

Recommending Tourist Locations Based on Data from Photo Sharing Service: Method and Algorithm

Andrew Ponomarev^{*†}

^{*}SPIIRAS, St.Petersburg, Russia

[†]ITMO University, St.Petersburg, Russia
ponomarev@iias.spb.su

Abstract—Tourists' information support is more actual than ever, objectively because tourism is one of the largest and fastest-growing economic sectors and subjectively because each tourist faces unfamiliar and dynamic environment, which he or she has to adapt to. One of the ways to deliver information support to tourist is various recommender systems. Classical way to build recommender systems requires either collection of ratings (collaborative filtering system) or extensive knowledge work on describing tourism domain and attractions of each area. However, there is another, more lightweight approach – to make recommendations based on social media analysis. This paper presents a method and an algorithm for identifying potentially interesting locations based on Flickr photo sharing site media stream. One of the particular problems addressed in this paper is to reduce the number of queries to the Flickr API.

I. INTRODUCTION

Tourism has become one of the largest and fastest-growing economic sectors in the world. Despite occasional shocks, it has shown virtually uninterrupted growth. International tourist arrivals have increased from 527 million in 1995 to 1133 million in 2014. International tourism receipts earned by destinations worldwide have surged from US\$ 415 billion in 1995 to US\$ 1245 billion in 2014. Moreover, the number of international tourist arrivals worldwide is expected to increase by an average of 3.3% a year over the period 2010 to 2030 [1].

On the other hand, there are some structural and behavioral changes in tourism highly connected to the development of Internet and Information Technologies. The increasing use of ICTs in tourism services allows tourists to take a more active role in the production of tourism products, being no longer satisfied with standardized products. The "postmodern tourist" with differentiated life-styles, individual motives and specific interests demands products tailored accordingly to stated preferences [2].

All that makes the problem of tourists' information support more actual than ever. Therefore, information (and search) services of all kinds that can help in collecting information about the trip being planned and provide tourist with information needed during the trip are becoming more and more popular. One of the functions typically provided by those services is recommendation of attractions based on tourist's preferences and current conditions (weather, transport, etc.).

Systems intended to mitigate a choice problem leveraging (implicit or explicit) subjective preferences received a name of "recommendation systems". The variety of techniques to

build, deploy and assess this kind of systems separated into a specific research area in the mid-90s of XX century.

Approaches to build recommendation systems are usually classified according to the kind of input data that is used for recommendations. Most popular are two of them [3]: collaborative filtering and content-based. In the former one the only information that is available are ratings that users assigns to objects. In the latter, input information is formed by structured representation of items and a vector of user's ratings. There are several more approaches: demographic recommendation systems, knowledge-based recommendation systems, social-based recommendation systems, but they are less used.

Dependence on a specific type of information causes limitations in applying each of recommendation techniques. For example, collaborative filtering cannot be used when the number of ratings is small, but just after start of any recommendation system the set of ratings is usually empty, hence the so called "cold start" problem. Similarly, the structured representation of items needed for content-based methods might be available for one regions (in tourist recommender systems) and be missing for others.

The classical approaches to build recommendation systems mentioned earlier are personalized. It means that they provide potentially precise recommendations matching one's preferences, but they require substantial amount of information about user to account for his/her preferences. This information is not always available, and that results in some immanent problems of these recommendation systems, e.g. "cold start" problem. For application contexts where there is lack of preferences information, there are other approaches providing lightweight non-personalized recommendations.

Non-personalized recommendations are based on visiting statistics data. There are three potential sources of these data: a) it can be collected by the tourist application itself; b) it can be queried for from be local authorities or POI administration; c) it can be mined from the global stream of public data. The source (a) is the most convenient as the data can be collected with all the needed context attributes and in the most appropriate form and granularity, however, it requires a huge number of users and cannot be employed by a newly created application. The source (b) relies on the communication with external entities (local authorities and museums administration) and is very laborious. It can be appropriate for

a local application, e.g. St. Petersburg local city guide, but hardly can be implemented for a global recommendation service that should work in every location worldwide. Moreover, it is not suitable for recommending architectural POIs, publicly available observation places, as there might be no administration to collect visiting statistics. With all the drawbacks of (a) and (b) for making a globally active POI recommendation service, the option (c) becomes viable. With the dissemination of camera and GPS-equipped mobile devices, widening mobile internet coverage and the forming of new information processing habits, publicly available stream of geo-tagged events is becoming more and more affluent. There are many scientific publications showing various ways of leveraging this live source of human activity: from events and opinion detection to, a more relevant to the topic of this paper, POI detection and recommendation (see, e.g. [4]). However, this stream obviously bears some bias that must be taken into account. E.g., it is produced by active users of social networks and owners of modern smart-phones. Target users of mobile tourist guide are obviously a subset of smart-phone owners, but actually not necessarily are active users of social networks, so there still is a chance of biased inference.

In some sense, point-of-interest detection and recommendation is media stream analysis can be interpreted as an “open form” of collaborative filtering. Openness here means that this technique does not implement functionality of user’s feedback collection typically present in recommendation systems; instead, it relies on some external feedback source, namely, social media event stream. A common idea of this kind of systems is that geo-tagged images are interpreted as some signs that a user who posted them enjoyed the place or view.

This paper follows this lightweight non-personalized recommendations approach. It proposes method and algorithm for identifying city locations that are likely to be most interesting for tourists based photos published on Flickr photo sharing service.

The rest of the paper is organized as follows. Section II presents the discussion of related work. Section III describes relevant fragments of the Flickr API. Section IV contains the description of the proposed algorithms, and Section V provides some data on performance and quality evaluation of the proposed method and algorithm.

II. RELATED WORK

Key premises to employ recommendation systems in some domain are, first, abundance of choice complicating decision-making, and second, significant subjectivism in decision-making. Tourist in an unfamiliar environment (city, country) frequently face both of the premises: abundance of attractions to visit and uncertainty in which of them to visit to gain most positive experience from the stay. This explains the attention that is paid to tourist information support in recommendation systems community. Besides, social sciences research reveals the importance of decision support systems in tourism, caused by large number of aspects that need to be paid attention to: tourist mobility, high risk and uncertainty in unfamiliar environment, distributed nature of information sources and several other factors [5].

Main directions and achievements in tourist recommendation systems design are summarized in review papers [6] (systems before fall 2009) and [7] (2008-2014). These studies reveal that nowadays in tourist recommendation systems all modern recommendation techniques are used. Collaborative filtering, content-based and demographic ones are the most widely employed.

There are several papers dedicated to point-of-interest (or landmark) detection based on media stream analysis.

In [8] methods are proposed to detect actual events taking place in city based on the Twitter stream.

Han & Lee use photo stream (Flickr photo sharing service, to be precise) for landmark detection [4]. Moreover, they employ some kind of additional analysis to make further inference. E.g., they try to distinguish images made by tourists (which are mostly relevant for making recommendations to tourists) from images made by local population by analyzing attributes of the image poster’s account.

Photo2Trip system makes step further; based on the analysis of sequences of geo-tagged photos from public photo sharing sites, Photo2trip identifies and recommends typical tourist trips [9], [10].

However, most of the cited works are designed for a predefined territory, so they can load all photos related to that territory in advance and process them with variety of “complete information” data processing methods. The problem is that getting a complete information (including coordinates) about each photo via Flickr API is time-consuming to the extent that makes it impractical for large territories and time-frames. The contribution of this paper is to build a non-personalized location recommendation system based on media stream analysis that could be used without restriction in any point on the Globe, therefore, there’s a need to address a speed impediment resulted in using Flickr API. Therefore, the paper proposes a method and an algorithm to find interesting locations without information about each particular photo.

III. BRIEF FLICKR API DESCRIPTION

Flickr photo sharing service provides extensive API [11] allowing to access many if not all features of the service from making queries of recent photos in the specified area to posting new ones. This section briefly describes only those capabilities of the API that are relevant to the problem of detection of interesting places by means of geo-tagged photo density analysis.

In general, Flickr API uses OAuth authentication, but some types of calls can do without it. The delineation is quite intuitive – those calls that deal with private information usually kept “within” user’s space require authentication, and those that are indifferent to specific users do not.

The most useful from the purpose of this paper Flickr API call is `flickr.photos.search`. This call returns a list of photos matching some criteria. The call may be used by either authenticated or not authenticated client. However, for the not authenticated client the returned list will consist of only public (visible to anybody) photos.

Criteria that can be specified as a parameter for this call include tags, visibility, content type and many other, but the most relevant to this paper are:

- Minimum and maximum date the photo is taken.
- Bounding box (in geographic coordinates) to limit the area where photos are searched.

The result of this call provides not the complete information about the photos found, but rather their descriptors. For example, a call might return an XML document like this:

```
<photos page="1" pages="89"
  perpage="2" total="881">
  <photo id="2636"
    owner="47058503995@N01"
    secret="a123456" server="2"
    title="test_04" ispublic="1"
    isfriend="0" isfamily="0" />
  <photo id="2635"
    owner="47058503995@N01"
    secret="b123456" server="2"
    title="test_03" ispublic="1"
    isfriend="0" isfamily="0" />
</photos>
```

The most important information in this XML is `total` attribute of the `photos` element, corresponding to the number of photos satisfying the criteria specified in the call. To get an additional information about photos one has to use other API calls providing the unique identifiers of photos (attribute `id` of `photo` element). For example, to get exact geographic coordinates associated with the photo `flickr.photos.geo.getLocation` can be used.

The call `flickr.photos.geo.photosForLocation` can be used as an alternative for `flickr.photos.search`, but it is less flexible as it doesn't allow to restrict the time range when selected photos were taken.

IV. PROBLEM DEFINITION AND A PROPOSED SOLUTION

The goal of the proposed system is to recommend interesting locations in the area unfamiliar to the user. Recommendation systems are usually based on some assumption that simplifies original recommendation problem and helps to build rigorous mathematical model. In this case, the assumption is that people tend to make and share photos of the places they find interesting and attractive. That is the same class of places that are usually recommended to the tourists visiting the area. Therefore, geo-tagged photos can be interpreted as some kind of "votes" for tourist attractiveness of the location. Of course, this assumption is not always true, however, validation results presented in Section V show that it can lead to rather useful results and allow to reach the goal of finding interesting locations to recommend with reasonable quality.

To fulfil the original goal there has to be a method and an algorithm to identify local clusters of photos taking into account that they are stored in Flickr photo sharing service and can be accessed only via Flickr API. It means that finding out precise coordinates of a photo is possible only by special query (one query for each photo), but it is also possible to easily (with

one query) to find out the number of photos in a rectangular area. Time required to execute an API call is high enough to make collecting precise coordinates of all photos impractical, therefore, local clusters should be approximately identified using aggregate data on rectangular areas.

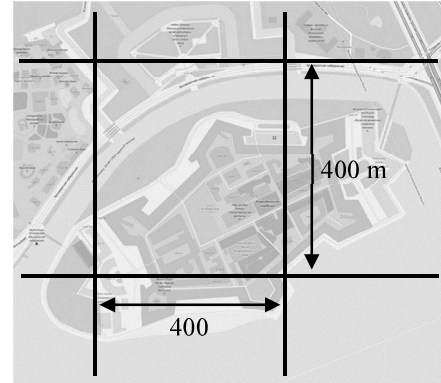


Fig. 1. Grid fragment example

All the examined region (currently, a city) is split into a number of rectangular areas forming grid with cell size of about 400×400 meters (see Fig. 1). The size is influenced by the following factors. The bigger the cell is the less calls to the Flickr API are needed, and it reduces the time required to obtain data. On the other hand, large cells bear too much uncertainty about actual location of interesting places. E.g. it is rather hard to find something interesting if all you know that it is somewhere in the square with side of 1 km. The selected size of 400 meters presents one of possible compromises. It is large enough, but also can be examined in several minutes of walk.

For each of the cells it is possible to obtain the number of photos in the cell via Flickr API. Then, the task is to find some kind of "outstanding cells" of the resulting matrix. However, in general case, the number of such cells (even as large as 400×400 meters) can be quite large. For example, the size of St.Petersburg and its suburbs is about 40×40 km, which results in 10000 cells and therefore 10000 calls to Flickr API. It may be acceptable for a system targeted to one city, but for universal system supporting many cities it becomes too time consuming and may result in blocking by Flickr for abusing.

Hence, there are two tasks:

- 1) To define a criterion for selecting cells as potentially interesting. It might be local maximum or something entirely different.
- 2) To develop a method for effective pruning of unnecessary calls to FlickrAPI (about the cells that most probably are not interesting according to the defined criterion).

For the first task, the proposed method is to find 10% of cells containing most photos. In other words, all cells belonging to the area being examined are sorted in descending order by the number of photos attributed to them. The first 10% of cells in this sequence are considered to contain some potentially interesting sites and are recommended to the user.

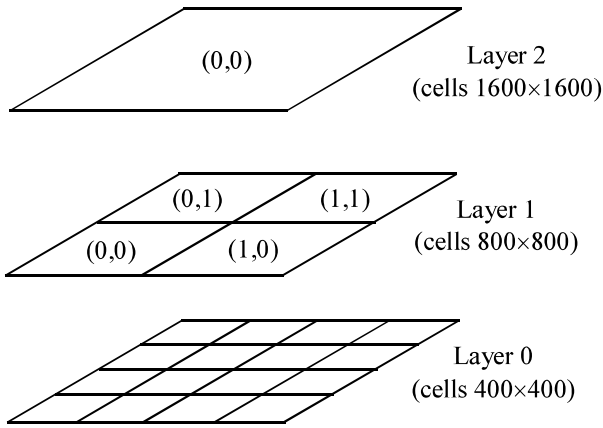


Fig. 2. Aggregate layers structure

For the second task, the proposed method is to consider several layers of aggregation over the initial layer consisting of 400x400 meters cells (layer 0). Each aggregated layer i also consists of square cells but the side of the cell of layer i twice as big as the side of layer $i-1$ (see Fig. 2). Moreover, cell bounds in different layers are aligned in such a way that each cell of layer i consists of exactly four cells of layer $i-1$. Photo counts of cells in aggregated layers can also be found out via Flickr API. Obviously, there is a simple relation between photo counts in different layers. Let all rows and columns of cells in each layer i be numbered from 0 to $n^{(i)}-1$. Let, also, $q_{ij}^{(k)}$ denote the number of photos in the cell in row j and column i located in the layer k . Then, by the construction of the aggregate layers:

$$q_{c,r}^{(k)} = \sum_{j=0}^1 \sum_{i=0}^1 q_{2c+j, 2r+i}^{(k-1)}.$$

As $q_{ij}^{(k)} \geq 0$, then:

$$q_{2c+j, 2r+i}^{(k-1)} \leq q_{c,r}^{(k)}, \quad j, i \in \{0, 1\}.$$

That gives an idea for possible pruning condition. If at some point of the search process we have identified the needed number of layer 0 cells (10% of the total number) and at the same time we have not evaluated all the cells in the layer 0, but have evaluated only cells of layers 1 and higher, and the respective values of that higher layer cells are less than the values of identified layer 0 cells, we can stop the search process.

So, the search procedure can be organized as a form of branch and bound search, when we start from the highest layer (say, 3rd, but it can be even 4th, depending on the actual size of the area being examined), evaluate all the values of $q_{ij}^{(3)}$. Then pick a cell to branch (bc , br), and branching in this context means descending to the lower layer, and evaluating cells $q_{2*bc+i, 2*br+j}^{(2)}$, $i, j \in \{0, 1\}$. Then again pick a cell to branch, and so on until values of the required number of layer 0 cells are known and the rest cell values (in any layer) are less than them.

The final thing to be defined is a heuristic to determine the cell to branch on (branching heuristic). This heuristic is important as it significantly affects the overall performance of the procedure. It should balance the necessity to evaluate the needed amount of layer 0 cells as soon as possible to reach the stop condition, and undesirability to evaluate many cells at all. Two heuristics are proposed and evaluated.

The first one (*PICK-MAX*) is to pick each time a cell with the highest number of photos no matter what layer this cell belongs.

The second one (*PICK-ADJ*) is to pick each time a cell with the highest normalized adjusted number of photos $a_{ij}^{(k)}$, where adjusted number is defined by the following equation:

$$a_{c,r}^{(k)} = \frac{q_{c,r}^{(k)}}{4^k}.$$

The purpose of the adjustment is to equalize cells of different layers; for higher layers it corresponds to the expected number of photos in the layer 0 cells covered by this cell assuming uniform distribution of photos.

Algorithm 1 Interesting cells search

Input:

- L – starting layer, $L \geq 0$

- $n^{(L)}$ – the number of rows/columns in the starting layer, $n^{(L)} > 0$

```

0:  $Q \leftarrow \emptyset$ 
1:  $S \leftarrow \lfloor n^{(L)} * n^{(L)} * 4^L / 10 \rfloor$ 
2:  $R \leftarrow \emptyset$ 
3: for  $c \in n^{(L)}$  do
4:   for  $r \in n^{(L)}$  do
5:      $Q \leftarrow Q \cup \{(c, r, L, count(c, r, L))\}$ 
6: while True do begin
7:    $(c, r, l, q_{c,r}^{(l)}) \leftarrow PICK-MAX(Q)$ 
8:   if  $|R| \geq S$  and  $ordered(R)[S][4] > q_{c,r}^{(l)}$  then
9:     return  $R$ 
10:   $(c, r, l, q_{c,r}^{(l)}) \leftarrow pick\_cell(Q)$ 
11:   $Q \leftarrow Q \setminus \{(c, r, l, q_{c,r}^{(l)})\}$ 
12:  if  $l \neq 0$  then begin
13:    if  $|R| \geq S$  and  $ordered(R)[S][4] > q_{c,r}^{(l)}$  then
14:      continue
15:    for  $j \in \{0, 1\}$  do
16:      for  $i \in \{0, 1\}$  do
17:         $Q \leftarrow Q \cup \{(2*c+j, 2*r+i, l-1, count(2*c+j, 2*r+i, l-1))\}$ 
18:  else
19:     $R \leftarrow top(S, R \cup \{(c, r, l, q_{c,r}^{(l)})\})$ 
20: end
    
```

In the Algorithm 1 Q is the queue of all cells that should be examined and either included to the result set R or pruned out. One element of Q is a tuple consisting of cell coordinates inside layer, layer number and the number of photos in the cell. S is the number of “interesting” layer 0 cells the algorithm is looking for. The algorithm depends on several functions, which

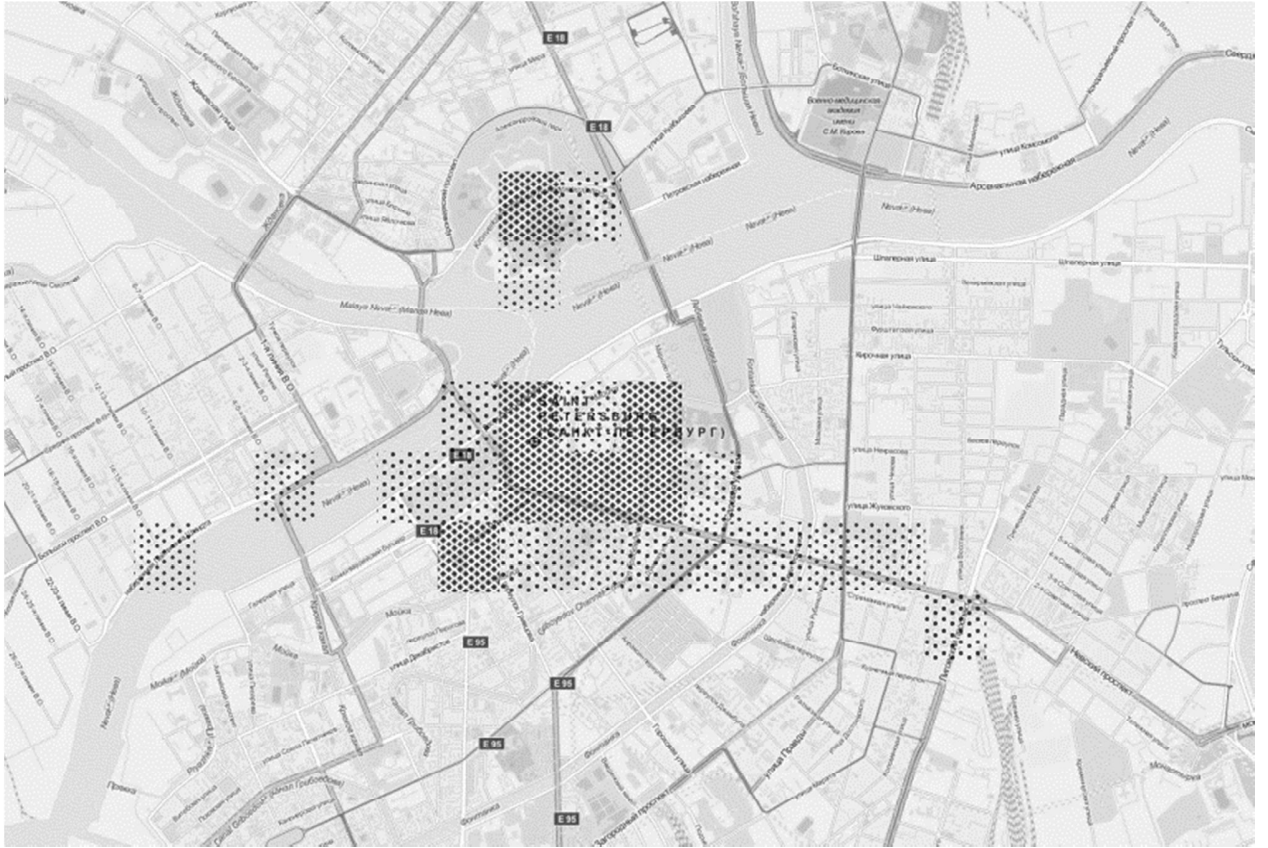


Fig. 3. Selected locations in St.Petersburg based on data from first half of 2015

are:

- $count(c, r, l)$ – function that evaluates $q_{c,r}^{(l)}$ via Flickr API;
- $pick_cell(Q)$ – selects an element of Q according to the employed heuristic, either *PICK-MAX* or *PICK-ADJ*. These heuristics are also implemented as functions – *PICK-MAX*(Q) retrieves a cell with the greatest number $q_{c,r}^{(l)}$ from the queue Q , *PICK-ADJ*(Q) retrieves a cell with the greatest ratio $q_{c,r}^{(l)}/4^l$;
- $top(S, Q)$ – a function that returns S elements with the highest values $q_{c,r}^{(l)}$ from the Q , or the entire Q , if $|Q| \leq S$.

Lines 3-5 of the algorithm add all cells of the highest (starting) layer to the queue to examine them later.

Lines 6-20 form the main part of the algorithm. On each iteration of the cycle finish condition is checked. Expression “ordered(R)[S][4]” means 4th element of the 4-tuple that is in the S th position in the R ordered by descending (of the 4th tuple elements). If all queued elements are less than the best S known cells of layer 0, then the algorithm is stopped and R is returned. Otherwise, an element is picked according to one of the examined heuristics and either added to the result (if it is layer 0 cell) or decomposed to cells of lower layer.

The algorithm is provided in general form, however if *PICK-MAX* heuristic is employed, lines 10, 13, and 14 become

unnecessary, as the cell with the largest count is selected in the line 7 to check the stop condition.

Obvious improvement of lines 15-17 is based on the relationship between counts of higher-level and lower-layer cells. If the count of an upper cell is known and counts of any three inner cells are known then the count of the fourth inner cell can be easily calculated without a call to FlickrAPI. The algorithm implementation used in experiments takes advantage of this improvement.

The algorithm depends on selection procedures from Q . Efficient implementation of these selection procedures can be based on heap data structure (see, e.g., [12]). However, as there are two criteria of ordering (photos number and adjusted photos number), there must be two heaps – one for each criterion – synchronized on modification.

V. EXPERIMENTS AND VALIDATION

The aim of the experimentation part is twofold. First, it should validate the whole idea of detecting interesting places of an area by selecting area rectangles that contain most photos from some popular photo sharing site (e.g., Flickr). Second, it should verify that the number of queries to the Flickr API is reduced due to the developed method and algorithm.

To validate the whole idea that lies behind the interesting places identification, a following experiment was performed. Two different cities were selected: St.Petersburg, a big cultural center (Fig.3) and Tyumen, a middle-sized regional center in

Siberia. The rationale of selecting two cities was that the detection procedure may depend significantly on the city size or cultural status.

In each of the cities, five “experts” were selected. The experts had higher education, mostly technical, but no special cultural education or training. Among experts, there were both male and female, and they belonged to rather wide age group of 25-60 years old. Each expert had lived in the city he/she was asked about for at least five years.

Each of the experts was asked to mark areas on a city map that they would recommend to city visitors. Then cells of the layer 0 were detected corresponding to the marking of each expert. As cell sets selected by different experts were different, and there was no reason to prefer one expert opinion to another five joint etalon sets were constructed. The first etalon set contained the cells selected by at least one expert, the second one contained the cells selected by at least two experts and so on with the fifth set containing cells selected by all five experts. Sizes of the sets are shown in the Table I.

TABLE I. SIZES OF THE JOINT ETALON SETS

Set number	St.Petersburg	Tyumen
1	70	29
2	36	14
3	27	9
4	15	6
5	7	3
Number of cells in the area	576	256

Quality of the output produced by the proposed algorithm was evaluated by widely used in information retrieval measures precision and recall. In the context of the considered task, precision is the probability that randomly chosen cell retrieved by the algorithm is among cells, selected by the experts. Recall is much like the opposite, it is the probability that randomly chosen cell from the set, selected by experts, is actually found by the algorithm. Recall and precision for both cities are presented in the Table II.

TABLE II. QUALITY MEASURES OF THE PROPOSED METHOD

Set number	St.Petersburg		Tyumen	
	Precision	Recall	Precision	Recall
1	0.68	0.56	0.6	0.52
2	0.52	0.83	0.4	0.71
3	0.40	0.85	0.32	0.89
4	0.22	0.87	0.2	0.83
5	0.12	1.0	0.12	1.0

It can be seen, that in both cities recall is rather high, which means that the proposed method was able to detect most of the places that were selected by experts. Precision, on the other hand, is not as high, which means that the method detects many cells (places) which are not marked as recommended to visit by human experts. To some extent, it can be controlled by the parameter of the selection criterion, i.e. instead of 10% of the cells with most photos one can use 5% or even 1%. However, that will inevitably affect recall.

F1 score is widely used as a single quality measure instead of precision/recall pair:

$$F1Score = 2 \frac{Precision * Recall}{Precision + Recall}.$$

In this paper, F1 score is used to check if the selection of top 10% (containing most photos) cells is adequate. For that purpose, F1 score of top n cells was evaluated on each etalon set for n ranging from 5 (five top cells of the area) to 10% of the area cells (57 for St.Petersburg and 25 for Tyumen). The results for St.Petersburg and Tyumen are in Fig. 4 and Fig. 5 respectively.

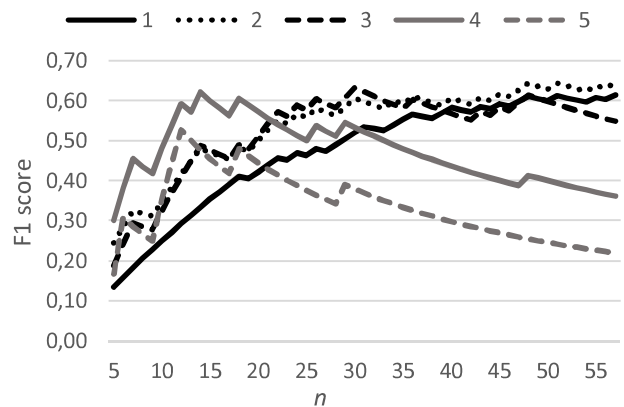


Fig. 4. F1 score for St.Petersburg

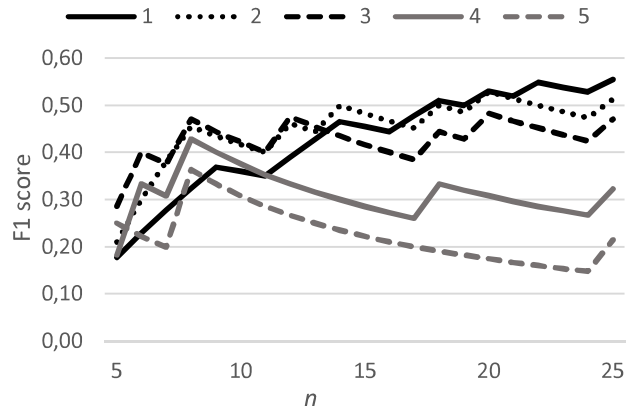


Fig. 5. F1 score for Tyumen

It can be seen from the figures, that stricter selection criteria will not improve overall quality (measured by F1 score).

For Tyumen, the results are worse than for St.Petersburg. Probably, that may be explained by the fewer active Flickr users, but further investigation needed.

To measure the efficiency of different branching heuristics the number of calls to Flickr API was assessed for values of selection criterion from 1% to 10%. The results are presented at Fig. 6. The experiment was performed in the St.Petersburg center area where the total number of cells was 576. It can be seen from the figure, that both branching heuristics give very similar results, *PICK-MAX* being not worse than *PICK-ADJ* for

all tested selection criterion values (and sometimes slightly better). Absolute numbers of calls to Flickr API for detection of top 10% cells is about half of the number of cells.

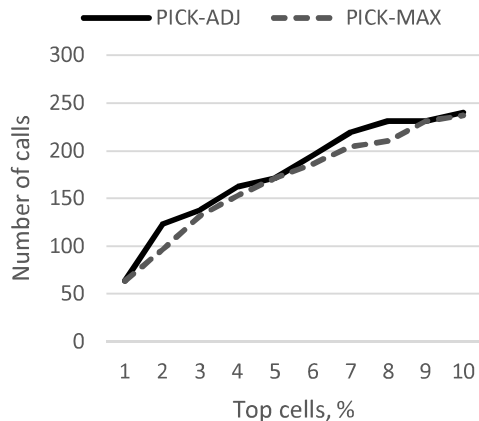


Fig. 6. Number of calls to Flickr API for different values of the selection criterion

VII. CONCLUSION

The paper proposes a method and an algorithm for interesting locations identification based on the publicly available photo stream of Flickr photo sharing service. One of the goals of this paper is to minimize the number of calls to Flickr API, as it is an expensive operation.

The proposed method and algorithm were validated in an experiment which showed that significant amount of city places marked as “interesting for city visitors” by human experts were correctly identified using Flickr data.

The proposed method and algorithm can be used alone, or, and it seems to be more fruitful, as a part of complex tourist information support suite (like one described in [13]), where the photo stream parameters can be joined with textual and structured descriptions of attractions and create an alternative for collaborative filtering systems suffering from the lack of user preference data.

Future work can be organized into four directions:

- 1) Tune and optimize parameters of the algorithm (examine other grid cells, other selection criterion parameters, and entirely other selection criteria).
- 2) Add context support (account for shot time and authors role).
- 3) Integration with tourist information systems. One possible way to do that is to select specific attractions situated in the most “popular” squares, detected by the algorithm presented in this paper.

4) Experiments with data sources other than Flickr.

ACKNOWLEDGMENT

The presented results are a part of the research carried out within the project funded by grants #16-07-00463 and #16-07-00466 of the Russian Foundation for Basic Research. This work was also partially financially supported by the Government of the Russian Federation, Grant 074-U01.

The author would like to thank Elena Puchkova for the help in interviewing Tyumen experts group.

REFERENCES

- [1] World Tourism Organization, “UNWTO Tourism Highlights 2015 Edition”, Web: <http://www.e-unwto.org/doi/pdf/10.18111/9789284416899>.
- [2] T. Berka, M. Plößnig “Designing recommender systems for tourism”, In *11th International Conference on Information Technology in Travel & Tourism*, ENTER, 2004, Cairo, Web: <http://195.130.87.21:8080/dspace/handle/123456789/583>.
- [3] G. Adomavicius, A. Tuzhilin “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions” // *IEEE Transactions on Knowledge and Data Engineering*, June 2005. vol. 17, no. 6, P. 734-749.
- [4] J. Han, H. Lee, “Adaptive landmark recommendations for travel planning: Personalizing and clustering landmarks using geo-tagged social media” // *Pervasive and Mobile Computing*, 2015, Vol. 18. pp. 4-17.
- [5] Gretzel U. Intelligent systems in tourism: A social science perspective. *Annals of Tourism Research*, 38(3), 2011, 757-779.
- [6] Kabassi K. Personalizing recommendations for tourists. *Telematics and Informatics*, 27(1), 2010, 51-66.
- [7] Borràs J., Moreno A., Valls A. Intelligent tourism recommender systems: A survey // *Expert Syst. Appl.* 2014. T. 41. C. 7370-7389.
- [8] A. Marcus, M.S. Bernstein, O. Badar, D.R. Karger, S. Madden, R.C. Miller, Twitinfo: aggregating and visualizing microblogs for event exploration, in *Proceedings of the 2011 Annual Conference on Human factors in computing systems*, CHI’11, ACM, New York, NY, USA, ISBN: 978-1-4503-0228-9, 2011, pp. 227-236.
- [9] X. Lu, C. Wang, J.-M. Yang, Y. Pang, L. Zhang, “Photo2trip: generating travel routes from geo-tagged photos for trip planning”, in *Proceedings of the International Conference on Multimedia*, MM’10, ACM, New York, NY, USA, ISBN: 978-1-60558-933-6, 2010, pp. 143-152.
- [10] H. Yin, X. Lu, C. Wang, N. Yu, L. Zhang, “Photo2trip: an interactive trip planning system based on geo-tagged photos”, in *Proceedings of the International Conference on Multimedia*, MM’10, ACM, New York, NY, USA, ISBN: 978-1-60558-933-6, 2010, pp. 1579-1582.
- [11] Flickr API Description, Web: <https://www.flickr.com/services/api/>
- [12] T.H. Cormen, C.E. Leiserson, R.L. Rivest, C. Stein, *Introduction to Algorithms* (3st ed.), MIT Press and McGraw-Hill, 2009.
- [13] Smirnov A., Kashevnik A., Ponomarev A., Shilov N., Shchekotov M., Teslya N., Smart Space-Based Intelligent Mobile Tourist Guide: Service-Based Implementation, *Proceedings of the 15th Conference of Open Innovations Association FRUCT*, Saint-Petersburg, Russia, ITMO university publisher house, 21-25 April 2014, pp. 126-134.