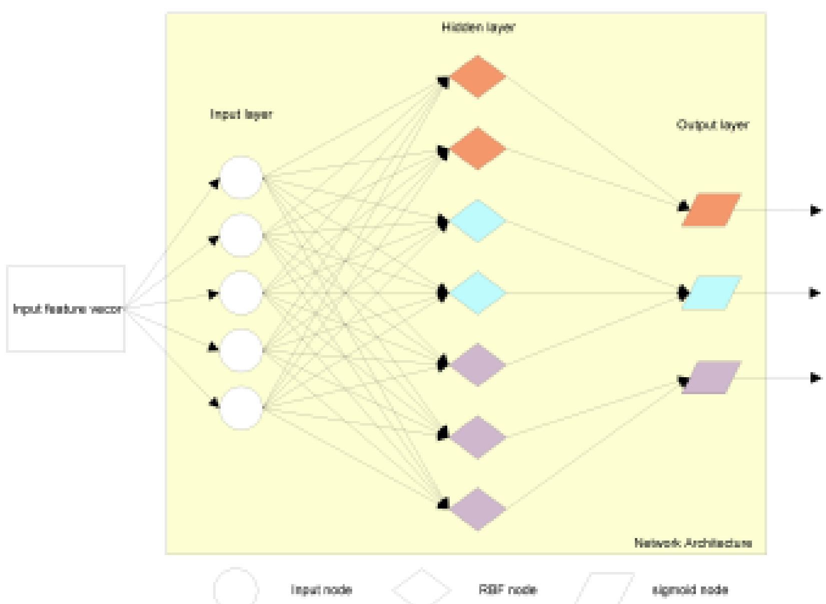


*Figure 2.* Architecture of the RAN.

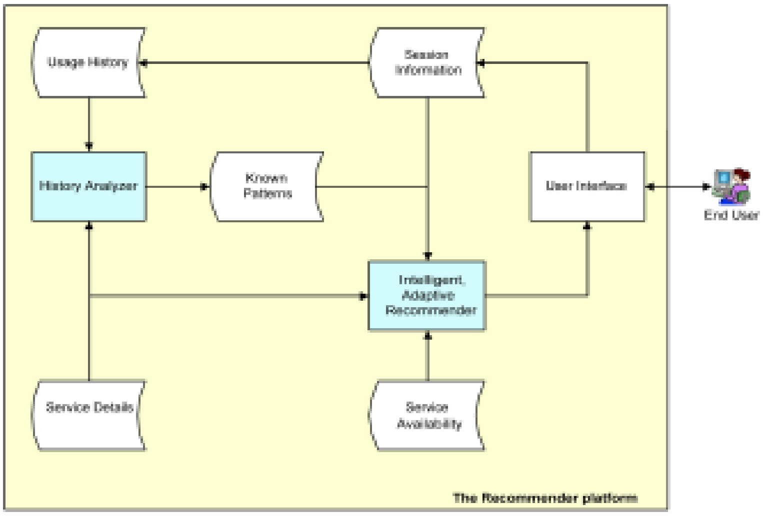


*Figure 3.* System architecture.

intelligent recommendation can be made; default options are presented. As the user forms an initial selection of a single travel service, typically tickets to some travel destination, this service is fed to the network as input. Output layer nodes that are activated to a degree that surpasses a predefined threshold are considered as possible typical travel plans that the user may be trying to form. This information is utilized in two ways.

Firstly, it is used to generate a personalized version of the network. Specifically, all output and hidden layer nodes that are not related to the activated ones are ignored, leading to the creation of a much smaller network. The fact that the network after the first user interaction can be limited in size in this way makes it possible for this adapted network to be dedicated to the specific user. Thus, it may be used for online processing of data with the aim of offering intelligent recommendations to the specific user.

Secondly, it is used to form the intelligent system recommendation to the user. Specifically, for each type of service, all candidate services are fed to the personalized version of the network. Each one of these services will activate one output layer nodes that have been kept if it is related to the typical travel plan that corresponds to it. In other words, the degree of activation of the output layer nodes can be used to perform a first ranking of the services, as services that activate the output layer are services that might be used by the user to continue forming a travel plan the system would consider as typical and thus probable.

Service availability is an optional input of the recommender. Using it, the system does not propose to its end users services that may not be available. Moreover, it promotes services that have received little attention, as recommendations do not only offer solutions to user needs, but also often participate in the formulation of user needs (users that were not aware of the service may decide they are interested in it after it has been recommended to them).

Following this process, the set of selected services is presented to the tourist as a system recommendation. The tourist may now decide to include one of them in the travel plan. Similar to the first choice, this comprises valuable information that may be used to improve the system's performance throughout the remaining of the session. This is accomplished in two ways.

Firstly, the service that is selected further specifies the behavior of the end user. Thus, output nodes that are activated by it continue to represent possible user behaviors, while the ones that are not activated (together with the corresponding hidden layer nodes) may be removed. This way, as the user selects more services to include in a plan, the user intention becomes better specified and thus the network becomes smaller and more efficient computationally wise.

Secondly, information about the services that were presented and selected by the tourist, as well as information about services that were selected but ignored, can be utilized to generate a neural network structure that is further adapted to the user. Thus, we may further refine the network, demanding that the selected service activates at least one of the output layer nodes, and that the services not selected do not activate any of the output nodes.

On the other hand, it is well known that classical neural network adaptation is a time-consuming process that may need numerous epochs to conclude; training data have to be fed to the network iteratively numerous times before the network parameters are properly adjusted to produce the desired output. Of course, this is not acceptable for a network that has to adapt to its new environment in an online mode of operation. An approach has to be followed that will allow for computationally efficient adaptation.

In our case this can easily be achieved if we consider that adaptation to a specific user will only cause a minor perturbation of the network parameters. With this in mind, although the output layer functions are nonlinear, we may consider them as “almost” linear in the neighborhood of the initial and adapted parameter values. This allows for a linearization of the retraining process, which can be concluded in only one epoch (Doulamis, Doulamis, & Kollias, 2000).

As has been said, the system keeps in the usage history a set of services, together the corresponding class (matching typical pattern), included in the following set $S_b = \{(\underline{S'}_1, \underline{d'}_1), \ldots, (\underline{S'}_{m_b}, \underline{d'}_{m_b})\}$, where vectors $\underline{S'}_i$ and $\underline{d'}_i$ with $i = 1,2, \ldots, m_b$ denote the $i$th input and corresponding desired output vectors. The network should have the ability to adapt its performance, taking into account both this former knowledge and the current formation of a travel plan.

Let us consider network adaptation in more detail. Let vector $\underline{w}_b$ include all adaptable network parameters before adaptation, and $\underline{w}_a$ be the new parameter vector, which is obtained through adaptation. A retraining set, $S_c$, is assumed to be extracted from the travel plan that the user forms, composed of, say, $m_c$ feature vectors; $S_c = \{(\underline{S}_1, \underline{d}_1), \ldots, (\underline{S}_{m_c}, \underline{d}_{m_c})\}$. Adaptation is performed using efficient network training ; new network parameters $\underline{w}_a$ are computed, by minimizing the following error criterion with respect to the parameters:

$$E_a = E_{c,a} + E_{f,a}$$

$$E_{c,a} = \frac{1}{2}\sum_{i=1}^{m_c} \left\| \underline{y}_a(\underline{S}_i) - \underline{d}_i \right\|_2$$

$$E_{f,a} = \frac{1}{2}\sum_{i=1}^{m_b} \left\| \underline{y}_a(\underline{S}_i) - \underline{d'}_i \right\|_2$$

where $E_{c,a}$ is the error performed over the retraining set $S_c$ (“current” travel plan), $E_{f,a}$ the corresponding error over the original “training” set (“former” knowledge); $\underline{y}_a(\underline{S}_i)$ and $\underline{y}_a(\underline{S'}_i)$, corresponding to the feature vectors of services $\underline{S}_i$ and $\underline{S'}_i$ respectively, are the outputs $\underline{y}(\underline{S}_i)$ and $\underline{y}(\underline{S'}_i)$ of the (adapted) network consisting of parameters $\underline{w}_a$. Similarly $\underline{y}_b(\underline{S}_i)$ would represent the output of the network, consisting of parameters $\underline{w}_b$, when accepting service $\underline{S}_i$ at its input. Parameter $\eta$ is a weighting factor accounting for the significance of the current retraining set compared to the former one and $\|\cdot\|_2$ denotes the $L_2$-norm.

At the end of the session all information about user selections is moved into the usage history. Thus, when plenty of new records have been inserted into the usage history, the latter may be analyzed again, producing a richer representation of typical tourist selections. The history analyzer uses the service details (i.e., the feature vectors of the services) to cluster travel plans in the usage history. The nontrivial detected clusters correspond to typical plans that users form often; they are stored, in the form of centers, spreads, and activations, as known patterns. A small set of representative members is also identified for each cluster/pattern and stored together with it; these data are utilized during the retraining of the neural network.

## Simulation Results

The overall system presented in the previous subsection comprises a variety of methodologies and subsystems, including context–adaptive agglomerative clustering, initialization of resource allocating neural networks, and utilization of neural networks to map tourist selections. To verify the simulation of the proposed system and to verify the effectiveness of the proposed methodology, we populated a synthetic database of travel-related services with three types of services, namely ship tickets, lodging, and transport from the port to the hotel or apartment.

Services, together with their ID, are represented by a position in the three-dimensional space. As has already been mentioned, the three axes correspond to relative price, time of year, and geographical location. All values range from 0 to 1. As far as relative price is concerned, 0 implies absence of the service, 0.5 the cheapest service (of this type) available, and 1 most expensive available. Time of year assumes value 0 for summer time and 1 for winter time. Depending on the exact data, intermediate values are also possible. Finally, geographical location takes values in {0,1}, each one corresponding to one of the two assumed destination islands (a small but simple variation of the distance metric is needed in the case that more than two destinations are assumed, as to allow for the distance between any two distinct destinations to be the same, regardless of their mapping on the {0,1} interval). Clearly, a tourist would need much more information to make a decision; these few features are kept here for the sake of clarity, but the methodology easily generalizes to richer service descriptions.

To initialize the experiment, the usage history had to be populated. A Gauss–based generator was constructed and used to generate random travel plans, with probabilities being highest for three typical travel plans; these are presented in Table 1. These were selected as three quite realistic user behaviors; they correspond to users that look for cheap holidays year round and in any location, users that enjoy luxurious holidays during the summer time and mainly in one of the two possible destinations, and to users that prefer save on their travel budget in order to have the option to spend some more during their stay, again in one of the two possible destinations.

Obviously, the location feature always assumes values in {0,1}. Value 0.5 in Table 1 indicates the generation of services for the two locations with equal probabilities. Additionally, values in the (0,0.5) for relative price were rounded to one of the two extremes, and values over 1 and below 0 were rounded to 1 and 0, respectively, for all features. Moreover, the same geographical location and time of year was applied to all services in the same plan. One hundred and fifty plans were generated and inserted in the usage history, 50 for each one of the above-mentioned typical behaviors. The analysis of usage history, via clustering of logged plans, successfully detected three nontrivial clusters; two clusters, of 2 and 3 travel plans, respectively, were ignored as trivial. The centers and spreads of the services of the nontrivial clusters of plans are presented in Table 2. The center and spread combinations can be used to verify that the detected clusters correctly represent the typical behaviors that were induced in the usage history.

Continuing, we utilized these to initialize an RBF neural network with 4 input layer nodes, 8 hidden layer nodes, and 3 output layer nodes; each output layer node corresponds to one of the detected typical plans, each hidden layer node to one of the detected typical services (one of the travel plans only has two hidden layer nodes, as it does not contain a transport service), and each input layer node corresponds to one of type of service, relative price, time of year, and location.

To test the performance of the system, we randomly generated descriptions of ship ticket services.

Table 1
The Typical User Behaviors Induced in the Sample Usage History.

| Behavior | Travel Tickets | Transport | Lodging |
|---|---|---|---|
| 1 | 0.5,0.1–0.5,0.5–0.5,0.5 | 0,0–0.5,0.5–0.5,0.5 | 0.5,0.1–0.5,0.5–0.5,0.5 |
| 2 | 1,0.1–0,0.3–1,0 | 1,0.2–0,0.3–1,0 | 1,0.1–0,0.3–1,0 |
| 3 | 0.5,0.3–0.2,0.4–1,0.2 | 0.5,0.1–0.2,0.4–1,0.2 | 0.8,0.2–0.2,0.4–1,0.2 |

Services are represented as center-spread triplets.

Table 2
Detected Typical Plans, With Centers and Spread Rounded to One Decimal Position

| Plan | Travel Tickets | Transport | Lodging |
|---|---|---|---|
| 1 | 0.6,0.2–0.6,0.5–0.4,0.5 | 0,0–0.6,0.5–0.4,0.5 | 0.7,0.1–0.6,0.5–0.4,0.5 |
| 2 | 0.9,0.1–0.1,0.2–1,0 | 0.8,0.2–0.1,0.2–1,0 | 0.9,0.1–0.1,0.2–1,0 |
| 3 | 0.7,0.3–0.4,0.3–0.9,0.1 | 0.6,0.1–0.4,0.3–0.9,0.1 | 0.7,0.2–0.4,0.3–0.9,0.1 |

Specifically, we randomly generated triplets of values, making sure that they were valid relative price, period, and location values. These were supposed to be the feature vectors of the initial user selection, concerning ticket reservations. We fed each one of these triplets (with ticket as type of service) to the above-mentioned neural network and removed all nodes that did not contribute to a significant output layer activation. Continuing, we retrained the network, including the considered triplet in the training set as the current travel plan.

Available services were generated similarly to the initial user selections and fed to the resulting neural network. Services that activated the output layer to a high degree were constantly found to match the initial user selection, according to one of the typical user behaviors that were used to generate the usage history. As an example, in Table 3 we provide one of the sample initial selections and the available services that activated the corresponding network's output layer to a high degree.

The initial user selection, for which we present results in Table 3, is a ticket service that matches those of the second typical behavior that was induced in the usage history. We can see that all selected services also match the corresponding services in the same typical behavior. These were selected out of all the randomly generated services (feature vectors in all valid ranges were created and fed to the neural network).

Table 3
The Output of the Neural Recommender for a Specific User Selection

| Service | Type of Service | Feature Vector |
|---|---|---|
| Initial | ticket | 0.9,0.2,1 |
| 1 | lodging | 0.91,0.23,1 |
| 2 | transport | 0.87,0.18,1 |
| 3 | lodging | 0.88,0.22,1 |
| 4 | lodging | 0.78,0.30,1 |

This obviously corresponds to the case when services of all kinds are available. The results in Table 3 indicate that the proposed system is able to successfully and efficiently determine: 1) to which typical behavior a specific users' actions correspond, and 2) out of a large set of diverse available services, which ones best match a specific typical behavior. Combining these two characteristics of the system, it is easy to see that it comprises a very efficient recommender for personalized travel planning and booking, when we wish to follow a collaborative learning approach.

## Conclusion

This article presented a methodology for intelligent and adaptive travel recommendations. The main innovation of the proposed approach is that complete travel plans, rather than individual services, are the elementary considered particles of the usage history. Thus, the resulting system is able to provide intelligent recommendations to tourists looking for any kind of tourist service, and moreover provide intelligent recommendations about one kind of service based on other, different services already selected by the tourist.

Specifically, when the usage history is clustered through a modified agglomerative clustering algorithm, tourist behaviors related to the formulation of complete travel plans may be extracted from the clusters. This information is mapped onto a neural network structure, which allows for online recommendations; service availability may also be considered during such recommendations.

All available feedback from the user is utilized to adapt the system to the user. Specifically, positive feedback is utilized to filter the neural network, while negative feedback, together with positive feedback and historical data, is utilized to refine the network's parameters. This last part is pursued through linearization of the retraining of the network, to allow for quick response times.

Finally, we have presented a system architecture that is able to exploit the benefits of the proposed algorithms and methodologies and provided a few indicative preliminary simulation results.

## Biographical Notes

Manolis Wallace is with the National Technical University of Athens, where he serves as a researcher of the Image, Video and Multimedia Systems Laboratory, and with the University of Indianapolis, Athens Campus, where he serves as an Assistant Professor. He graduated from the National Technical University of Athens in 2001 and he is currently pursuing a Ph.D. degree at the same University. His research interests lie in the fields of data mining and intelligent information retrieval, in which he has published more than 14 peer-reviewed papers and has chaired a session at an international conference. He is a member of the IEEE, a reviewer for 3 Transactions and various other journals and conferences.

Ilias Maglogiannis is with University of the Aegean. He received a Diploma in Electrical & Computer Engineering and a Ph.D. in Biomedical Engineering from the National Technical University of Athens (NTUA) Greece in 1996 and 2000, respectively, with scholarship from the Greek Government. From 1996 until 2000 he worked as a Researcher in the Biomedical Engineering Laboratory in NTUA and he has been engaged to several European and National Projects. From 1998 until 2000 he was also the head of the computer department, responsible for the modernization of the Dept. of Civil Engineering in the National Technical University of Athens. Since February of 2001 he has been a Lecturer in the Dept. of Information and Communication Systems at the University of the Aegean. His published scientific work includes 4 lecture notes (in Greek) on Biomedical Engineering and Multimedia topics, 6 journal papers, and more than 20 national and international conference papers. He has served on program and organizing committees of national and international conferences and he is a reviewer for several scientific journals. His scientific activities include biomedical engineering, image processing, multimedia, and human–computer interaction.

Kostas Karpouzis was born in Athens, Greece, in 1972. He graduated from the Department of Electrical and Computer Engineering, the National Technical University of Athens in 1998 and received his Ph.D degree in 2001 from the same university. His current research interests lie in the areas of human computer interaction, image and video processing, 3D computer animation, and virtual reality. He is a member of the Technical Chamber of Greece and a member of ACM SIGGRAPH and SIGCHI societies. Dr. Karpouzis has published 5 papers in international journals and more than 15 in proceedings of international conferences. He is a member of the technical committee of the International Conference on Image Processing (ICIP). Since 1995 he has participated in seven research projects at Greek and European level.

George Kormentzas was born in Athens, Greece, in 1973. He received a Diploma in Electrical and Computer Engineering and a Ph.D. in Computer Science both from the National Technical University of Athens (NTUA), Greece, in 1995 and 2000, respectively. From 1996 to 2001 he was a research associate with the Institute of Informatics & Telecommunications of the Greek National Center for Scientific Research "Demokritos" working in a number of EU-funded research projects in the area of high-speed wired and wireless networks. Since 2001, he has been a lecturer in the University of the Aegean. He is a member of the Technical Chamber of Greece and WSEAS (World Scientific and Engineering Society). His research interests are in the fields of traffic analysis, network control, resource management, and quality of service in broadband networks. He has over 30 publications in books, journals, and international conference proceedings in the above areas.

Stefanos Kollias was born in Athens in 1956. He obtained his Diploma from NTUA in 1979, his M.Sc. in Communication Engineering in 1980 from UMIST in England, and his Ph.D. in Signal Processing from the Computer Science Division of NTUA. He has been with the Electrical Engineering Department of NTUA since 1986 where he serves now as a Professor. Since 1990 he has been Director of the Image, Video and Multimedia Systems Laboratory of NTUA. He has published more than 120 papers in the above fields, 50 of which were in international journals. He has been a member of the Technical or Advisory Committee or invited speaker in 40 international conferences. He is a reviewer of 10 IEEE Transactions and of 10 other journals. Ten graduate students have completed their doctorate under his supervision, while another 10 are currently performing their Ph.D. thesis. He and his team have been participating in 38 European and National projects.

## References


Ardissono, L., & Goy, A. (2000). Tailoring the interaction with users in web stores. *User Modeling and User-Adapted Interaction, 10*, 251–303.

Balabanovic, M., & Shoham, Y. (1997). Content-based, collaborative recommendation. *Communications of the ACM, 4*, 66–72.

Baudisch, P., & Terveen, L. (Eds.). (1999). *Proceedings from the CHI '99 Workshop Interacting with Recommender Systems*, Pittsburgh, PA [Online].

Benaki, E., Karkaletsis, V., & Spyropoulos, C. D. (1997). Integrating user modelling into information extraction: The umie prototype. In *Proceedings Sixth International*

*Conference on User Modeling (UM'97)* (pp. 55–57).

Brajnik, G., & Tasso, C. (1994). A shell for developing non-monotonic user modelling systems. *International Journal of Human–Computer Studies, 40*, 31–62.

Burke, R. (2002). Hybrid recommender systems: Survey and experiments. *User Modeling and User Adapted Interaction, 12*(4), 331–370.

Chen, P. M., & Kuo, F. C. (2000). An information retrieval system based on a user profile. *Journal of Systems and Software, 54*, 3–8.

Chin, D. N. (1989). KNOME: Modelling what the user knows. In A. Kobsa & W. Wahlster (Eds.), *User models in dialog systems* (pp. 74–107). Berlin: Springer.

Delgado, J., & Davidson, R. (2002). Knowledge bases and user profiling
in travel and hospitality recommender systems. In *Proceedings of the*
*ENTER 2002 Conference* (pp. 1–16). Innsbruck, Austria: Springer Verlag.

Doulamis, N., Doulamis, A., & Kollias, S. (2000). On-line retrainable neural networks: Improving performance of neural networks in image analysis. *IEEE Transactions on Neural Networks, 11*(1), 1–20.

Franke, T. (2003). Enhancing an online regional tourism consulting system with extended personalized services. *Information Technology & Tourism, 5*(3), 135–150.

Kautz, H. (Ed.). (1998). *Recommender systems.* Papers from the 1998 Workshop, Technical Report WS-98-08. Menlo Park, CA: AAAI Press.

Kay, J. (1995). The um toolkit for cooperative user modelling. *User Modelling and User Adapted Interaction, 4*, 149–196.

Langley, P. (1999). User modelling in adaptive interfaces. *Proceedings Seventh International Conference on User Modelling (UM'99)* (pp. 357–370), Banff, Canada.

Maglogiannis, I., Kormentzas, G., & Panagiotarakis, N. (2002). The business model of a distributed Virtual Travel Agency. *15th EEEE Conference in Operational Research*, Tripoli Greece.

Maglogiannis, I., Kormentzas, G., & Panagiotarakis, N. (2003). Using information and communication technologies in maritime travel business. *Proccedings of the 10th International Conference on Information Technology and Travel & Tourism*, Helsinki, Finland.

O'Connor, P., & Frew, A. (2000). Evaluating electronic channels of distribution in the hotel sector: A Delphi study. *Information Technology & Tourism, 3*(3/4), 177–193.

Paliouras, G., Papatheodorou, C., Karkaletsis, V., & Spyropoulos, C. D. (2002). Discovering user communities on the Internet using unsupervised machine learning techniques. *Interacting With Computers, 14*, 761–791.

Resnick, P., & Varian, H. (Eds.). (1997). Special issue on recommender systems. *Communications of the ACM, 40*(3), 56–89.

Ricci, F., Arslan, B., Mirzadeh, N., & Venturini, A. (2002). ITR: A case-based travel advisory system. In S. Craw & A. Preece (Eds.), *6th European Conference on Case Based Reasoning (ECCBR)*, Aberdeen, Scotland.

Shafer, J. B., Konstan, J. A., & Riedl, J. (2001). E-commerce recommendation applications. *Data Mining and Knowledge Discovery, 5*, 115–153.

Soboroff, I., Nicholas, C., & Pazzani, M. J. (Eds.). (1999). *Proceedings from the ACM SIG-IR'99 Workshop on Recommender Systems, Algorithms and Evaluation*, University of California, Berkeley [Online].

Theodoridis, S., & Koutroumbas, K. (1998). *Pattern recognition.* Academic Press.

Tsapatsoulis, N., Wallace, M., & Kasderidis, S. (2003). Improving the performance of resource allocation networks through hierarchical clustering of high-dimensional data. *Proceedings of the International Conference on Artificial Neural Networks (ICANN)*, Istanbul, Turkey.

Yager, R. R. (2000). Intelligent control of the hierarchical agglomerative clustering process. *IEEE Transactions on Systems, Man and Cybernetics, Part B, 30*(6), 835–845.

Varian, H. (Ed.). (1996). Collaborative filtering. *Proceedings from the Berkeley Workshop on Collaborative Filtering*, University of California, Berkeley.

Wallace, M., & Kollias, S. (2003). Soft attribute selection for hierarchical clustering in high dimensions. *Proceedings of the International Fuzzy Systems Association World Congress (IFSA)*, Istanbul, Turkey.