

Design & Development of a TAST Model to Generate Personalized Travel Package Recommendation

C.Srujana¹; S.Anitha²; M.Laxman³ & Prof.Dr.G.Manoj Someswar⁴

1. M.Tech.(CSE) from Kshatriya College of Engineering, Affiliated to JNTUH, Telangana, India.
2. Assistant Professor, Dept. of CSE, Kshatriya College of Engineering, Affiliated to JNTUH, Telangana, India.
3. Associate Professor & HOD, Dept. of CSE, Kshatriya College of Engineering, Affiliated to JNTUH, Telangana, India.
4. Principal & Professor, Dept. of CSE, Anwar-ul-uloom College of Engineering & Technology, Vikarabad, RR District, Affiliated to JNTUH, Telangana, India.

ABSTRACT:

Recent years have witnessed an increased interest in recommender systems. Despite significant progress in this field, there still remain numerous avenues to explore. Indeed, this research paper provides a study of exploiting online travel information for personalized travel package recommendation. A critical challenge along this line is to address the unique characteristics of travel data, which distinguish travel packages from traditional items for recommendation. To that end, in this research paper, we first analyze the characteristics of the existing travel packages and develop a tourist-area-season topic (TAST) model. This TAST model can represent travel packages and tourists by different topic distributions where the topic extraction is conditioned on both the tourists and the intrinsic features (i.e., locations, travel seasons) of the landscapes. Then, based on this topic model representation, we propose a cocktail approach to generate the lists for personalized travel package recommendation. Furthermore, we extend the TAST model to the tourist-relation-area-season topic (TRAST) model for capturing the latent relationships among the tourists in each travel group. Finally, we evaluate the TAST model, the TRAST model, and the cocktail recommendation approach on the real-world travel package data. Experimental results show that the TAST model can effectively capture the unique characteristics of the travel data and the cocktail approach is, thus, much more effective than traditional recommendation techniques for travel package recommendation. Also, by considering tourist relationships, the TRAST model can be used as an effective assessment for travel group formation.

Keywords: Tourist relation area season topic; Chi Square Automatic Interaction Detection; Classification and Regression Trees; Rule induction; Intrinsic Complex Spatio-temporal Relationships; Data visualization

INTRODUCTION

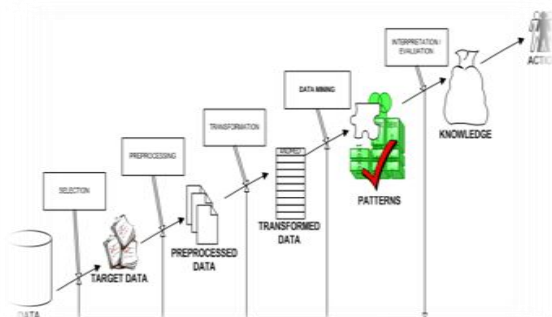


Figure 1: Structure of Data Mining

Generally, data mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.[1]

While large-scale information technology has been evolving separate transaction and analytical systems, data mining

provides the link between the two. Data mining software analyzes relationships and patterns in stored transaction data based on open-ended user queries. Several types of analytical software are available: statistical, machine learning, and neural networks. **Generally, any of four types of relationships are sought:**

- **Classes:** Stored data is used to locate data in predetermined groups. For example, a restaurant chain could mine customer purchase data to determine when customers visit and what they typically order. This information could be used to increase traffic by having daily specials.[2]
- **Clusters:** Data items are grouped according to logical relationships or consumer preferences. For example, data can be mined to identify market segments or consumer affinities.
- **Associations:** Data can be mined to identify associations. The beer-diaper example is an example of associative mining.[3]
- **Sequential patterns:** Data is mined to anticipate behavior patterns and trends. For example, an outdoor equipment retailer could predict the likelihood of a backpack being purchased based on a consumer's purchase of sleeping bags and hiking shoes.

Data mining consists of five major elements:

- 1) Extract, transform, and load transaction data onto the data warehouse system.
- 2) Store and manage the data in a multidimensional database system.
- 3) Provide data access to business analysts and information technology professionals.
- 4) Analyze the data by application software.
- 5) Present the data in a useful format, such as a graph or table.[4]

Different levels of analysis are available:

- **Artificial neural networks:** Non-linear predictive models that learn through training and resemble biological neural networks in structure.
 - **Genetic algorithms:** Optimization techniques that use process such as genetic combination, mutation, and natural selection in a design based on the concepts of natural evolution.
 - **Decision trees:** Tree-shaped structures that represent sets of decisions. These decisions generate rules for the classification of a dataset. Specific decision tree methods include Classification and Regression Trees (CART) and Chi Square Automatic Interaction Detection (CHAID). CART and CHAID are decision tree techniques used for classification of a dataset. They provide a set of rules that you can apply to a new (unclassified) dataset to predict which records will have a given outcome. CART segments a dataset by creating 2-way splits while CHAID segments using chi square tests to create multi-way splits. CART typically requires less data preparation than CHAID.[5]
 - **Nearest neighbor method:** A technique that classifies each record in a dataset based on a combination of the classes of the k record(s) most similar to it in a historical dataset (where $k=1$). Sometimes called the k -nearest neighbor technique.
 - **Rule induction:** The extraction of useful if-then rules from data based on statistical significance.
 - **Data visualization:** The visual interpretation of complex relationships in multidimensional data. Graphics tools are used to illustrate data relationships.
- Characteristics of Data Mining:**
- **Large quantities of data:** The volume of data so great it has to be analyzed by automated techniques e.g. satellite information, credit card transactions etc.
 - **Noisy, incomplete data:** Imprecise data is the characteristic of all data collection.



- **Complex data structure:** conventional statistical analysis not possible
- **Heterogeneous data stored in legacy systems**

Benefits of Data Mining:

- 1) It's one of the most effective services that are available today. With the help of data mining, one can discover precious information about the customers and their behavior for a specific set of products and evaluate and analyze, store, mine and load data related to them.[6]
- 2) An analytical CRM model and strategic business related decisions can be made with the help of data mining as it helps in providing a complete synopsis of customers.
- 3) An endless number of organizations have installed data mining projects and it has helped them see their own companies make an unprecedented improvement in their marketing strategies (Campaigns).
- 4) Data mining is generally used by organizations with a solid customer focus. For its flexible nature as far as applicability is concerned is being used vehemently in applications to foresee crucial data including industry analysis and consumer buying behaviors.
- 5) Fast paced and prompt access to data along with economic processing techniques have made data mining one of the most suitable services that a company seek.

Advantages of Data Mining:

Marketing / Retail:

Data mining helps marketing companies build models based on historical data to predict who will respond to the new marketing campaigns such as direct mail, online marketing campaign...etc. Through the results, marketers will have appropriate approach to sell profitable products to targeted customers.

Data mining brings a lot of benefits to retail companies in the same way as marketing. Through market basket analysis, a store can have an appropriate production arrangement in a way that customers can buy frequent buying products together with pleasant. In addition, it also helps the retail companies offer certain discounts for particular products that will attract more customers.[7]

1. Finance / Banking

Data mining gives financial institutions information about loan information and credit reporting. By building a model from historical customer's data, the bank and financial institution can determine good and bad loans. In addition, data mining helps banks detect fraudulent credit card transactions to protect credit card's owner.

2. Manufacturing

By applying data mining in operational engineering data, manufacturers can detect faulty equipments and determine optimal control parameters. For example semi-conductor manufacturers has a challenge that even the conditions of manufacturing environments at different wafer production plants are similar, the quality of wafer are lot the same and some for unknown reasons even has defects. Data mining has been applying to determine the ranges of control parameters that lead to the production of golden wafer. Then those optimal control parameters are used to manufacture wafers with desired quality.[8]

3. Governments

Data mining helps government agency by digging and analyzing records of financial transaction to build patterns that can detect money laundering or criminal activities.

4. Law enforcement:

Data mining can aid law enforcers in identifying criminal suspects as well as apprehending these criminals by examining trends in location, crime type, habit, and other patterns of behaviors.[9]

5. Researchers:

Data mining can assist researchers by speeding up their data analyzing process; thus, allowing those more time to work on other projects.



LITERATURE SURVEY

Future computing environments will free the user from the constraints of the desktop. Applications for a mobile environment should take advantage of contextual information, such as position, to offer greater services to the user. In his paper, we present the Cyber guide project, in which we are building prototypes of a mobile context-aware tour guide. Knowledge of the user's current location, as well as a history of past locations, are used to provide more of the kind of services that we come to expect from a real tour guide. We describe the architecture and features of a variety of Cyber guide prototypes developed for indoor and outdoor use on a number of different hand-held platforms. We also discuss the general research issues that have emerged in our context-aware applications development in a mobile environment.

We propose fLDA, a novel matrix factorization method to predict ratings in recommender system applications where a "bag-of-words" representation for item meta-data is natural. Such scenarios are commonplace in web applications like content recommendation, ad targeting and web search where items are articles, ads and web pages respectively. Because of data sparseness, regularization is key to good predictive accuracy. Our method works by regularizing both user and item factors simultaneously through user features and the bag of words associated with each item. Specifically, each word in an item is associated with a discrete latent factor often referred to as the topic of the word; item topics are obtained by averaging topics across all words in an item. Then, user rating on an item is modeled as user's affinity to the item's topics where user affinity to topics (user factors) and topic assignments to words in items (item factors) are learned jointly in a supervised fashion. To avoid overfitting, user and item factors are regularized through Gaussian linear regression and Latent Dirichlet Allocation (LDA) priors respectively. We show our model is accurate, interpretable and handles both cold-start and warm-start scenarios seamlessly through a single model. The efficacy of our method is illustrated on benchmark datasets and a new dataset from Yahoo! Buzz where fLDA provides superior predictive accuracy in cold-start scenarios and is comparable to state-of-the-art methods in warm-start scenarios. As a by-product, fLDA also identifies interesting topics that explains user-item interactions. Our method also generalizes a recently proposed technique called supervised LDA (sLDA) to collaborative filtering applications. While

sLDA estimates item topic vectors in a supervised fashion for a single regression, fLDA incorporates multiple regressions (one for each user) in estimating the item factors.[10]

Recommender systems are information search and decision support tools used when there is an overwhelming set of options to consider or when the user lacks the domain-specific knowledge necessary to take autonomous decisions. They provide users with personalized recommendations adapted to their needs and preferences in a particular usage context. In this research paper, we present an approach for integrating recommendation and electronic map technologies to build a map-based conversational mobile recommender system that can effectively and intuitively support users in finding their desired products and services. The results of our real-user study show that integrating map-based visualization and interaction in mobile recommender systems improves the system recommendation effectiveness and increases the user satisfaction.

When visiting cities as tourists, most of the times people do not make very detailed plans and, when choosing where to go and what to see they tend to select the area with the major number of interesting facilities. Therefore, it would be useful to support the user choice with contextual information presentation, information clustering and comparative explanations of places of potential interest in a given area. In this paper we illustrate how MyMap, a mobile recommender system in the Tourism domain, generates comparative descriptions to support users in making decisions about what to see, among relevant objects of interest.[11]

The increasing availability of large-scale location traces creates unprecedented opportunities to change the paradigm for knowledge discovery in transportation systems. A particularly promising area is to extract energy-efficient transportation patterns (green knowledge), which can be used as guidance for reducing inefficiencies in energy consumption of transportation sectors. However, extracting green knowledge from location traces is not a trivial task. Conventional data analysis tools are usually not customized for handling the massive quantity, complex, dynamic, and distributed nature of location traces.[12] To that end, in this paper, we provide a focused study of extracting energy-efficient transportation patterns from location traces. Specifically, we have the initial focus on a sequence of

mobile recommendations. As a case study, we develop a mobile recommender system which has the ability in recommending a sequence of pick-up points for taxi drivers or a sequence of potential parking positions. The goal of this mobile recommendation system is to maximize the probability of business success. Along this line, we provide a Potential Travel Distance (PTD) function for evaluating each candidate sequence. This PTD function possesses a monotone property which can be used to effectively prune the search space. Based on this PTD function, we develop two algorithms, LCP and Sky Route, for finding the recommended routes. Finally, experimental results show that the proposed system can provide effective mobile sequential recommendation and the knowledge extracted from location traces can be used for coaching drivers and leading to the efficient use of energy.

SYSTEM STUDY

FEASIBILITY STUDY

The feasibility of the project is analyzed in this phase and business proposal is put forth with a very general plan for the project and some cost estimates. During system analysis the feasibility study of the proposed system is to be carried out. This is to ensure that the proposed system is not a burden to the company. For feasibility analysis, some understanding of the major requirements for the system is essential.

Three key considerations involved in the feasibility analysis are

- ◆ ECONOMICAL FEASIBILITY
- ◆ TECHNICAL FEASIBILITY
- ◆ SOCIAL FEASIBILITY

ECONOMICAL FEASIBILITY

This study is carried out to check the economic impact that the system will have on the organization. The amount of fund that the company can pour into the research and development of the system is limited. The expenditures must be justified.[13] Thus the developed system as well within the budget and this was achieved because most of the

technologies used are freely available. Only the customized products had to be purchased.

TECHNICAL FEASIBILITY

This study is carried out to check the technical feasibility, that is, the technical requirements of the system. Any system developed must not have a high demand on the available technical resources. This will lead to high demands on the available technical resources. This will lead to high demands being placed on the client. The developed system must have a modest requirement, as only minimal or null changes are required for implementing this system.

SOCIAL FEASIBILITY

The aspect of study is to check the level of acceptance of the system by the user. This includes the process of training the user to use the system efficiently.[14] The user must not feel threatened by the system, instead must accept it as a necessity. The level of acceptance by the users solely depends on the methods that are employed to educate the user about the system and to make him familiar with it. His level of confidence must be raised so that he is also able to make some constructive criticism, which is welcomed, as he is the final user of the system.

SYSTEM DESIGN

SYSTEM ARCHITECTURE:

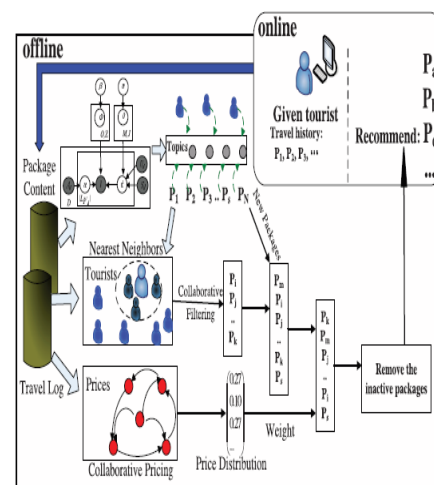


Figure 2: System Architecture

DATA FLOW DIAGRAM:

1. The DFD is also called as bubble chart. It is a simple graphical formalism that can be used to represent a system in terms of input data to the system, various processing carried out on this data, and the output data is generated by this system.
2. The data flow diagram (DFD) is one of the most important modeling tools. It is used to model the system components. These components are the system process, the data used by the process, an external entity that interacts with the system and the information flows in the system.
3. DFD shows how the information moves through the system and how it is modified by a series of transformations. It is a graphical technique that depicts information flow and the transformations that are applied as data moves from input to output.
4. DFD is also known as bubble chart. A DFD may be used to represent a system at any level of abstraction. DFD may be partitioned into levels that represent increasing information flow and functional detail.

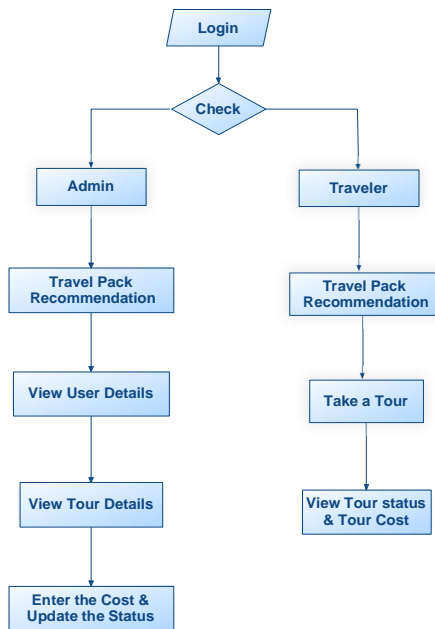


Figure 3: Data Flow Diagram

UML DIAGRAMS

UML stands for Unified Modeling Language. UML is a standardized general-purpose modeling language in the field of object-oriented software engineering. The standard is managed, and was created by, the Object Management Group.

The goal is for UML to become a common language for creating models of object oriented computer software. In its current form UML is comprised of two major components: a Meta-model and a notation. In the future, some form of method or process may also be added to; or associated with, UML.

The Unified Modeling Language is a standard language for specifying, Visualization, Constructing and documenting the artifacts of software system, as well as for business modeling and other non-software systems.

The UML represents a collection of best engineering practices that have proven successful in the modeling of large and complex systems.

The UML is a very important part of developing objects oriented software and the software development process. The UML uses mostly graphical notations to express the design of software projects.

GOALS:

The Primary goals in the design of the UML are as follows:

1. Provide users a ready-to-use, expressive visual modeling Language so that they can develop and exchange meaningful models.
2. Provide extendibility and specialization mechanisms to extend the core concepts.
3. Be independent of particular programming languages and development process.
4. Provide a formal basis for understanding the modeling language.
5. Encourage the growth of OO tools market.
6. Support higher level development concepts such as collaborations, frameworks, patterns and components.
7. Integrate best practices.

USE CASE DIAGRAM:

A use case diagram in the Unified Modeling Language (UML) is a type of behavioral diagram defined by and created from a Use-case analysis. Its purpose is to present a graphical overview of the functionality provided

by a system in terms of actors, their goals (represented as use cases), and any dependencies between those use cases. The main purpose of a use case diagram is to show what system functions are performed for which actor. Roles of the actors in the system can be depicted.

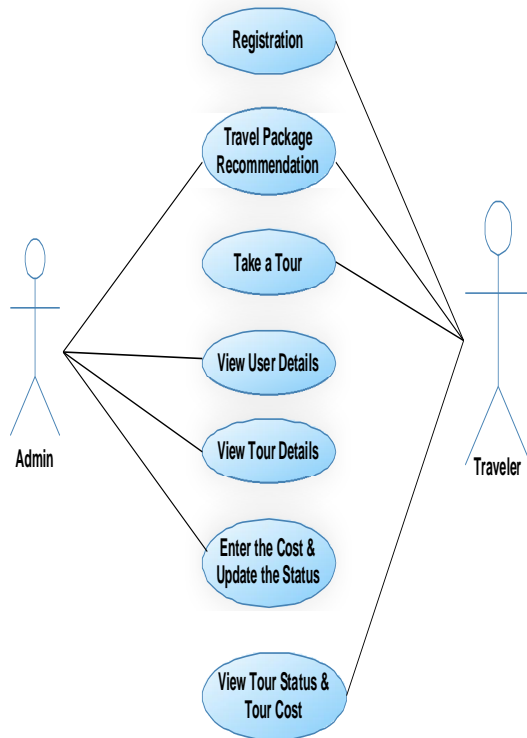


Figure 4: Use Case Diagram

CLASS DIAGRAM:

In software engineering, a class diagram in the Unified Modeling Language (UML) is a type of static structure diagram that describes the structure of a system by showing the system's classes, their attributes, operations (or methods), and the relationships among the classes. It explains which class contains information.

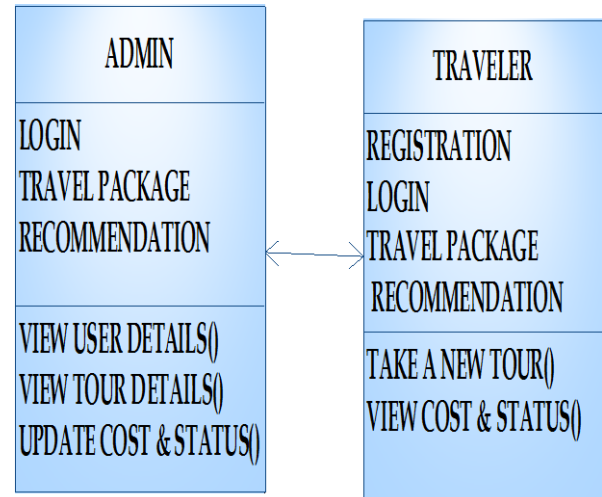


Figure 5: Class Diagram

SEQUENCE DIAGRAM:

A sequence diagram in Unified Modeling Language (UML) is a kind of interaction diagram that shows how processes operate with one another and in what order. It is a construct of a Message Sequence Chart. Sequence diagrams are sometimes called event diagrams, event scenarios, and timing diagrams.

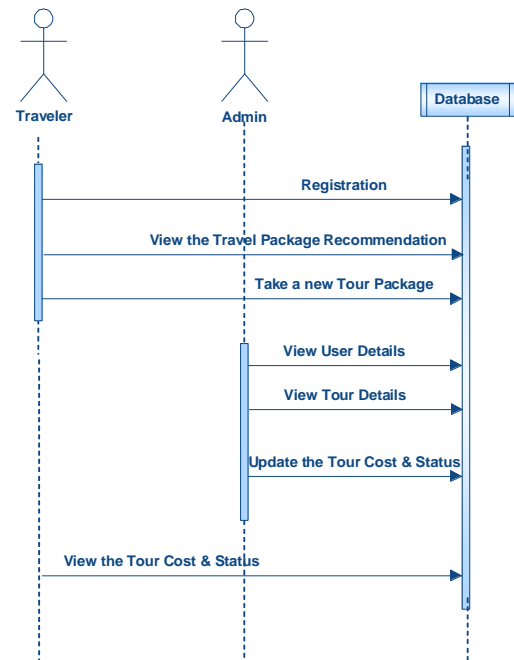


Figure 6: Sequence Diagram

ACTIVITY DIAGRAM:

Activity diagrams are graphical representations of workflows of stepwise activities and actions with support for choice, iteration and concurrency. In the Unified Modeling Language, activity diagrams can be used to describe the business and operational step-by-step workflows of components in a system. An activity diagram shows the overall flow of control.

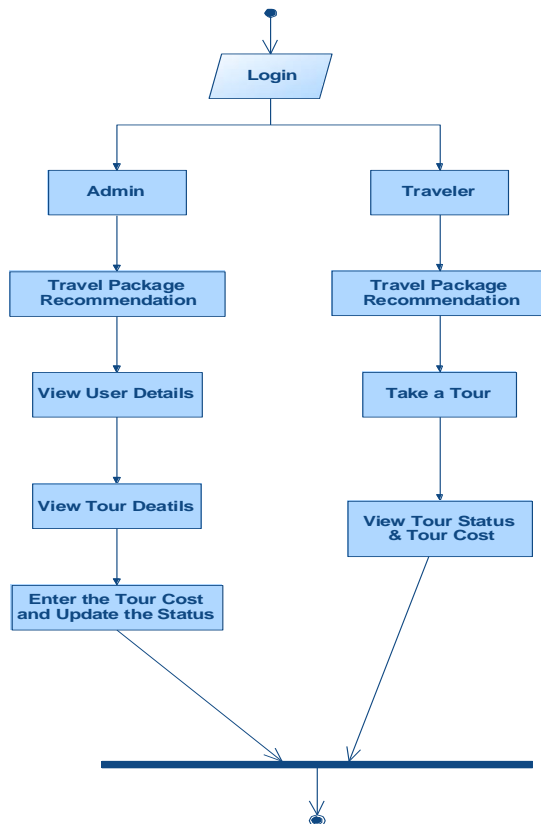


Figure 7: Activity Diagram

INPUT DESIGN

The input design is the link between the information system and the user. It comprises the developing specification and procedures for data preparation and those steps are necessary to put transaction data in to a usable form for processing can be achieved by inspecting the computer to read data from a written or printed document or it can occur by having people keying the data directly into the system. The design of

input focuses on controlling the amount of input required, controlling the errors, avoiding delay, avoiding extra steps and keeping the process simple. The input is designed in such a way so that it provides security and ease of use with retaining the privacy. Input Design considered the following things:

- What data should be given as input?
- How the data should be arranged or coded?
- The dialog to guide the operating personnel in providing input.
- Methods for preparing input validations and steps to follow when error occur.

OBJECTIVES

1. Input Design is the process of converting a user-oriented description of the input into a computer-based system. This design is important to avoid errors in the data input process and show the correct direction to the management for getting correct information from the computerized system.

2. It is achieved by creating user-friendly screens for the data entry to handle large volume of data. The goal of designing input is to make data entry easier and to be free from errors. The data entry screen is designed in such a way that all the data manipulates can be performed. It also provides record viewing facilities.

3. When the data is entered it will check for its validity. Data can be entered with the help of screens. Appropriate messages are provided as when needed so that the user will not be in maize of instant. Thus the objective of input design is to create an input layout that is easy to follow

OUTPUT DESIGN

A quality output is one, which meets the requirements of the end user and presents the information clearly. In any system results of processing are communicated to



the users and to other system through outputs. In output design it is determined how the information is to be displaced for immediate need and also the hard copy output. It is the most important and direct source information to the user. Efficient and intelligent output design improves the system's relationship to help user decision-making.

1. Designing computer output should proceed in an organized, well thought out manner; the right output must be developed while ensuring that each output element is designed so that people will find the system can use easily and effectively. When analysis design computer output, they should Identify the specific output that is needed to meet the requirements.

2. Select methods for presenting information.

3. Create document, report, or other formats that contain information produced by the system.

The output form of an information system should accomplish one or more of the following objectives.

- ❖ Convey information about past activities, current status or projections of the
- ❖ Future.
- ❖ Signal important events, opportunities, problems, or warnings.
- ❖ Trigger an action.
- ❖ Confirm an action.

SYSTEM ANALYSIS

EXISTING SYSTEM:

There are many technical and domain challenges inherent in designing and implementing an effective recommender system for personalized travel package recommendation.

1. Travel data are much fewer and sparser than traditional items, such as movies for recommendation, because the costs for a travel are much more expensive than for watching a movie.
2. Every travel package consists of many landscapes (places of interest and attractions), and, thus, has intrinsic complex spatio-temporal relationships. For example, a travel package only includes the landscapes which are geographically colocated together. Also, different travel packages are usually developed for different travel seasons. Therefore, the landscapes in a travel package usually have spatial temporal autocorrelations.
3. Traditional recommender systems usually rely on user explicit ratings. However, for travel data, the user ratings are usually not conveniently available.

DISADVANTAGES OF EXISTING SYSTEM:

- Recommendation has a long period of stable value.
- To replace the old ones based on the interests of the tourists.
- A values of travel packages can easily depreciate over time and a package usually only lasts for a certain period of time

PROPOSED SYSTEM:

In this paper, we aim to make personalized travel package recommendations for the tourists. Thus, the users are the tourists and the items are the existing packages, and we exploit a real-world travel data set provided by a travels for building recommender systems. we develop a tourist-area-season topic (TAST) model, which can represent travel packages and tourists by different topic distributions. In the



TAST model, the extraction of topics is conditioned on both the tourists and the intrinsic features (i.e., locations, travel seasons) of the landscapes. Based on this TAST model, a cocktail approach is developed for personalized travel package recommendation by considering some additional factors including the seasonal behaviors of tourists, the prices of travel packages, and the cold start problem of new packages.

ADVANTAGES OF PROPOSED SYSTEM:

- Represent the content of the travel packages and the interests of the tourists.
- TAST model can effectively capture the unique characteristics of travel data.
- The cocktail recommendation approach performs much better than traditional techniques.

SYSTEM TESTING

The purpose of testing is to discover errors. Testing is the process of trying to discover every conceivable fault or weakness in a work product. It provides a way to check the functionality of components, sub assemblies, assemblies and/or a finished product. It is the process of exercising software with the intent of ensuring that the Software system meets its requirements and user expectations and does not fail in an unacceptable manner. There are various types of test. Each test type addresses a specific testing requirement.

TYPES OF TESTS

Unit testing

Unit testing involves the design of test cases that validate that the internal program logic is functioning properly, and that program inputs produce valid outputs. All decision branches and internal code flow should be validated. It is the testing of individual software units of the application. It is done after the

completion of an individual unit before integration. This is a structural testing, that relies on knowledge of its construction and is invasive. Unit tests perform basic tests at component level and test a specific business process, application, and/or system configuration. Unit tests ensure that each unique path of a business process performs accurately to the documented specifications and contains clearly defined inputs and expected results.

Integration testing

Integration tests are designed to test integrated software components to determine if they actually run as one program. Testing is event driven and is more concerned with the basic outcome of screens or fields. Integration tests demonstrate that although the components were individually satisfactory, as shown by successfully unit testing, the combination of components is correct and consistent. Integration testing is specifically aimed at exposing the problems that arise from the combination of components.

Functional test

Functional tests provide systematic demonstrations that functions tested are available as specified by the business and technical requirements, system documentation, and user manuals.

Functional testing is centered on the following items:

Valid Input : identified classes of valid input must be accepted.

Invalid Input : identified classes of invalid input must be rejected.

Functions : identified functions must be exercised.

Output : identified classes of application outputs must be exercised.



Systems/Procedures: interfacing systems or procedures must be invoked.

Organization and preparation of functional tests is focused on requirements, key functions, or special test cases. In addition, systematic coverage pertaining to identify Business process flows; data fields, predefined processes, and successive processes must be considered for testing. Before functional testing is complete, additional tests are identified and the effective value of current tests is determined.

System Test

System testing ensures that the entire integrated software system meets requirements. It tests a configuration to ensure known and predictable results. An example of system testing is the configuration oriented system integration test. System testing is based on process descriptions and flows, emphasizing pre-driven process links and integration points.

White Box Testing

White Box Testing is a testing in which in which the software tester has knowledge of the inner workings, structure and language of the software, or at least its purpose. It is used to test areas that cannot be reached from a black box level.

Black Box Testing

Black Box Testing is testing the software without any knowledge of the inner workings, structure or language of the module being tested. Black box tests, as most other kinds of tests, must be written from a definitive source document, such as specification or requirements document, such as specification or requirements document. It is a testing in which the software under test is treated, as a black box .you cannot “see” into it. The test provides inputs and responds to outputs without considering how the software works.

Unit Testing:

Unit testing is usually conducted as part of a combined code and unit test phase of the software lifecycle, although it is not uncommon for coding and unit testing to be conducted as two distinct phases.

Test strategy and approach

Field testing will be performed manually and functional tests will be written in detail.

Test objectives

- All field entries must work properly.
- Pages must be activated from the identified link.
- The entry screen, messages and responses must not be delayed.

Features to be tested

- Verify that the entries are of the correct format
- No duplicate entries should be allowed
- All links should take the user to the correct page.

Integration Testing

Software integration testing is the incremental integration testing of two or more integrated software components on a single platform to produce failures caused by interface defects.

The task of the integration test is to check that components or software applications, e.g. components in a software system or – one step up – software applications at the company level – interact without error.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.



Acceptance Testing

User Acceptance Testing is a critical phase of any project and requires significant participation by the end user. It also ensures that the system meets the functional requirements.

Test Results: All the test cases mentioned above passed successfully. No defects encountered.

IMPLEMENTATION

MODULES:

1. User Module.
2. Server Module.
3. Package recommendations.
4. TAST Model

MODULES DESCRIPTION:

User Module

In this module, Users are having authentication and security to access the result from the system. Before accessing or searching the details user should have the account in that otherwise they should register first.

Server Module

In this module, provide the detailed information about the unique characteristics of travel package data. We aim to make personalized travel package recommendations for the tourists. Thus, the users are the tourists and the items are the existing packages, and we exploit a real-world travel data set provided by a travel company in China for building recommender systems.

Package recommendations

We collect some unique characteristics of the travel data. First, it is very sparse, and each tourist has only a few travel records. The extreme sparseness of the data

leads to difficulties for using traditional recommendation techniques, such as collaborative filtering. For example, it is hard to find the credible nearest neighbours for the tourists because there are very few co-travelling packages.

TAST Model:

First, it is necessary to determine the set of target

tourists, the travel seasons, and the travel places. Second, one or multiple travel topics (e.g., “The Sunshine Trip”) will be chosen based on the category of target tourists and the scheduled travel seasons. Each package and landscape can be viewed as a mixture of a number of travel topics. Then, the landscapes will be determined according to the travel topics and the geographic locations. Finally, some additional information (e.g., price, transportation, and accommodations) should be included. According to these processes, we formalize package generation as a What-Who-When-Where (4W) problem.

RESULTS & CONCLUSION

In this research paper, we present study on personalized travel package recommendation. Specifically, we first analyzed the unique characteristics of travel packages and developed the TAST model, a Bayesian network for travel package and tourist representation. The TAST model can discover the interests of the tourists and extract the spatial-temporal correlations among landscapes. Then, we exploited the TAST model for developing a cocktail approach on personalized travel package recommendation. This cocktail approach follows a hybrid recommendation strategy and has the ability to combine several constraints existing in the real-world scenario. Furthermore, we extended the TAST model to the TRAST model, which can capture the relationships among tourists in each travel group. Finally, an empirical study was conducted on real-world travel data. Experimental results demonstrate that the TAST model can capture the unique characteristics of the travel packages, the cocktail



approach can lead to better performances of travel package recommendation, and the TRAST model can be used as an effective assessment for travel group automatic formation. We hope these encouraging results could lead to many future work.

REFERENCES

- [1] G.D. Abowd et al., "Cyber-Guide: A Mobile Context-Aware Tour Guide," *Wireless Networks*, vol. 3, no. 5, pp. 421-433, 1997.
- [2] G. Adomavicius and A. Tuzhilin, "Toward the Next Generation of Recommender Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Trans. Knowledge and Data Eng.*, vol. 17, no. 6, pp. 734-749, June 2005.
- [3] D. Agarwal and B. Chen, "fLDA: Matrix Factorization through Latent Dirichlet Allocation," *Proc. Third ACM Int'l Conf. Web Search and Data Mining (WSDM '10)*, pp. 91-100, 2010.
- [4] O. Averjanova, F. Ricci, and Q.N. Nguyen, "Map-Based Interaction with a Conversational Mobile Recommender System," *Proc. Second Int'l Conf. Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM '08)*, pp. 212-218, 2008.
- [5] D.M. Blei, Y.N. Andrew, and I.J. Michael, "Latent Dirichlet Allocation," *J. Machine Learning Research*, vol. 3, pp. 993-1022, 2003.
- [6] R. Burke, "Hybrid Web Recommender Systems," *The Adaptive Web*, vol. 4321, pp. 377-408, 2007.
- [7] B.D. Carolis, N. Novielli, V.L. Plantamura, and E. Gentile, "Generating Comparative Descriptions of Places of Interest in the Tourism Domain," *Proc. Third ACM Conf. Recommender Systems (RecSys '09)*, pp. 277-280, 2009.
- [8] F. Cena et al., "Integrating Heterogeneous Adaptation Techniques to Build a Flexible and Usable Mobile Tourist Guide," *AI Comm.*, vol. 19, no. 4, pp. 369-384, 2006.
- [9] W. Chen, J.C. Chu, J. Luan, H. Bai, Y. Wang, and E.Y. Chang, "Collaborative Filtering for Orkut Communities: Discovery of User Latent Behavior," *Proc. ACM 18th Int'l Conf. World Wide Web (WWW '09)*, pp. 681-690, 2009.
- [10] N.A.C. Cressie, *Statistics for Spatial Data*. Wiley and Sons, 1991.
- [11] J. Delgado and R. Davidson, "Knowledge Bases and User Profiling in Travel and Hospitality Recommender Systems," *Proc. ENTER 2002 Conf. (ENTER '02)*, pp. 1-16, 2002.
- [12] U.M. Fayyad and K.B. Irani, "Multi-Interval Discretization of Continuous-Valued Attributes for Classification Learning," *Proc. Int'l Joint Conf. Artificial Intelligence (IJCAI)*, pp. 1022-1027, 1993.
- [13] F. Fouss et al., "Random-Walk Computation of Similarities between Nodes of a Graph with Application to Collaborative Recommendation," *IEEE Trans. Knowledge and Data Eng.*, vol. 19, no. 3, pp. 355-369, Mar. 2007.
- [14] Y. Ge et al., "Cost-Aware Travel Tour Recommendation," *Proc. 17th ACM SIGKDD Int'l Conf. Knowledge Discovery and Data Mining (SIGKDD '11)*, pp. 983-991, 2011.