International Journal of Innovative Computing, Information and Control
Volume **17**, Number **1**, February **2021**

ICIC International ©2021 ISSN 1349-4198
pp. 345–353

CONSUMER BEHAVIOR ANALYZER IN INTERNET OF THINGS (IOT) ENVIRONMENTS

Swe Nwe Nwe Htun[1], Thi Thi Zin[2] and Pyke Tin[3]

[1]Interdisciplinary Graduate School of Agriculture and Engineering
[2]Graduate School of Engineering
[3]International Relation Center
University of Miyazaki
1-1, Gakuenkibanadai-Nishi, Miyazaki 889-2192, Japan
{ swenwenwehtun; pyketin11 }@gmail.com; thithi@cc.miyazaki-u.ac.jp

Received August 2020; revised December 2020

Abstract. *This paper proposes an analyzer of consumer behavior in Internet of Things (IoT) environments. This analyzer is most useful in predicting the intentions of users during searches, and especially during image searches. Since most technologies are connected on the Internet, search results can be characterized using image-similarity measures. In this paper, information on image similarities is extracted using a Convolutional Neural Network (CNN) in IoT environments. In this proposed consumer behavior analyzer, the similarity measures characterizing the relationships between images are transformed into Markov Chain transition probabilities, and their stationary probabilities are then analyzed to describe the priority order for search results conforming with consumer intentions. In order to confirm the validity of the proposed method, the Yelp public dataset was used. The outcomes using this analyzer are promising, and this analyzer might be instrumental in making further improvements in practical applications of consumer technologies.*

Keywords: Consumer behavior, IoT environment, Markov Chain, Convolutional neural network

1. **Introduction.** Today's consumer technologies are emerging fast and effectively into the world of the Internet of Things (IoT), providing great potential for consumers by utilizing the technology in smart ways [1,2]. As these technologies advance, the behavior of consumers in utilizing and seeking information about them is also changing. The users of these technologies may have different intentions for using the same product depending on the situation. User intent must be taken into account in developing new technologies. A variety of methods can be used to study user intent, but this paper focuses on user intent within the framework of information searching behavior, keeping in mind the interaction of consumers and technology. On the other hand, perceived benefits of IoT networks have motivated consumers to interact with these technologies more and more, wanting to get the most out of their searches for information on the web. The approach for improving the search function must be specific and will require fine tuning. Again, understanding the user's intention is an important part of improving the retrieval of information for the user. It is worthwhile to mention that information is not only available in text, but also as images on the Internet. Effectively searching for image-based information could provide great benefit to consumers. However, though much has been published on search systems, no satisfactory way of improving the search function has yet been developed, especially for image searches.

DOI: 10.24507/ijicic.17.01.345

Only interpreting the user's search interests through textual keywords is not the right path, as it provides noisy search results, and deviation from the user's intent [1]. In spite of the huge amount of information retrieved in a web search, people usually view only the topmost relevant results for their query [3]. The traditional approach to performing a search involves text processing methods, such as for the text in the body of a page, the webpage address, keywords, and hyperlinks [4]. Search results are displayed according to the page rank. However, users do not always get satisfactory results because of deviations from their intentions, and the fact that visual similarities are not taken into account. In addition, pages are mostly ranked using hyper-link analysis, which is not always useful. Previous ranking methods depended on a PageRank algorithm developed for large-scale search engine, which makes heavy use of this hyperlink structure [5]. A gap exists between research into the use of different weights on the types of links (inter-site links versus intra-site links), and research into the probability of being voted important. An argument can be made that this complex link structure is the most common reason for deviation from the users' intent.

In most conventional image-search engines, user intentions are represented by textual query. The purpose of inputting a textual query to find a visual image is to extract results and ranks most relevant to the query. However, some limitations with such systems are apparent due to the structure of the model, especially in the classification process. Deviation from the user intent may result, as expressed in the natural characteristics of images, and in differing visual representations of object categories. Text-based queries have multiple visual heterogeneous concepts in object categories. Let us take the query "Apple" as an example. Some are expecting images of an "Apple" computer, while others are expecting images of fruit [6]. This type of situation occurs frequently, particularly for queries with no specific concepts. Understanding user intentions in the face of such linguistic ambiguity will remain a problem for web image search systems [7,8].

A great source of information on user intentions can be regarded as the feedback people provide on social networks. Social networks capture people's attention, which generates much information on search behavior as people navigate a vast sea of information. To make search engines more effective, user interest can be researched on ever-growing information sharing networks. Useful information includes posted pictures, likes, and shares by users [9]. Publicly posts with photos on social network get more likes than those without since images are subject to more interpretations than mere text. With this as a background, we have established a consumer behavior analyzer to improve the relevancy between returned images and user intentions. Specifically, we propose a system that takes account of user intentions in searching and re-ranking consumer information on a community-specific platform. For this purpose, we utilize a correlated Markov Chain model, and a state-of-the-art Convolutional Neural Network (CNN) architecture, along with the procedures for assessing image similarities, and determining user intentions.

2. **Proposed Method.** An overview of the proposed system is described in Figure 1. In the first module, the consumers' interest group is established by collecting data from the Yelp real-world dataset [10]. Yelp maintains a social network of users who evaluate the businesses by giving a star rating from 1 to 5 as shown in upper part of Figure 1. Shared on the network, this rating system provides the great value for users. Users can post photos, tagged with additional information on the business and their status. Moreover, friends and other users can vote on the usefulness of recommendations and can provide additional comments regarding their satisfaction with the business. These reviews and star ratings can effectively promote businesses through social networks. The Yelp dataset provides a wide range of opportunities for studying user intentions. We included such data in our

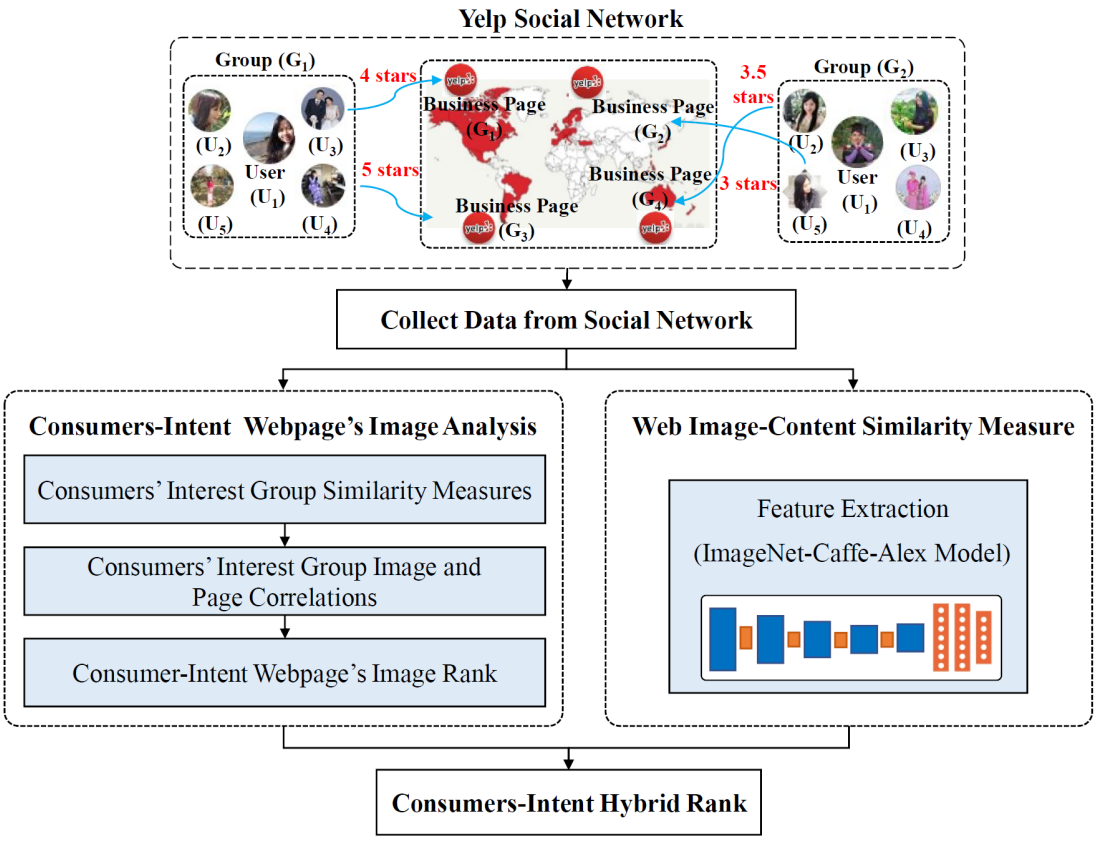

FIGURE 1. Overview of the proposed approach

research on how well information ranking in image searches reveals information on Internet users. We then describe similarity measurements between these established groups and their correlations, and present step-by-step computation procedure. After that, we embed these similarity measures and correlation structures into a correlated Markov Chain model. A transition probability matrix is derived that represents the changes in user intentions from one image to the next. The second module presents an analysis of image-content by applying a convolutional neural network model for visual similarity matching. These two modules are combined to produce the optimized result. The following sections provide technical details for the proposed computation processes. The theoretical concepts used in this system are based on those described in [11,12].

2.1. **Consumers-intent webpage's image analysis.** In this module, we first describe the assumptions for consumers' interest group similarity measures, and then embed these similarity measures into the Markov Chain method. By applying the outcome of Markov transition matrix, we produce the correlation structure matrix for the interest group.

2.1.1. *Consumers' interest group similarity measures.* Suppose there are N groups in which the consumers share the same interests. We assume that corresponding groups are labelled as $G_1, G_2, G_3, \ldots, G_N$ and common ratios of A and B are defined as $CR(A, B)$ as in (1).

$$CR(A, B) = \frac{\text{number of elements in } (A \cap B)}{\text{number of elements in } (A \cup B)} \tag{1}$$

After that, the similarity measures within the groups G_i and G_j are formulated as in (2).

$$S_{ij} = \beta_1 CR(U(G_i), U(G_j)) + \beta_2 CR(P(G_i), P(G_j)) + (1 - \beta_1 - \beta_2) CR(I(G_i), I(G_j)) \tag{2}$$

where we denote that $U(G)$ is the set of users, $P(G)$ the set of corresponding pages, and $I(G)$ the set of corresponding images in those pages where $0 < \beta_1, \beta_2 < 1$. We then obtain the embedded Markov Chain with the transition matrix as shown in (3).

$$Q = [q_{ij}], \text{ where } q_{ij} = S_{ij} \Big/ \sum S_{ij} \tag{3}$$

Let $X = [x_1, x_2, x_3, \ldots, x_N]$ be a column vector, which satisfies Equation (4) described as follows, where X is needed to compute the correlations of images among interest groups.

$$X = QX \tag{4}$$

2.1.2. *Consumers' interest group image and page correlations.* To form the consumers' interest image group correlations, three groups are considered. The correlation of images and pages among the groups G_i, G_j and G_k is denoted by $\rho(i, j, k)$ and defined as described in Equation (5).

$$\rho(i, j, k) = (S_{ki} + S_{kj})S_{ij}x_ix_j \tag{5}$$

where x_i and x_j are computed in (4).

2.1.3. *Consumers-intent webpage's image rank.* Before defining the consumers-intent webpage's image rank, the mathematical notations are defined to understand assumptions involved in taking user interests into account between groups. The link from an image I_p to an interest group G_x is denoted as in (6).

$$L(p, x) = \begin{cases} 1, & \text{if } I_p \in G_x, \\ 0, & \text{otherwise} \end{cases} \tag{6}$$

Therefore, the degrees of links between the images and the corresponding pages which satisfy the conditions for consumers intentions can be seen as described in (7).

$$\text{degree}(I_p, I_q) = \sum \sum \sum L(p, x)L(q, y)\rho(x, y, z) \tag{7}$$

After that, the degrees of links between the images are normalized, and the consumers-intent image similarity matrix is obtained as in (8).

$$CI = [c_{pq}], \text{where } c_{pq} = \frac{\text{degree}(I_p, I_q)}{\sum \text{degree}(I_p, I_q)} \tag{8}$$

From the Markov transition matrix called *CI*, we then construct a stationary distribution vector, and use these probabilities that describe the image rank with respect to user intentions. Then, the Consumers-Intent Webpage's Image Rank (CWI Rank) is obtained.

2.2. **Web image-content similarity measure.** In this module, we analyze the visual contents of images for ranking similarities between images. The model used for image feature extraction (imagenet-caffe-alex) [13] was trained using a deep learning framework: ImageNet images with Caffe [14]. This model was constructed using 1,000 categories of ILSVRC2012, using 1.2 million training images. The output of the coupling layer is taken as the feature of the image. The dimensions of the extracted feature vectors are 4,096, which can be used for similarity measures. The result is regarded as the Web Image-content Similarity Rank (WIS Rank).

2.3. **Consumers-intent hybrid rank.** The optimized consumers-intent information is derived by combining Consumers-Intent Webpage's Image Rank (CWI Rank) and Web Image-content Similarity Rank (WIS Rank) as shown in (9).

$$\text{Consumer-Intent Hybrid Rank} = \alpha\text{CWI Rank} + (1 - \alpha)\text{WIS Rank} \tag{9}$$

where $0 < \alpha < 1$, and α is defined for taking consumers' intention into account equally in both modules.

3. Main Results.

3.1. **Experimental results by CWI analysis module.** According to analysis, the dataset includes 291,824 user groups, each containing a different number of users. We performed step-by-step procedures for understanding user intentions through the first two groups of users: Harald (G_1), and Christine (G_2). Two interest groups were assumed: G_1 and G_2. The first group was composed of 3,249 users, 37 pages, and 843 images. The second group was composed of 683 users, 68 pages, and 916 images. In the two interest groups, 3,811 users, 103 pages and 1,540 images were used. Therefore, we assumed the groups were as in (10).

$$G_j = \{U(G_j), P(G_j), I(G_j)\} \tag{10}$$

where $U(G_j)$ stands for the sets of users, $P(G_j)$ stands for the sets of pages, and $I(G_j)$ stands for the sets of images.

We next defined the similarity measures between the two-user interest groups and computed these similarity measures to form a Markov transition matrix, as in Equations (2) through (4). After that, we solved Equation (5) to correlate images among the interest groups in the next interest group correlation structure. By applying Equations (2) through (4), the results were obtained as shown in the following.

$$S_{11} = 1, \quad S_{12} = 0.0612, \quad S_{21} = 0.0612, \quad S_{22} = 1$$

The computations to form Markov transaction matrix were described as shown below.

$$h_{11} = s_{11}/(s_{11} + s_{12}) = 0.9423, \quad h_{12} = s_{12}/(s_{12} + s_{22}) = 0.0577$$

$$h_{21} = s_{21}/(s_{11} + s_{21}) = 0.0577, \quad h_{22} = s_{22}/(s_{12} + s_{22}) = 0.9423$$

After that, matrix Q was obtained. Then, for the condition $X = QX$, we let X be a column vector, and solved the equation to obtain the Eigen vector X with the approximation. After substituting these values in Equation (5), the values for user interest group correlations were obtained as follows.

$$\rho(1,1,1) = 0.44, \quad \rho(1,1,2) = 0.03, \quad \rho(1,2,1) = 0.01, \quad \rho(1,2,2) = 0.01,$$

$$\rho(2,2,2) = 0.44, \quad \rho(2,2,1) = 0.03, \quad \rho(2,1,1) = 0.01, \quad \rho(2,1,2) = 0.01$$

The consumers' interests were computed by substituting these values in (6) for weighted links among the images. By doing this, we obtained the similarity matrix. By then normalizing the matrix, we obtained the user-intent image-similarity matrix. From the Markov transition matrix after normalizing, we derived the stationary distribution vector, and used values for these probabilities for ranking images according to user intentions. The result was the Consumers-Intent Webpage's Image Rank (CWI Rank).

3.2. **Experimental results of WIS measure module.** This section demonstrates the experimental results for web image-content similarity measure proposed in this paper. The method used for feature extraction in this implementation was a deep learning framework: ImageNet with Caffe. Our experimental data contained 206,949 images in five categories, including 132,354 food images, 6,620 drink images, 47,959 indoor images, 19,122 outdoor images, and 894 menu images. In the experimental results, 20 images were randomly selected from each class as query images. Figure 2 shows the image retrieval results between the two groups, including 1,540 images. In order to evaluate the effectiveness of the method, mean average precision evaluations were adopted for evaluating image-retrieval performance as in Table 1. We here described the results using 20 images from each of the five classes, and four query images were used from each of these classes. Then, mean average precision is calculated for each class. The data contained 1,540

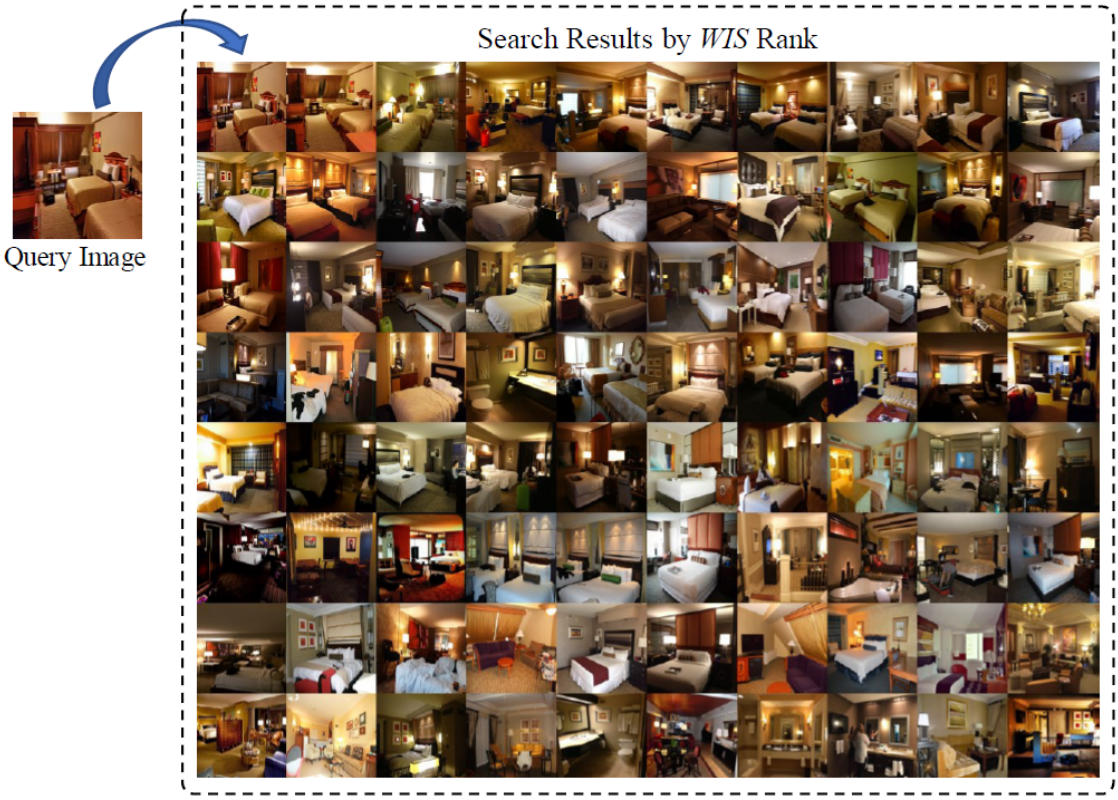

FIGURE 2. Results for web image-content similarity rank between two groups (inside images); Lists of ranking: from left to right and top to bottom

TABLE 1. Mean average precision for web image-content similarity measure

Class of images	Average precision (WIS Rank)
Food	0.8059
Drink	0.6163
Menu	0.3642
Inside	0.7914
Outside	0.8590

images, including 211 food images, 18 drink images, five menu images, 542 indoor, and 769 outdoor images.

3.3. **Experimental results by consumers-intent hybrid rank.** To obtain the consumers-intent hybrid rank, user interest was evaluated using the star ratings given by users. This user-centric concept was demonstrated by investigating the star ratings in pages with respect to the most similar images. Then, the star ratings on those pages were counted in descending order. Table 2 shows how the images were ranked by WIS, CWI and re-ranked by the consumers-intent hybrid rank. Figure 3 provides the comprehensive experimental results for the user interface. When the user uploaded a query and searched an image on the user interface, the top relevant results were retrieved according to combined consumers-intent webpage's image rank and image content similarity rank. The summarized result page indicated that the most relevant images most closely matched the user's query. By using the consumers-intent information ranking system, the user

Table 2. Consumers-intent hybrid rank

WIS Rank	CWI Rank based on star rating	Consumers-intent hybrid rank
Rank 1 Image	2.5 stars	Rank 1 Image (2.5 stars) → Rank 1
Rank 2 Image	4.5 stars	Rank 152 Image (4.5 stars) → Rank 2
Rank 3 Image	3.5 stars	Rank 1510 Image (4.5 stars) → Rank 3
Rank 4 Image	4.0 stars	Rank 1509 Image (4.5 stars) → Rank 4
Rank 5 Image	4.0 stars	Rank 1504 Image (4.5 stars) → Rank 5
Rank 6 Image	3.0 stars	Rank 1501 Image (4.5 stars) → Rank 6
Rank 7 Image	4.0 stars	Rank 1494 Image (4.5 stars) → Rank 7
Rank 8 Image	4.0 stars	Rank 1463 Image (4.5 stars) → Rank 8
Rank 9 Image	3.0 stars	Rank 1456 Image (4.5 stars) → Rank 9
Rank 10 Image	3.0 stars	Rank 1452 Image (4.5 stars) → Rank 10
Rank 11 Image	3.5 stars	Rank 1443 Image (4.5 stars) → Rank 11
Rank 12 Image	3.0 stars	Rank 1430 Image (4.5 stars) → Rank 12
Rank 13 Image	4.0 stars	Rank 1424 Image (4.5 stars) → Rank 13
Rank 14 Image	3.0 stars	Rank 1397 Image (4.0 stars) → Rank 14
Rank 15 Image	4.0 stars	Rank 1397 Image (4.0 stars) → Rank 15
Rank 16 Image	4.5 stars	Rank 1387 Image (4.0 stars) → Rank 16
Rank 17 Image	4.0 stars	Rank 1362 Image (4.0 stars) → Rank 17
Rank 18 Image	3.0 stars	Rank 1361 Image (4.0 stars) → Rank 18
–	–	–
Rank (i) Image	(v = values) stars	Rank (i) Image (v stars) → Re-rank

obtained the desired information from simple images. The system provided the most relevant information link when the user found the contents of an image.

We described the calculation for obtaining the Consumers-Intent Webpage's Image Rank (CWI Rank), as well as the processing of Web Image-content Similarity Rank (WIS Rank). Consumers-intent hybrid rank was calculated by integrating both of these features to optimize the results. We confirmed that the consumers-intent hybrid rank is more effective than only using image-content similarity. The two results are identified as result 1 and result 2. Result 1 only represents the WIS Rank (see Figure 2) and result 2 represents the consumers-intent hybrid rank (see Figure 3). In order to evaluate system performance, a survey was taken of 10 users for top-ranked 10 query (Q) images, who preferred one of the two results for each query. After taking the survey, the average values for results 1 and 2 for each of image query were computed as shown in Table 3. The survey results prove that the consumers-intent hybrid rank provided the results that most closely matched the users' intent. Finally, the average accuracies of results 1 and 2 for each query are obtained as described in Table 3.

4. **Conclusions.** This paper has reported on using a search behavior analyzer to investigate image searching behavior using consumer technologies. This analyzer makes use of stochastic analysis and takes advantage of recent developments in IoT environments. The proposed method mostly exploited visual information for consumers-intent in prioritizing the importance of the various technologies. In doing so, image content analysis and image similarity measures were carried out using Markov Chain and CNN models. The algorithms and implementations were written in MATLAB R2017a, and Apache Tomcat Server 7.1. As one of the main advantages of the approach used, an automatic retrieval

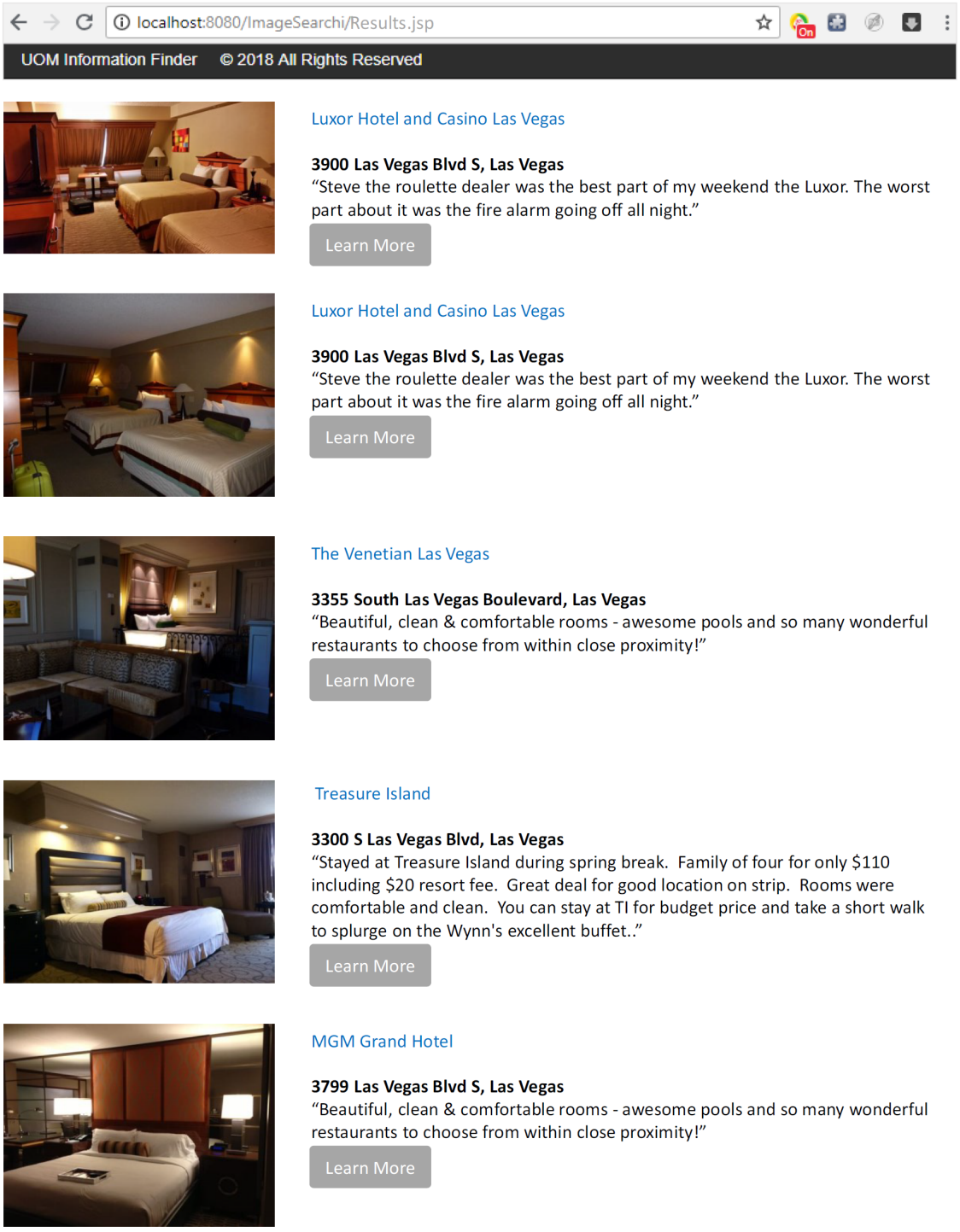

FIGURE 3. Result page for consumers-intent hybrid rank

TABLE 3. Average accuracy of consumers-intent hybrid rank

Result	Q_1	Q_2	Q_3	Q_4	Q_5	Q_6	Q_7	Q_8	Q_9	Q_{10}
1	0.6	0.6	0.3	0.4	0.2	0.3	0.2	0.5	0.6	0.4
2	0.4	0.4	0.7	0.6	0.8	0.7	0.8	0.5	0.4	0.6

process is possible, as opposed to the traditional keyword-based approach, which usually requires a laborious, time-consuming previous annotation of database images. Future

research in developing real-life consumer behavior analyzer will be enhanced by combining more user-oriented social media signals and feature extraction for image contents similarity measures on web database. The proposed method could make an important contribution to further development of consumer technologies in the era of the Internet of Things (IoT).

REFERENCES

[1] P. Anand, M. Sonal, K. Sudarshan and S. Komal, One click intent image search, *Intl. J. of Recent Technology and Engineering (IJRTE)*, vol.3, no.6, 2015.

[2] J. Lee and Y. S. Hong, Improving flight search engine by learning consumer preference function from choice data, *ICIC Express Letters*, vol.12, no.10, pp.1047-1052, 2018.

[3] A. Akusok et al., Arbitrary category classification of websites based on image content, *IEEE Comput. Intell. Magazine*, vol.10, no.2, pp.30-41, 2015.

[4] D. Cai, X. He, Z. Li, W.-Y. Ma and J.-R. Wen, Hierarchical clustering of WWW image search results using visual, textual and link information, *Proc. of ACM Intl. Conf. on Mult.*, pp.952-959, 2004.

[5] P. S. Sharma, D. Yadav and P. Garg, A systematic review on page ranking algorithms, *Int. J. Inf. Tecnol.*, vol.12, pp.329-337, DOI: 10.1007/s41870-020-00439-3, 2020.

[6] S. Liu, P. Cui, H. Luan, W. Zhu, S. Yang and Q. Tian, Social-oriented visual image search, *Comput. Vision and Image Understanding*, vol.118, pp.30-39, DOI: 10.1016/j.cviu.2013.06.011, 2014.

[7] K. Hassani and W.-S. Lee, Visualizing natural language descriptions: A survey, *ACM Comput. Surv.*, vol.49, pp.1-34, DOI: 10.1145/2932710, 2016.

[8] J. Zakraoui, M. Saleh and J. Al Ja'am, Text-to-picture tools, systems, and approaches: A survey, *Multimed. Tools Appl.*, vol.78, pp.22833-22859, DOI: 10.1007/s11042-019-7541-4, 2019.

[9] T. Geng, T. T. Zin, P. Tin, T. Toriu and H. Hama, Simultaneous visual ranking and clustering using weighted multiple features, *ICIC Express Letters*, vol.5, no.10, pp.3773-3778, 2011.

[10] *https://www.yelp.com/dataset*, Accessed in 2018.

[11] P. Tin, T. T. Zin, T. Toriu and H. Hama, A novel hybrid approach to image ranking system, *ICIC Express Letters, Part B: Applications*, vol.6, no.3, pp.743-748, 2015.

[12] S. N. N. Htun, T. T. Zin, M. Yokota and K. M. M. Tun, User-intent visual information ranking system, *Proc. of the 5th IEEE Global Conference on Consumer Electronics*, Kyoto, Japan, pp.1-2, DOI: 10.1109/GCCE.2016.7800316, 2016.

[13] *https://www.vlfeat.org/matconvnet/pretrained/*, Accessed in 2018.

[14] F.-F. Li, ImageNet: Crowdsourcing, benchmarking and other cool things, *CMU VASC Seminar*, vol.16, pp.18-25, 2010.