Visually Interpreting Names as Demographic Attributes by Exploiting Click-Through Data

[†‡]**Yan-Ying Chen,** [†]**Yin-Hsi Kuo,** [†]**Chun-Che Wu,** [†]**Winston H. Hsu**
†National Taiwan University, Taipei, Taiwan
‡FX Palo Alto Laboratory, Inc., California, USA
{yanying,kuonini,kenwu0721}@gmail.com, whsu@ntu.edu.tw

Abstract

Name of an identity is strongly influenced by his/her cultural background such as gender and ethnicity, both vital attributes for user profiling, attribute-based retrieval, etc. Typically, the associations between names and attributes (e.g., people named "Amy" are mostly females) are annotated manually or provided by the census data of governments. We propose to associate a name and its likely demographic attributes by exploiting click-throughs between name queries and images with automatically detected facial attributes. This is the first work attempting to translate an abstract name to demographic attributes in visual-data-driven manner, and it is adaptive to incremental data, more countries and even unseen names (the names out of click-through data) without additional manual labels. In the experiments, the automatic name-attribute associations can help gender inference with competitive accuracy by using manual labeling. It also benefits profiling social media users and keyword-based face image retrieval, especially for contributing 12% relative improvement of accuracy in adapting to unseen names.

Introduction

Human attributes (gender, age, ethnicity, etc.) are vital to semantically characterize a person or demographics of a community. This nature makes it valuable for marketing (Aaker 1997), personalization (Cheng et al. 2011), surveillance system (Vaquero et al. 2009), face image search (Kumar, Belhumeur, and Nayar 2008), social computing (Liu, Zamal, and Ruths 2012) and more human-centric research. The preliminary studies have achieved convincing results in attribute prediction of a person by analyzing his/her tweets (Burger et al. 2012), face images (Cheng et al. 2011) and user names (Mislove et al. 2012); however, in general circumstances people may hide most attribute contexts because of privacy issues.

Since users tend to keep their online profiles private (Dey, Jelveh, and Ross 2012), name is the most reachable piece of personal information among these contexts. The problem we address is – given a name, associating and predicting its likely demographic attributes, particularly, the gender

Figure 1: We aim at associating a human name such as "Amy Liu" with its likely visual attributes by exploiting click-throughs in web image search logs under a novel visual-data-driven manner without intensive labeling efforts and adaptive to unseen names.

and ethnicity attributes discussed in this paper. As shown in Figure 1[1], given a person named "Amy Liu," the person is likely an Asian female. Name makes the first impression of a person because naming conventions are strongly influenced by culture, e.g., first name and gender, last name and location of origin. Typically, the associations between names and the two attributes are available by referring to demographics maintained by governments (Burger et al. 2012; Mislove et al. 2012) and manually labeling attributes based on profile photos of sampled social media users (Liu, Zamal, and Ruths 2013). The former is limited in regional census data and not general to more countries. The latter has major concerns in time and cost when it adapts to large-scale data.

Different from the traditional manners, we propose to associate name and attribute by exploiting click-throughs between text queries and face images in web search logs, where the names are extracted from queries and the attributes are detected from face images automatically. In this paper, a *click-through* means when one of the URLs returned by a text query has been clicked to view a web image it di-

[1]The face images presented in this paper are under Creative Commons Licenses.

rects to (Beeferman and Berger 2000). The mechanism delivers two messages, (1) the association between a query and an image is based on viewers' clicks, that is, human intelligence from web-scale users; (2) users may have considerable knowledge to the associations because they might be partially aware of what they are looking for and search engines are getting much better at satisfying user intent. Both characteristics of click-throughs reduce the concerns of incorrect associations. Meanwhile, the Internet users' knowledge enables discovering name-attribute associations with high generality to more countries and without relying on regional census data that might not be freely accessible.

Because click-throughs may suffer from data sparsity problem (Craswell and Szummer 2007; Cao et al. 2010) and always have new instances out of the existing coverage, both leading to a large proportion of unseen names – the names lacking associations with any images in click-through data. We further propose to predict attributes of unseen names by using search snippets, the informative texts that appear under every search result and summarize what and why the page is relevant to the given query. Note that, the labels for training attribute classifiers of unseen names solely rely on the automatic association by using click-throughs. For evaluations, we only use English names; however, we believe that the method can scale to more communities by incorporating natural language processing for more languages.

In the experiments, we demonstrate how the learned attribute classifiers of names can contribute to keyword-based face image retrieval and cross-domain gender inference. The accuracy is increasing with the incremental knowledge of crowds taken into account. In summary, our contributions include, (1) proposing a novel visual-data-driven approach to associate a human name with its likely demographic attributes by using click-throughs; (2) proposing to predict attributes of unseen names by using search snippets; (3) showing the effectiveness of name-attribute associations in improving keyword-based face image retrieval especially for the queries comprising unseen names; (4) demonstrating the capability of the name-attribute associations in cross-domain gender inference.

The system overview is illustrated in Figure 2. In the rest of the paper, the literature survey is introduced first, followed by the description of click-through data. The proposed name-attribute association and attribute prediction for unseen names are then presented. Finally, the experiments are discussed, closing with a conclusion.

Related Work

Our main idea is motivated by these preliminary studies: (1) a click-through can be viewed as a weak indication of relevance with user judgment (Song, Miao, and Shen 2011; Joachims 2002); (2) certain human attributes are intuitively understandable through visual appearance of a person (Liu, Zamal, and Ruths 2013); (3) semantic attribute is cross-domain and adaptive to unseen (or new) classes (Lampert, Nickisch, and Harmeling 2014). We first review how click-through data are used in previous literature, followed by the challenges in obtaining attribute labels and identifying new classes by semantic attributes.

Click-through data have been leveraged for optimizing information retrieval (Joachims 2002; Cao et al. 2010), interpreting implicit feedback (Joachims 2005; Chen et al. 2011) and query suggestion (Beeferman and Berger 2000). It acts as additional and free contexts beyond content and associates relevant items even across multiple modals such as image and text. Song, Miao, and Shen (2011) propose an efficient multiple instance learning approach to automatically produce high relevance labels for pairs of query and URL. Jain and Varma (2011) propose to improve the performance of keyword-based image retrieval by click-through data. Yao et al. (2013) propose an economic way for video tagging by understanding user searching behavior through clicks. Different from the previous work where click-throughs only contribute document (image) level association, we propose to leverage visual-based face attribute detection to achieve attribute level association with name queries.

Attribute labels are not easily available on the Internet because users would like to protect their privacy and hide the details of their profiles (Dey, Jelveh, and Ross 2012). Some studies have tried to empirically determine attributes of a user by observing the user generated textual content in one form or another (Burger et al. 2012; Pennacchiotti and Popescu 2011). Liu, Zamal, and Ruths (2013) propose to collect gender labels of Twitter users by hiring Turkers to label each user's gender based on his/her profile photo. However, paid crowdsourcing is relatively harder to scale up to big data. Conversely, the public click-throughs are nearly free, and with more clicks the incorrect annotation can be reduced considerably. Meanwhile, visual content, particularly face, has become more important for knowledge extraction and representation with the prevalence of images and videos on the Internet (Hasan and Pal 2014).

Attributes that carry semantics are beneficial for zero-shot learning which targets the classification when no training data for a new class are available. Lampert, Nickisch, and Harmeling (2014) propose to identify unseen objects based on high-level semantic attributes such as the object's color or shape. Chen, Gallagher, and Girod (2013) propose attribute-based approach to predict the most possible name given a face image. In contrast to recognizing attributes or names for given image content, we intend to translate a name (class) to semantic attributes by how the crowd think it should look like.

Data Collection

We need massive annotated data for inferring the associations between names and demographic attributes. Instead of collecting data manually, we leverage click-throughs in web search logs that can be viewed as weak annotations. In the experiments, we utilize a new large-scale real-world click data set publicly released by Microsoft Research and Bing (Hua et al. 2013). They sampled from one-year click logs of the Bing image search engine and formed the dataset – **Clickture-Lite**, which contains 23 million click data with 1 million unique images. Each of the click-through data (θ) consists of $<$image ID, query text, number of clicks$>$ (cf. Figure 2) with 11.7 million unique query texts.

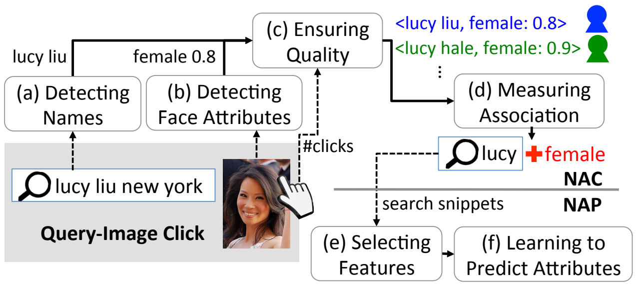

Figure 2: The proposed framework with two major parts – name-attribute association by clicks (NAC) and attribute prediction for unseen names (NAP). Given a query-image click, (a) the name in the query and (b) the face attributes in the image are detected automatically to form an association. (c) The clicks contributed by the Internet users are leveraged to ensure quality of association. (d) Assuming each full name represents one person, the association of a (first/last) name and an attribute is measured by considering the attribute probability of persons with the same first/last name. For unseen names, (e) we collect search snippets as features and exploit the associations obtained by NAC (a to d) as labels for learning attribute detectors (f).

Name-Attribute Association by Clicks (NAC)

Seeing the importance of semantic attributes and their close relationship with a name, we would like to address the problem – given a name n_i, associating its likely gender and ethnicity attributes. We leverage our click-through data θ as observation to measure the name-attribute association $P(a_j|n_i;\theta)$ of each attribute a_j given a name n_i. The name-attribute association $P(a_j|n_i;\theta)$ is a statistical measurement that can be interpreted as a prior probability of attributes for an identity with that name. As shown in Figure 2, four major steps (a) detecting names, (b) detecting face attribute, (c) ensuring quality by clicks and (d) measuring name-attribute associations are proposed to infer the likely attributes of a given name.

Detecting Names: We construct a name dictionary to detect names in queries. First, we collect a list of celebrity names available on the Internet [2], totally 2,221 names denoted as C. Each name is split into words, typically a first name and a last name, resulting in 1,164 first names and 1,681 last names denoted as F and L, respectively. To enrich the name dictionary, we combine each first name in F and each last name in L, which can cover nearly two million name combinations. Given a query, we then extract any names in C or any consecutive words where the former one in F and the latter one in L. The queries comprising any of the two million names in our dictionary are used to infer associations. Without loss of generality, we further address the unseen names out of this coverage in the next section.

Detecting Face Attributes: On the part of image, we leverage a visual-based approach to automatically detect face attributes. First, we localize the faces in an image by face detection (Viola and Jones 2001), totally 495,585 faces detected from 328,838 images. For each of the detected faces, we use the four face attribute classifiers $a_j \in$ {gender, Asian, African, Caucasian} (Kumar, Belhumeur, and Nayar 2008), to predict the probability $P_v(a_j|d_t)$ for a binary outcome of each attribute a_j appearing in a given face d_t. Each attribute classifier is learned by Support Vector Machines (SVMs) using four kinds of low-level features, Local Binary Pattern (LBP), Gabor filter, Histogram of Oriented Gradient and Grid Color Moments extracted from face images. The average accuracy of face attribute detection is around 80% (Kumar, Belhumeur, and Nayar 2008).

Ensuring Quality of Query-Image Pairs: The ambiguous mappings between multiple names in a query and multiple faces in an image may induce incorrect name-attribute associations. To eliminate the ambiguity, we only keep the query-image pair if the query has exactly one name n_i and the image has exactly one face d_t (cf. the example query-image pair in Figure 2), each corresponding to the probabilities of attributes $P_v(a_j|d_t)$ detected from the face d_t. In addition, the number of clicks for a query-image pair can be thought as the agreement among users such that it can be used to confirm the correctness of a name-attribute association. In our experiments, a query-image pair should be clicked at least twice to ensure the quality of association. The converted name-face pair is referred as $< n_i, d_t >$.

Measuring Probability of Association: We measure the probability of a name-attribute association $P(a_j|n_i)$ of a name n_i and an attribute a_j by averaging attribute probability $P_v(a_j|d_t)$ of all images for a name.

$$P(a_j|n_i) = \frac{\sum_{d_t \in d'} P_v(a_j|d_t)}{|d'|}, \tag{1}$$
$$d' = \{d_t | \forall < n_h, d_t >, n_h = n_i\}.$$

A name n_i may consist of a first name $n_i^{(f)}$ and a last name $n_i^{(l)}$; for example, "Michael Jordan" comprise the first name "Michael" and the last name "Jordan." In such case, we separate $P(a_j|n_i)$ to $P(a_j|n_i^{(f)})$ and $P(a_j|n_i^{(l)})$. Like the example in Figure 2, there may exist more than one full names with the same first name, e.g., "Lucy Liu," "Lucy Hale" and "Lucy Gordon" all have the same first name "Lucy." Assuming that a full name represents a person, the association probability $P(a_j|n_i^{(f)})$ of a first name $n_i^{(f)}$ and an attribute a_j can be measured by averaging $P(a_j|n_i)$ of the full names n' comprising the same first name $n_i^{(f)}$.

$$P(a_j|n_i^{(f)}) = \frac{\sum_{n_h \in n'} P(a_j|n_h)}{|n'|}, \tag{2}$$
$$n' = \{n_h | \forall n_h, n_h^{(f)} = n_i^{(f)}\}.$$

The association probability for last name $P(a_j|n_i^{(l)})$ follows the same formulation. The higher $|n'|$ means more reference persons are sampled as the observations, that is, the association probability is more statistically convincing. The number of reference persons is sort of confidence for the association and its impact is discussed in the experiments.

[2]http://www.posh24.com/celebrities/

Attribute Prediction for Unseen Names (NAP)

We further target predicting attributes of the unseen names which lack click-throughs with any images in observation data. Because of limited information in the very short text of a name, we crawl additional contexts of a name to further predict its attributes. The query expansion technique (Sahami and Heilman 2006) is adopted to acquire the snippets of each name returned by the search engine as the features (Figure 2 (e)). The automatic name-attribute associations via click-throughs are exploited as the pseudo labels for learning attribute classifiers of names (Figure 2 (f)). The output is a probability $\hat{P}(a_j|n_i)$ for a binary classifier of each attribute a_j given a name n_i.

Selecting Differential Features: We use the approach extensively used in prior studies (Burger et al. 2012; Liu, Zamal, and Ruths 2012) to select discriminative text features for specified classes. We convert all the snippets to lowercase and split words with space and punctuation. All the split words are then lemmatized to group the different inflected forms of a word as a single item (Bird 2006). The split words are formed as a corpus for extracting unigrams and bigrams. The k most differential unigrams and bigrams found in the snippets of each class are selected as features. For a binary attribute classifier with two classes, the most differential features have higher difference of term frequency in the snippets for one class versus those in the opposite one. Some examples of the selected features are presented in Figure 6. Finally, the selected differential unigrams and bigrams are collected as a vocabulary S for feature extraction.

Learning to Predict Attributes: Based on the vocabulary S, we extract the features for search snippets of each name. The features of a name constitute an indicator vector, each indicating the presence of a term in S. The attribute label of each name in training data is annotated by automatic associations by using NAC. The binary attribute classifier of a name is then learned by SVM (Chang and Lin 2011) with Radial Basis Function as kernel. The cost and gamma parameters are optimized by grid search. Given a new coming name which has not yet been clicked, we can predict its likely attributes by feeding the features extracted from its search snippets to the learned attribute classifier.

Experiments

Evaluation Data Sets and Metrics

We organize the evaluation with four major parts (1) the accuracy of the name-attribute associations, (2) the accuracy of attribute prediction for unseen names, (3) the ranking quality of keyword-based image retrieval and (4) the accuracy of cross-domain gender inference improved by the automatic name-attribute associations. All of the experiments are evaluated in publicly available benchmarks introduced as below.

Demographics from Government (GT): We refer to the 1,000 popular female names and 1,000 popular male names from the real demographics provided by the U.S. Social Security Administration [3]. The associations of names and ethnicity are collected from the U.S. Census [4], totally 1,000 popular last names.

[3]http://www.ssa.gov/OACT/babynames/

Figure 3: The tag cloud of male (left) and female (right) names. The larger size the name has, the higher probability the persons with the name are likely of that gender.

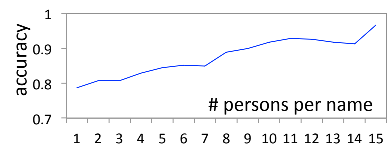

Figure 4: The accuracy of name-gender associations by clicks. The results show that the association accuracy improves saliently if more reference persons (more web search logs contributed by crowds) are taken into account.

Bing Image Search Development Data (Bing-D): The Clickture-Lite provides a development set with relevance labels for evaluating image retrieval. Based on the development set, we construct 4 subsets – CQ-S, CQ, PQ, PQ-S – for detailed evaluations. CQ-S retains the queries comprising at least one full name appearing in the name dictionary. CQ further includes the queries comprising **Partial Names** – at least one first or last name in the dictionary. In addition to CQ, PQ includes **Unseen Names** that are out of the dictionary. The **Ambiguous Names**, e.g., multiple names or names of cartoon and game characters, are excluded from PQ to form PQ-S. The performance is evaluated by discounted cumulative gain at 25 (DCG_{25}) for each query. The setting and detailed numbers are presented in Table 2.

Twitter User Data (TU): We use the canonical gender-labeled Twitter data (Liu, Zamal, and Ruths 2013) to evaluate how much our gender-name associations can contribute to gender inference for Twitter users. Totally 10,788 users and their user names are included for evaluation, totally 6,903 females and 3,885 males. Each gender label is decided by Turkers based on the user's profile photo.

Evaluation of Name-Attribute Association (NAC)

Figure 3 shows the tag clouds of male (left) and female (right) names, where the size of names is decided by the name-male and name-female probabilities $P(a_j|n_i^{(f)})$ (cf. Eq. 2), respectively. The top 3 names having the highest associations with the male attribute are "William," "Thomas" and " Russell" while the top 3 ones for female are "Ann,"

[4]http://www.census.gov/

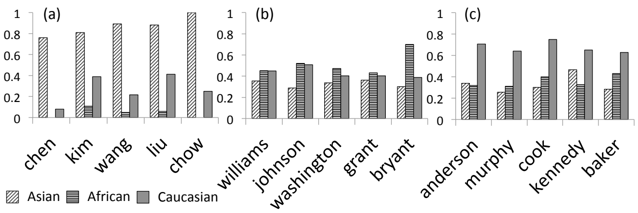

Figure 5: The attribute probabilities given the five sampled popular last names of three different races in the U.S. government data (GT). For the Asian last names (a), the estimated association probability with the Asian attribute is much higher than that with the other ethnicity attributes.

(a)	his, actor, online, game, he, company, team, men, sport	latest news, his stage, stock chart, name after, against other
(b)	her, she, actress, facebook, photo, model, girl, like, fashion, woman	american actress, her father, her work, show off, photo gallery

Figure 6: The example differential terms in the snippets searched by male names (a) and female names (b). The left and right are unigram and bigram features, respectively.

"Danielle," and "Kimberly." The persons who have these names are likely to be males/females. It is intuitive that more reference persons are considered in statistics, more accurate results we might have. We use the (GT) data set as the ground truth and control the minimum number of reference persons for a name (n' in Eq. 2) from 1 to 15. As shown in Figure 4, the names with at least one reference person can reach 79.64% in accuracy and the accuracy is constantly increasing for the names with more reference persons.

We present our results in terms of ethnicity including Asian, African and Caucasian. The race distribution are much more unbalanced because we now only target English language communities. One more problem is visual-based attribute detection for three races are more challenging compared to gender attribute detection. The results among the most race-specific last names still tell the interesting relations between last name and race. We select 5 popular names for Asian, African and Caucasian, respectively, based on the real demographics in (GT) data. Note that, most of them are relatively popular names rather than race-specific names, because most of the last names in the U.S. census data are not that specific to a certain race. Figure 5 reports the probability yielded by three binary classifiers, each corresponding to a race attribute. The probability of association with the Asian attribute (a) is much higher than that with the other race attributes. For the African (b) and Caucasian (c), the differences are not that significant, perhaps because the mix-race groups are more common for these two race communities in our training data.

Evaluation of Attribute Prediction for Unseen Names (NAP)

We present the experiments for predicting attributes for an unseen name, *where its name-attribute association in the training data is not available*. However, because the association between name and race is not that differentiable, it is less effective to train attribute classifiers by the less reliable associations. Therefore, we only address how well the gender attribute of an unseen name can be predicted.

Table 1: Attrbuite prediction for unseen names. The accuracy of NAP can reach 75% by cross-validation (NAP-CV) and 77% by testing the government data (NAP-GT) that is comparative to automatic associations by clicks (NAC-GT).

method	NAP-CV	NAP-GT	NAC-GT
Accuracy	75.76%	77.25%	79.64%

We first select the differential words from the search snippet data crawled by using name as query. The differential terms in snippets searched by male and female names are presented in Figure 6 (a) and (b). Most snippets comprise the short introduction or profiles of celebrities, brands or social media users. The snippets regarding celebrities usually have gender-specific pronouns like "he" or "she" as well as their professions such as "actress." A name might represent a brand, which may not be directly related to human. However, an interesting finding is that these brand names are human-attribute-specific. For example, the name "Laura" is a brand name of woman clothing such that its descriptions comprise the word "woman" much frequently. For the snippets from user names in social media, the content may reveal their interests and favors. Among these interests, "game," "sport" and "fashion" are more differential over gender groups.

To evaluate the accuracy of the proposed prediction approach (NAP), we conduct experiments over our snippet data set labeled by automatic name-gender association. Table 1 shows the accuracy reaches 75.76% in cross-validation manner (NAP-CV). We again use the (GT) data as the ground truth to evaluate how precisely the predicted gender can match the real demographics. The accuracy (NAP-GT) reaches 77.25% which is comparative to name-gender association (NAC-GT).

Image Retrieval by Name Attributes

Translating a name into semantic attributes poses a new opportunity for keyword-based image retrieval. The attributes are helpful to rerank the retrieved images; for example, given a query with only female names, suppressing the images with males can improve the precision in top ranks. We compare our methods with the two previous studies, (1) **IR** (Hua et al. 2013) is the text-based initial ranking provided by (Bing-D); (2) **FR** (Ahonen, Hadid, and Pietikäinen 2004) is the ranking measured by face recognition, totally 6,762 identity classifiers trained by the web images weakly labeled by names in queries. As Table 2 shows, the proposed NAC, NAP and their combination NAC+NAP (using NAP for unseen names; otherwise, using NAC) can consistently improve image retrieval accuracy compared to the baseline IR in all the experiment settings. Although the DCG_{25} of FR is better than NAC and NAP, FR only works for the queries

Table 2: The setting of test sets (Bing-D) and the average DCG_{25}. 'yes' and 'no' indicates whether a test set comprises the Partial, Unseen or Ambiguous Names. 'N/A' means the method cannot deal with the corresponding test set. The results show that the proposed gender inference (NAC+NAP) can deal with general test cases while the baseline (FR) cannot. It also consistantly improves the retrieval accuracy compared to the baseline (IR), even reaching 12% relative gain.

Data Set	CQ-S	CQ	PQ	PQ-S
#queries	62	92	300	153
#images	5,825	7,599	30,835	22,573
Partial Name	no	yes	yes	yes
Unseen Name	no	no	yes	yes
Ambiguous Name	yes	yes	yes	no
IR	0.491	0.481	0.436	0.501
FR	0.511	N/A	N/A	N/A
NAC	0.500	0.499	N/A	N/A
NAP	0.498	0.502	0.467	0.565
NAC+NAP	0.500	0.503	0.468	0.563

with full names in the name dictionary while our method NAC+NAP can cope with more queries (e.g., PQ and PQ-S) even they comprise Partial Names, Unseen Names or Ambiguous Names. For the queries without Ambiguous Names, the relative improvement of NAC+NAP can further reach 12% compared to the baseline IR.

The example search results are shown in Figure 7. Because Eva Amurri Martino participates in a comedy film starring Adam Sandler and Andy Samberg, the initial ranking results contain lots of images related to the film (left of (a)). By considering attributes, the images with female are promoted to top ranks (right of (a)). However, we might not be able to improve human-related queries beyond face such as (c). Nevertheless, the proposed method can provide an essential cue for further improvement and is complementary to the existing methods that consider visual similarity only.

User Profiling by Name Attributes

The name-gender associations by clicks in Bing image search logs also benefit gender inference for Twitter users. Twitter user name is not of strict naming convention like personal name. It could be a simple name or name combination with special symbols and of more free forms. To extract the informative parts, we split a user name by any punctuation or space to get one or more terms and transform them in lower case. Then we match each term with the name set included in the automatic name-gender associations by NAC and use the corresponding probability to represent the likely gender of the user name. If a Twitter user name is matched with multiple names, we assign it with the average attribute probability of the names that it is comprised of.

To demonstrate the accuracy of gender inference, we use the publicly available gender-labels for Twitter users (TU) as the ground truth. We remove the user names without any

Figure 7: Examples for image ranking results. The blue rectangles indicate correct results. (a)(b) Based on the name attribute prediction, we can estimate whether the query is related to female and filter out those irrelevant male images. (c) Our proposed method might fail if the name query intended for information beyond face but it might be enhanced combining with intention classification methods.

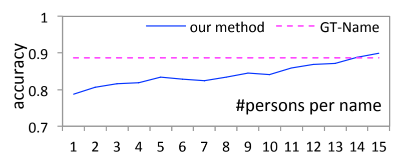

Figure 8: Gender inference for Twitter users. By considering the associations with more reference persons, the accuracy of our method is steadily increasing and comparable with the baseline (GT-Name) that uses data from U.S. census.

terms matched with our name set, totally 5,320 user names remained for evaluation. As shown in Figure 8, our method can reach more than 80% in accuracy. By considering the name-gender associations with more reference persons, the accuracy is steadily increasing and comparative with the accuracy of the baseline (GT-Name) by using name-gender associations from the U.S. government (Liu, Zamal, and Ruths 2013). That again confirms that the proposed name-gender association by click-throughs is convincing to measure prior knowledge for gender inference even in different domains (names in web search logs vs. Twitter user names).

Conclusion and Future Work

To sum up, we propose the first work to interpret a name to demographic attributes in visual-data-driven manner by using publicly available web search logs. The proposed framework is adaptive to incrementally updated data and general to names for more countries and even unseen names lacking click-throughs with any images. Finally, we demonstrate how the automatic associations can improve keyword-based face image retrieval and cross-domain gender inference. With the idea of visual interpretation, we are going to address more names beyond human, e.g., animal, plants and places. For example, the flower species named with "Grevillea" are mostly brightly colored and petal-less (visually perceivable attributes) due to the common genus in biology.

References

Aaker, J. L. 1997. Dimensions of brand personality. In *Journal of Marketing Research*.

Ahonen, T.; Hadid, A.; and Pietikäinen, M. 2004. Face recognition with local binary patterns. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

Beeferman, D., and Berger, A. 2000. Agglomerative clustering of a search engine query log. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*.

Bird, S. 2006. Nltk: the natural language toolkit. In *Proceedings of the COLINGACL on Interactive presentation sessions*.

Burger, J. D.; Henderson, J.; Kim, G.; and Zarrella, G. 2012. Discriminating gender on twitter. In *Proceedings of the Conference on Empirical Methods on Natural Language Processing (EMNLP)*.

Cao, B.; Shen, D.; Wang, K.; and Yang, Q. 2010. Click-through log analysis by collaborative ranking. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*.

Chang, C.-C., and Lin, C.-J. 2011. LIBSVM: A library for support vector machines. *ACM Transactions on Intelligent Systems and Technology*. Software available at http://www.csie.ntu.edu.tw/∼cjlin/libsvm.

Chen, W.; Ji, Z.; Shen, S.; and Yang, Q. 2011. A whole page click model to better interpret search engine click data. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*.

Chen, H.; Gallagher, A. C.; and Girod, B. 2013. What's in a name? first names as facial attributes. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Cheng, A.-J.; Chen, Y.-Y.; Huang, Y.-T.; Hsu, W. H.; and Liao, H.-Y. M. 2011. Personalized travel recommendation by mining people attributes from community-contributed photos. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*.

Craswell, N., and Szummer, M. 2007. Random walks on the click graph. In *Proceedings of the ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR)*.

Dey, R.; Jelveh, Z.; and Ross, K. 2012. Facebook users have become much more private: A large-scale study. In *Proceedings of the IEEE International Workshop on Security and Social Networking (SESOC)*.

Hasan, M. K., and Pal, C. 2014. Experiments on visual information extraction with the faces of wikipedia. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*.

Hua, X.-S.; Yang, L.; Wang, J.; Wang, J.; Ye, M.; Wang, K.; Rui, Y.; and Li, J. 2013. Clickage: Towards bridging semantic and intent gaps via mining click logs of search engines. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*.

Jain, V., and Varma, M. 2011. Learning to re-rank: Query-dependent image re-ranking using click data. In *Proceedings of the International World Wide Web Conference (WWW)*.

Joachims, T. 2002. Optimizing search engines using click-through data. In *Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining (SIGKDD)*.

Joachims, T. 2005. Accurately interpreting clickthrough data as implicit feedback. In *Proceedings of the ACM SIGIR Conference on Research & Development on Information Retrieval (SIGIR)*.

Kumar, N.; Belhumeur, P.; and Nayar, S. 2008. Facetracer: A search engine for large collections of images with faces. In *Proceedings of the European Conference on Computer Vision (ECCV)*.

Lampert, C. H.; Nickisch, H.; and Harmeling, S. 2014. Attribute-based classification for zero-shot visual object categorization. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

Liu, W.; Zamal, F. A.; and Ruths, D. 2012. Using social media to infer gender composition of commuter populations. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM) Workshop on When the City Meets the Citizen*.

Liu, W.; Zamal, F. A.; and Ruths, D. 2013. What's in a name? using first names as features for gender inference in twitter. In *Proceedings of the AAAI Spring Symposium*.

Mislove, A.; Lehmann, S.; Ahn, Y.-Y.; Onnela, J.-P.; and Rosenquist, J. N. 2012. Understanding the demographics of twitter users. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Pennacchiotti, M., and Popescu, A.-M. 2011. A machine learning approach to twitter user classification. In *Proceedings of the International AAAI Conference on Weblogs and Social Media (ICWSM)*.

Sahami, M., and Heilman, T. D. 2006. A web-based kernel function for measuring the similarity of short text snippets. In *Proceedings of the International World Wide Web Conference (WWW)*.

Song, H.; Miao, C.; and Shen, Z. 2011. Generating true relevance labels in chinese search engine using clickthrough data. In *Proceedings of the AAAI Conference on Artificial Intelligence (AAAI)*.

Vaquero, D. A.; Feris, R. S.; Tran, D.; Brown, L.; Hampapur, A.; and Turk, M. 2009. Attribute-based people search in surveillance environments. In *Proceedings of the IEEE Winter Conference on Applications of Computer Vision (WACV)*.

Viola, P., and Jones, M. 2001. Rapid object detection using a boosted cascade of simple features. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

Yao, T.; Mei, T.; Ngo, C.-W.; and Li, S. 2013. Annotation for free: Video tagging by mining user search behavior. In *Proceedings of the ACM International Conference on Multimedia (ACM MM)*.