Tour Miner:Mining System of Tour Plans from SNS
— Mining the User Experience into Travel Records —

Shingo Yamaguchi, Katakura Shouta
Graduate School of Sciences and Technology for Innovation
Yamaguchi University
2-16-1 Tokiwadai, Ube 755-8611, Japan
Email: { fshingo, s055ffg } @yamaguchi-u.ac.jp

Thanawit Gerdprasert, Anawat Kaenthong,
Merisa Puengsawad
Department of Computer Engineering, Faculty of Engineering
Kasetsart University
Bangkok 10900, Thailand
Email : { thanawit.g, anawat.k, merisa.p }@ku.th

Abstract—**Tour Miner is the process that consist of mining the tour plans from popular social network system when the given keyword is presented, the system then find the travel record from the according input and then return the Tour Record back to the user. While the previous paper, the research focus primarily on the Smelting process that use the heuristic mining in order to output the best possible tour record, in this paper we primarily focused on the mining function. We create the offline database which stored the user travel record in order to access it faster, the user input will be used to generate the travel record which will be used in the Smeltering function in the further part.**

I. Introduction

With the popularity of the internet becoming more and more popular, the social network become relevant to our daily day life. People start using more and more social network system in order to share their own experience. This is where we proposed the Framework called Tour Miner[2] (See Fig. 1) which try to create the tour record based on the using the user experience from Social Network System.

Arai et al. [1] proposed the method of getting the travel record from social network system called Twitter and its support platform called Four Square in order to gather the user experience from the twitter using various keywords and notable term to improve the quality of the travel records by distributing the evaluation point into the user experience. The model from this research using the data from the twitter as the quality based, which result in the data that is very detailed and very valuable with the tradeoff to the lack of data amount.

For our research purpose, we want to emphasize on getting the numerous of travel record to improve the effective of the smeltering phase of the research. Alternatively, rather than gathering the record at the real time, we try to create the offline database in order to make our service process faster and suitable for recommending the tour plan for the user. In the final product, we plan to illustrate how the process work by develop the full-stack web application.

II. Preminary

1) Tour Miner: Tour Miner is the process which has the goal to mines the tour plans from SNS. The framework consist of two main process, web mining and process mining. Both of
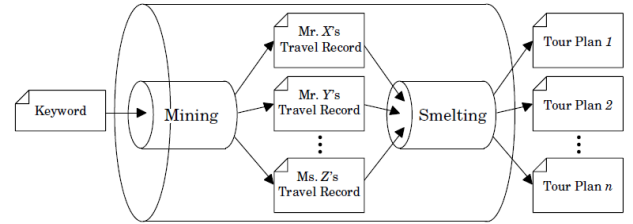

Fig. 1. Overall framework of the Tour Miner. This paper focus on the first part of the process called Web Mining.

the process will be combined in order to generating the tour plan and deliver the result to the user.

2) Web Mining: Web Mining is initial process of the tour miner system. The goal of this phase is to crawl and retrive the user experience from the SNS. The data preparation technique is then apply to the retrived data in order to make the data ready for the create and update the offline database. Offline database will allow us to perform querying with fast and effective to retrieve numerous of data to pass on the process mining.

3) Process Mining: Process mining techniques extract the knowledge from event logs to discover, monitor and improve processes. Process discovery is the most important technique. Distory technique takes an event log and produces a process model such as Petri Net and BPMN.

III. Mining Function from Social Network System

In this section, we elucidate the implementation of the mining function. We first illustrate the outline of the function and then detail of each step and detail of the function.

1) Outline: This section will explain the process of web mining which consist of three major step (See Fig. 2). First, the getting user data and the create the travel record, and closing up with storing the information into the database. Majority of data preparing technique, such as the removing of dimensionality, is applied but omitted in the paper due to the complexity and redundancy.
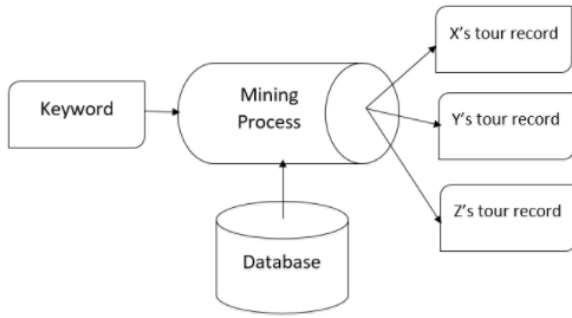
Fig. 2. Outline concept of the mining part of tour miner. This is the process of the tour miner system.

Fig. 3. All the user who most likely to travel's information will be stored for further use in the next step.

2) Getting Traveling User from SNS: First, we choose the Twitter as the platform to gathering user experience from. Our method start by crawling through the user who most likely to travel to different places in the past by using the specific keyword.We use Twitter API's search function by looking through the whole Twitter server with the our specified parameter, then retrieve the qualified information (See Fig. 3).

3) Create the Traveling Record for each User: After we gather the numerous number of user from the previous process, we will focus primarily on their timeline. We utilize the Twitter's API timeline function in order to looking through each user timeline and collect each of their Tweets that contain the geolocation. Each of geolocation indicate that the user has been on that location and post on their social media which mean that we can connect we can use these information to generate the traveling record.

4) Preparing and Storing into the databasee: After we retrieve each of the user travel record from crawling through their timeline, we need to clean the data in order to simplify its to be able to function efficiently in the process mining part. The data will be mapped with the Google Map API reverse geolocation technique to transform the geolotaion information into the real world location. We then stripped the specific location into the well-known location, in our case we use prefecture form in order to minimize the noise. The example output will look like the point to point connection (See Fig. 4)., this is due to the fact that our database is in the form of

Fig. 4. The final format for the information before storing it into the database.

MySQL which require the specific row and column in order to function properly.

IV. IMPLEMENTATION

We create the website in order to implement how our function work by allowing any user input starting prefecture location into our website. After the user input the keyword, the website user interface will send the data to the server side, the server then query all the record that consist of selected prefecture and presented to the user. The example output from website can be seen in Fig. 5

V. CONCLUSION AND FUTURE WORK

In this paper, we propose a method to getting the traveling record from Twitter using the keyword and store it in the database. The strong point of our method is that the offline database allow fast response query and exponential growth for information as the time goes on.

As the future work, we are going to develop more interest based or more specific information for the user input and also connect the web mining function to process mining function.

REFERENCES

[1] K. Arai , "A method of sightseeing route recommendation using Twitter' DEIM Forum 2015 G7-6.
[2] T. Takuma, Y. Shingo, M. Bundit, R. Arnon, L. Pattara , "Tour Miner: Mining System of Tour Plans from SNS'
[3] Z. Gou, S. Tsugawa, A. Korkaew, S. Yamaguchi, T. Khamket, B. Manaskasemsak,A. Rungsawang, An Interest-based Tour Planning Tool by Process Mining From Twitter,

Fig. 5. The final format for the information before storing it into the database.