

Social Affinity Filtering

Recommendation through Fine-grained Analysis of User Interactions and Activities

Suvash Sedhain
ANU & NICTA
Canberra, Australia
ssedhain@nicta.com.au

Riley Kidd
ANU
Canberra, Australia
rileyjkidd@gmail.com

Scott Sanner
NICTA & ANU
Canberra, Australia
ssanner@nicta.com.au

Khoi-Nguyen Tran
ANU
Canberra, Australia
kndtran@cs.anu.edu.au

Lexing Xie
ANU & NICTA
Canberra, Australia
lexing.xie@anu.edu.au

Peter Christen
ANU
Canberra, Australia
peter.christen@anu.edu.au

ABSTRACT

Content recommendation in social networks poses the complex problem of learning user preferences from a rich and complex set of interactions (e.g., likes, comments and tags for posts, photos and videos) and activities (e.g., favourites, group memberships, interests). While many social collaborative filtering approaches learn from aggregate statistics over this social information, we show that only a small subset of user interactions and activities are actually useful for social recommendation, hence learning *which* of these are most informative is of critical importance. To this end, we define a novel social collaborative filtering approach termed social affinity filtering (SAF). On a preference dataset of Facebook users and their interactions with 37,000+ friends collected over a four month period, SAF learns which fine-grained interactions and activities are informative and outperforms state-of-the-art (social) collaborative filtering methods by over 6% in prediction accuracy; SAF also exhibits strong cold-start performance. In addition, we analyse various aspects of fine-grained social features and show (among many insights) that interactions on video content are more informative than other modalities (e.g., photos), the most informative activity groups tend to have small memberships, and features corresponding to “long-tailed” content (e.g., music and books) can be much more predictive than those with fewer choices (e.g., interests and sports). In summary, this work demonstrates the substantial predictive power of fine-grained social features and the novel method of SAF to leverage them for state-of-the-art social recommendation.

Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Information Filtering

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.
COSN '13 BOSTON, USA
Copyright 2013 ACM 978-1-4503-2084-9/13/10 ...\$15.00
<http://dx.doi.org/10.1145/2512938.2512947>

Keywords

social networks; collaborative filtering; recommender systems

1. INTRODUCTION

Online social networks such as Facebook record a rich set of user preferences (likes of links, posts, photos, videos), user traits, interactions and activities (conversation streams, tagging, group memberships, interests, personal history, and demographic data). This presents myriad new dimensions to the recommendation problem by making available a rich labeled graph structure of social interactions and content from which user preferences can be learned and new recommendations can be made.

Most existing recommendation methods for social networks aggregate this rich social information into a simple measure of user-to-user interaction [9, 19, 22, 23, 32, 20, 21]. But in aggregating all of these interactions and common activities into a *single* strength of interaction, we ask whether important preference information has been discarded? Indeed, the point of departure for this work is the hypothesis that different fine-grained interactions (e.g. commenting on a wall or getting tagged in a video) and activities (e.g., being a member of a university alumni group or a fan of a TV series) *do* represent different preferential *affinities* between users, and moreover that effective *filtering* of this information (i.e., learning which of these myriad fine-grained interactions and activities are informative) will lead to improved accuracy in social recommendation.

To quantitatively validate our hypotheses and evaluate the informativeness of different fine-grained features for social recommendation, we have built a Facebook App to collect detailed user interaction and activity history available through the Facebook Graph API along with user preferences solicited by the App on a daily basis. Given this data, (1) we define a novel recommendation method called *social affinity filtering (SAF)* that learns to predict whether a user (*ego*) will like an item based on the surrogate item preferences of others (*alters*) who share fine-grained interactions or activities with the *ego*, and (2) we analyse the relative informativeness of these fine-grained interaction and activity features across a variety of dimensions.

In the four months that our App was active, we collected data for a set of Facebook app users and their full interactions with 37,000+ friends along with 22 distinct types of interaction and users activity for 3000+ groups, 4000+ favourites, and 10,000+ pages. In subse-

quent sections that outline our experimental methodology and results in detail, we make the following critical observations:

- **Overall performance:** We found that SAF significantly outperforms numerous state-of-the-art collaborative filtering and social recommender systems by over 6% in accuracy using just *page* (like) features.
- **Privacy vs. performance:** Because the reluctance of a user to install an App increases with the number of permissions requested, the above results suggest that an SAF-based social recommendation App need only request permissions for a user’s *page* likes in order to achieve state-of-the-art recommendation accuracy.
- **Big data scalability:** We implement SAF as a simple linear classifier that can be globally optimised with a variety of classification methods (e.g., naive Bayes, logistic regression, SVM) and online training algorithms amenable to real-time, big data settings.
- **Cold-start capable:** Since SAF trains a single model for all users and does not require a user’s preferences in order to recommend for them, we show that SAF exhibits strong *cold-start* performance for users *without* expressed item preferences as long as those users have interactions or shared activities with users who have expressed item preferences.
- **Interaction analysis:** Among *interactions*, we found that those on videos are more predictive than those on other content types (photos, post, link), and that outgoing interactions (performed by the ego on the alter’s timeline) are more predictive than incoming ones (performed by alters on the ego’s timeline), although the level of exposure of an ego to an alter’s preferences is often more important than the directionality, modality, or action underlying the interaction with the alters.
- **Activity analysis:** The most predictive activity SAGs tend to have small memberships indicating that these informative activities represent highly specialised interests. We also found features corresponding to “long-tailed” dynamic content (such as music and books) can be more predictive than those with fewer choices that add little new content over time (e.g. interests or sports).
- **Importance of social data beyond friends:** We found that *groups*, *pages*, and *favourites* make for more informative SAGs than those defined by user-to-user interactions. This is likely because the former can be applied to SAGs *over the entire Facebook population* rather than just a user’s friends (where the available preference data is considerably more sparse).
- **Social activity and item popularity vs. performance:** We analyse *how many* shared activities are needed for good performance and observe that increased activity membership correlates with increased recommendation accuracy. However, excessive item popularity among activities hurts the discriminative power of SAF to make good recommendations.
- **Fine-grained vs. aggregate social data:** Among activity features, a small subset proved to be much more informative than the rest. This suggests the value of learning *which* fine-grained features are predictive and sheds doubt on the efficacy of existing social recommendation methods that aggregate social information between two users into a single numerical value.

Subsequent sections demonstrate these findings in detail.



Figure 1: Overview of *social affinity filtering* (SAF): A *social affinity group* (SAG) of user *u* (ego) consists of a set of alternate users $\{v\}$ (alters) who have a certain *interaction* or share an *activity* membership with *u*. SAF learns to classify whether user *u* will like item *i* based on the observed preferences of members of each SAG of user *u* toward item *i*.

2. SOCIAL AFFINITY FILTERING

As illustrated in Fig 1, the high-level objective of this work is to predict whether or not a user *u* (ego) will like an item *i*. Specifically, the Facebook App we have built for our experimentation collects explicit like and dislike feedback for links posted on Facebook (e.g., Youtube video, news or blog item, etc.) leading to the following preference data:

$$likes(u, i) := \begin{cases} true & u \text{ clicked } like \text{ for } i \\ false & u \text{ clicked } dislike \text{ for } i \\ unknown & u\text{'s preference for } i \text{ is unobserved} \end{cases}$$

From the observed data, *social affinity filtering* (SAF) learns to predict $likes(u, i)$ based on the surrogate link preferences $likes(v, i)$ of sets of other Facebook users *v* who have at least one interaction or activity in common with *u*. The details of SAF are outlined in the following subsections.

2.1 Interactions and Activities on Facebook

In the context of Facebook, we use the term *interactions* and *activities* to refer to the range of user-user and user-community actions, respectively.

Interactions describes communication between Facebook users and can be broken down into the following dimensions:

- **Modality:** (4 possibilities) User *u* can interact with another user *v* via *links*, *posts*, *photos* and *videos* that appear in either user’s timeline.
- **Action type:** (3 possibilities) A user *u* can *comment* or *like* user *v*’s item. He/she can also *tag* user *v* on an item, often indicating that user *v* is present when the content is created (for photo/video/post), or to explicitly raise user *v*’s attention for a post — with one exception in Facebook that *u* cannot tag a link with users.
- **Directionality:** (2 possibilities) We look at *incoming* and *outgoing* interactions, i.e., if user *u* comments on, tags, or likes user *v*’s item, then this is an *outgoing* interaction for

u , and an *incoming* interaction for v . Although high correlation between *incoming* and *outgoing* interactions has been observed [26], whether interaction direction affects user preferences differently is still an open question we wish to answer in this work.

Overall there are 22 possible interaction types, namely the cross-product of modalities, actions and directions, minus the special cases of *link-tag-{incoming, outgoing}* since links cannot be tagged.

Activities are user interactions with Facebook communities like groups, pages, and favourites defined as follows:

- **Groups** on Facebook¹ are analogous to real-world community organisations. They allow users to declare membership and support people to organise activities, to post related content, and to have recurring discussions about them. Examples of groups include *Stanford Thai* (Fig 1 bottom left), or *Harvard Debate Club*.
- **Pages** on Facebook² are analogous to the homepages of people, organisations and events on the world-wide-web. They are publicly visible, and users can subscribe to the updates on the page, and also engage in discussions. Example pages include *DARPA* (an organisation, Fig 1 bottom middle), or *Beyonce* (a singer).
- **Favourites** are analogous to bookmarks (on physical books or on the web browser). They are a user-created list containing various items such as Facebook apps, books, music, and many other types of items (even pages) to indicate their interest. Example favourites include *Big Bang Theory* (TV series), or *FC Barcelona* (soccer club). Fig 1 bottom right shows a Facebook screenshot when a user adds a favourite.³

Our evaluation includes 3000+ *group*, 4000+ *page* and 10000+ *favourite* features as detailed in Sec 3.1.

2.2 Social Affinity Groups (SAGs)

With *interactions* and *activities* now defined, we proceed to define two types of *social affinity groups* (SAGs) of a user u that will be used as proxies for u 's preferences:

- **Interaction Social Affinity Groups (ISAGs)**: Let the set of ISAGs be the cross-product of interaction modality, action,

¹From Facebook Blog: <http://www.facebook.com/blog/blog.php?post=324706977130>, "Groups are the place for small group communication and for people to share their common interests and express their opinion. Groups allow people to come together around a common cause, issue or activity to organise, express objectives, discuss issues, post photos and share related content."

²From Facebook Blog: (<http://www.facebook.com/blog/blog.php?post=324706977130>) "Facebook Pages enable public figures, businesses, organisations and other entities to create an authentic and public presence on Facebook. Facebook Pages are visible to everyone on the Internet by default. Facebook users can connect with these Pages by becoming a fan and then receive their updates and interact with them."

³According to Facebook Blog, (<https://www.facebook.com/help/232262810142682>) "Facebook facilitates a wide variety of user selected favourites (Activities, Favorite Athletes, Books, Interests, Movies, Music, Sports, Favorite Teams, Television). These favourites allow a user to associate themselves with other people who share their same favourite tendencies."

and direction:

$$\begin{aligned} \text{Interaction-Classes} &:= \{link, post, photo, video\} \\ &\quad \times \{likes, tag, comment\} \\ &\quad \times \{incoming, outgoing\} \end{aligned}$$

Then for $k \in \text{Interaction-Classes}$ we define

$$ISAG(u, k) := \{v \mid \text{user } v \text{ has had interaction } k \text{ with } u\}$$

For example,

- $ISAG(u, link-like-incoming)$ is the set of all users who have liked a link posted by user u , and
- $ISAG(u, photo-comment-outgoing)$ is the set of all users whose photos user u has commented on.

- **Activity Social Affinity Groups (ASAGs)**: We define ASAGs based on group membership, page likes and user favourites (of which there are over 17000 distinct activities in our data set). For any one of these activities $k \in \text{Activity-Groups}$ we define:

$$ASAG(k) := \{v \mid \text{user } v \text{ has taken part in activity } k\}$$

For example,

- $ASAG(page-Beyonce)$ is the set of all users who have liked *Beyonce's* Facebook page, and
- $ASAG(group-Harvard Debate Club)$ is the set of all users who have joined the Facebook group for the *Harvard Debate Club*.

2.3 Social Affinity Filtering (SAF)

With SAGs now defined, we can use them to build features for a classification-based approach to social recommendation that we term *social affinity filtering* (SAF). In SAF, our goal is to predict $likes(u, i)$ for user u and item i . As features $X_k^{u,i}$ for this classification task, we can use the observed preferences of members of each SAG k as proxies for $likes(u, i)$. Formally, we define such features as follows:

- **Interaction Social Affinity Features (ISAFs)**: We define feature $X_k^{u,i} \in \{true, false\}$ for user u , item i and interaction $k \in \text{Interaction-Classes}$ as

$$X_k^{u,i} := \begin{cases} true & \exists v \in ISAG(u, k) \wedge likes(v, i) = true \\ false & \text{otherwise} \end{cases}$$

In short, $X_k^{u,i}$ is *true* if any user sharing interaction k with u liked i . Here, v is implicitly limited to u 's Facebook friends (with whom u may interact).

- **Activity Social Affinity Features (ASAFs)**: We define feature $X_k^{u,i} \in \{true, false\}$ for user u , item i and activity $k \in \text{Activity-Groups}$ as

$$X_k^{u,i} := \begin{cases} true & u \in ASAG(k) \wedge \\ & \exists v \in ASAG(k) \wedge likes(v, i) = true \\ false & \text{otherwise} \end{cases}$$

In short, $X_k^{u,i}$ is *true* if both u and some other v are a member of activity k and v has liked i . Here, v may range over all Facebook users, i.e., v need not be a friend of u to share the same public activity k .

| | | Social Affinity Features | | | | | | | | | | | | |
|------------------|--|--------------------------|--------------------|--------------------|--------------|-----------------|----------|----------|-----------|-----------------|--|--|---------------|--|
| | | ISAF | | | | ASAF (Page) | | | | | | | | |
| | | link-like-incoming | link-like-outgoing | video-tag-outgoing | Friend liked | Big Bang Theory | Facebook | Shakira | ANU CECs | I Love Canberra | | | User/Dislike? | |
| User-Item | | X_1 | X_2 | X_{22} | X_{23} | X_1 | X_2 | X_2 | X_{k-1} | X_k | | | | |
| u_1, i_5 | | 1 | 0 | 0 | 1 | 0 | 0 | 1 | 0 | 0 | | | 1 | |
| u_1, i_8 | | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 0 | 0 | | | 0 | |
| u_2, i_5 | | 1 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 1 | | | 1 | |
| u_2, i_7 | | 0 | 1 | 1 | 0 | 1 | 0 | 0 | 1 | 0 | | | 1 | |
| \vdots | | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | \vdots | | | \vdots | |
| u_{90}, i_{65} | | 0 | 0 | 1 | 0 | 0 | 0 | 1 | 0 | 0 | | | 0 | |

Figure 2: SAF training data example: each row corresponds to a training data sample for a specific user-item pair (u, i) for which the prediction target $likes(u, i)$ is observed (last column). All other columns represent the value of ISAF or ASAF features evaluated relative to the (u, i) label of each row. All columns are binary-valued (0 = false, 1 = true).

While other non-binary definitions of ISAFs and ASAFs are certainly possible (e.g., the count or fraction of members in a SAG who like the item), simple binary features provided the best performance in our experimental evaluation.

Combining these ISAFs and ASAFs into feature vector $\mathbf{X}(u, i) = \langle \dots, X_k^{u,i}, \dots \rangle$ for $k \in \text{Interaction-Classes} \cup \text{Activity-Groups}$ (or any subset thereof), a SAF classifier is then simply a function

$$f : \mathbf{X}(u, i) \rightarrow likes(u, i),$$

where we restrict $likes(u, i) \in \{true, false\}$.⁴ Given a dataset of historical observations $D = \{\mathbf{X}(u, i) \rightarrow likes(u, i)\}$, we can train f using any existing classification method; in this work we consider linear classifiers trained by an SVM, logistic regression, or naïve Bayes. For prediction, given user u and item i , we build the feature vector $\mathbf{X}(u, i)$ and predict $likes(u, i) = f(\mathbf{X}(u, i))$ using the trained classifier f .⁵

To understand how SAF works, it helps to visualise the training data as shown in Fig 2.

2.4 SAF vs. Other Filtering Methods

While a classification approach to recommendation might evoke comparisons to *content-based filtering* (CBF) [18], we remark that CBF is not a *social* recommendation approach and unlike CBF, SAF does not require explicit user features (e.g., age, gender, location, etc.) or item descriptors (link text, link genre, etc.); in contrast, SAF uses interaction and/or activity data for social network users to define SAGs and learns the affinities between a user (ego) and the different set of alters as defined by these SAGs. Additionally, unlike state-of-the-art *social collaborative filtering* approaches [9, 19, 20, 21, 22, 23, 32], SAF does not aggregate user-user interaction and shared activity data into a single aggregate statistic, instead it uses fine-grained distinctions in this social data to define a large

⁴For training purposes, we omit any unobserved cases for which the class label $likes(u, i) = unknown$. At prediction time, the binary classifier must always select a class label of *true* or *false*.

⁵Since almost all classification methods provide a score (or probability) of a classification, we can also generate the top- n item recommendations for a user u by sorting items by their score.

| | App Users | Ego network of App Users |
|--------|-----------|--------------------------|
| Total | 119 | 37,872 |
| Male | 85 | 20,840 |
| Female | 34 | 17,032 |

Table 1: App user demographics. The *ego network* is the friend network of the App users.

| App Users | Tags | Comments | Likes |
|--------------------------|-----------|-----------|-----------|
| Post | 7,711 | 22,388 | 15,999 |
| Link | — | 7,483 | 6,566 |
| Photo | 28,341 | 10,976 | 8,612 |
| Video | 2,525 | 1,970 | 843 |
| Ego network of App Users | Tags | Comments | Likes |
| Post | 1,215,382 | 3,122,019 | 1,887,497 |
| Link | — | 891,986 | 995,214 |
| Photo | 9,620,708 | 3,431,321 | 2,469,859 |
| Video | 904,604 | 486,677 | 332,619 |

Table 2: Statistics on user *interactions*.

| | App Users | Ego Network of App Users |
|------------|-----------|--------------------------|
| Groups | 3,469 | 373,608 |
| Page Likes | 10,771 | 825,452 |
| Favourites | 4,284 | 892,820 |

Table 3: Statistics on user *actions*, counted for *Groups*, *Pages* and *Favourites* over the App users and their ego network.

number of SAGs and learns which of these SAGs are informative for recommendation.

3. EVALUATION

3.1 Data Description

We built a Facebook App to collect explicit user like and dislike preferences back for links posted on Facebook (e.g., Youtube video, news or blog item, etc.) as well as detailed user interaction and activity history available through the Facebook Graph API. The data collection was performed with full permission from the user and in accordance with an approved Ethics Protocol #2011/142 from the Australian National University.

Our App requested to collect information on profiles (including activity memberships) and timelines (interactions) for the App users *and* their friends as required by Sec 2.2 and Sec 2.3. With such expressive permissions, many potential users were hesitant to install the App — after an intensive one month user drive at our University, we were able to attract 119 App users allowing us to collect activity and interaction data for a combined 37,872 users.⁶

We summarise basic statistics of the data in Tables 1–4. Table 1 presents user and friend demographics. Table 2 summarises the number of records for each item modality (row) and action (col-

⁶The issue of low App user uptake with such expressive App permissions underscores the importance of identifying the *minimal* set of permissions to obtain good recommendation performance — a question we address in our subsequent analysis.

| | Friend recommendation | Non-Friend recommendation |
|---------|-----------------------|---------------------------|
| Like | 1392 | 1127 |
| Dislike | 895 | 2111 |

Table 4: Dataset breakdown of prediction target $like(u, i)$ by the source of the link (Friend/Non-friend) and rating (Like= $true$, Dislike= $false$).

umn) combination. Table 3 shows the group membership, page like and favourite counts for users.

Our App recommends three links to App users each day, which the users may optionally like or dislike. Recommended links are harvested from *both* friends’ and non-friends’ timelines. We display only three links per day in order to avoid rank-bias with preferences; each link could be independently rated. Table 4 shows App user link preference statistics.

All subsequent experiments use offline batch data stored and analysed *after* a four month data collection period.

3.2 SAF Comparison

In this section, we compare novel SAF-based methods with a variety of (social) collaborative filtering baselines:

1. **Most Likely Class Constant Predictor (Const)**
2. **Nearest Neighbor (NN)** [6]
3. **Matrix Factorization (MF)** [27]
4. **Social Matchbox (SMB)** [23]

Here, Const serves as a lower bound on performance, NN and MF are two well-known state-of-the-art *non-social* collaborative filtering algorithms, and SMB is a state-of-the-art *social* collaborative filtering algorithm employing matrix factorization with social regularisation.

Among the novel SAF methods, we analyse four different sets of social affinity features:

1. **Interaction Social Affinity Features (ISAF)**
2. **Activity-based Social Affinity Features (ASAF)** for
 - (a) **Group Memberships**
 - (b) **Page Likes**
 - (c) **Favourites**

Furthermore, for these four classes of features, we train one of three classifier types, leading to the following classes of SAF recommenders evaluated in our experiments:

1. **Naïve Bayes (NB-ISAF, NB-ASAF)**
2. **Support Vector Machines (SVM-ISAF, SVM-ASAF)**
3. **Logistic Regression (LR-ISAF, LR-ASAF)**

NB uses a standard Naïve Bayes implementation, SVM and LR are both implemented using *LIBLINEAR* [10].

In all experiments, we report average classification accuracy (fraction of correct classifications in held-out test data) using 10-fold cross validation and provide standard error bars corresponding to 95% confidence intervals.

Fig 3 compares the above baselines and SAF algorithms. In all of these experiments, SAF variants performed statistically significantly better than the best baseline (SMB), except for NB-ASAF

which we conjecture is due to violation of feature independence assumptions that become more pronounced as the number of features increases (n.b., NB-ISAF uses 22 features while NB-ASAF uses 1000’s of features).

In terms of the best recommenders, we observe that LR-ASAF and SVM-ASAF perform comparably to each other and learn quite well despite the large size of this ASAF feature set. Overall, *LR-ASAF performs 6% better than the best baseline for page likes*. We combined all four features sets in a fifth experiment (not shown) and remark that none of NB, LR, or SVM with all features outperformed LR-ASAF with just page likes. We also note that all ISAF variants statistically significantly outperform all (social) collaborative filtering baselines. Hence, w.r.t. this Facebook dataset, we conclude that (a) *SAF with any available feature set* is sufficient to outperform existing (social) collaborative filtering baselines and that (b) if one wanted to minimise the permissions an App requests then it seems *SAF with page likes* alone is sufficient to outperform all other feature sets (alone or combined).

It is important to consider why ASAFs outperform ISAFs. We conjecture the reasons for this are quite simple: ISAFs can only see the *friends* of user u whereas ASAFs are able to look at all users, independent of u ’s friends. Hence, given the relative sparsity of friend-only data in Facebook compared to the greater Facebook population (at least the subset of the population the App collected) and also the relative number of ISAFs compared to ASAFs, ASAFs appear to draw on a much larger set of SAGs that in turn draw on a much larger user population. Among ASAFs, page likes are the most predictive followed by group membership and favourites. This reinforces our conjecture that data sparsity can hurt SAF since we note from Table 3 that page likes are more prevalent than groups and favourites.

Comparing SAF to the state-of-the-art in social collaborative filtering as represented by Social Matchbox (SMB) [23], we observe that SAF consistently outperforms it. We note that the key difference of SAF vs. SMB is that SAF exploits the predictiveness of fine-grained interactions and activities, whereas most social collaborative filtering approaches [9, 19, 20, 21, 22, 23, 32] instead collapse the diverse set of interactions into a single aggregate statistic for each pair of users. The performance of SAF-based recommenders suggests that the aggregation of all social information into aggregate statistics (without learning which interactions or activities are most informative) may not distinguish informative parts of the social signal from the noise.

On the computational side, we remark that SAF is implemented as a simple linear classifier that can be used in conjunction with a variety of classification methods (e.g., naïve Bayes, logistic regression, SVM) and online training algorithms amenable to real-time, big data settings. Furthermore, the linear classification methods used in SAF admit global convex optimisation w.r.t. a variety of training objectives (e.g., log loss in logistic regression, or hinge loss in SVMs) unlike (social) collaborative filtering approaches based on matrix factorization that use non-convex objectives and lack training optimality guarantees.

3.3 Cold-start Evaluation

Many collaborative filtering algorithms (e.g., NN and MF) suffer from the user *cold-start* problem, i.e., when no preference data is available for a user, these algorithms cannot perform better than the constant (most likely class) predictor since they have no way of generalising to a new unseen user. Since SAF trains a single model for all users and does *not* require a user’s preferences in order to recommend for them, SAF can be used in a *cold-start* setting to recommend for users *without* expressed item preferences as long

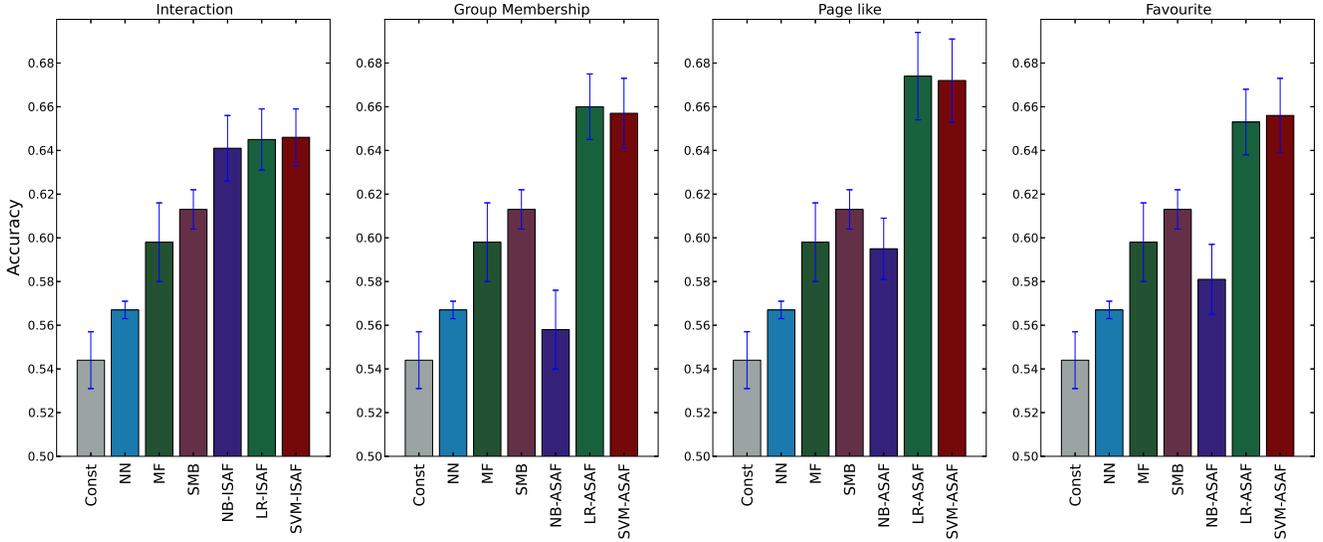


Figure 3: Comparison of a simple baseline (Const), two collaborative filtering baselines (NN and MF), a social collaborative filtering baseline (SMB) and novel SAF recommenders using different feature sets (one ISAF and three ASAF sets) and classifiers (NB,LR,SVM). The best SAF-based model (LR-ASAF) — for Page likes — significantly outperforms all baselines by at least 6%. Combining all four feature sets (not shown) does not lead to improvement over Page likes features alone.

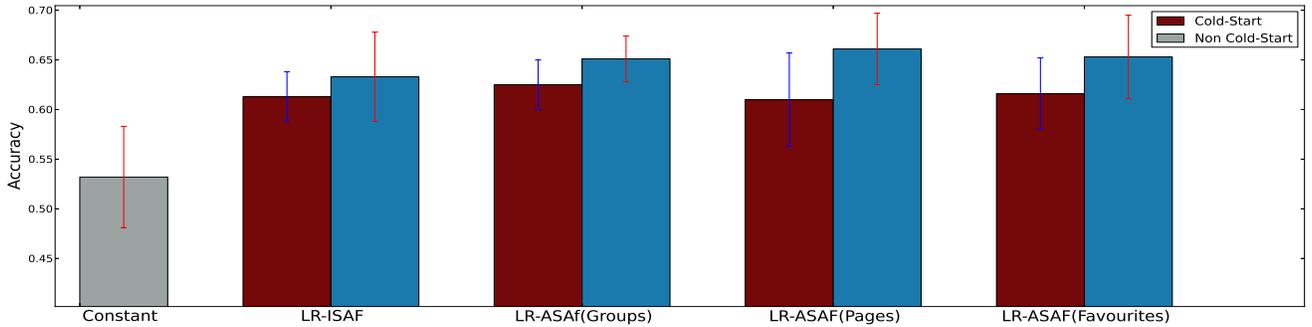


Figure 4: Cold-start evaluation of SAF: accuracy evaluated on cold-start users outperforms the most likely class (Constant) predictor baseline and is somewhat comparable to the non cold-start case when all test user data is *not* withheld from training.

as those users have interactions or shared activities with users who have expressed item preferences.

To quantitatively evaluate the cold-start performance of SAF, we run 10-fold cross validation with specially constructed folds. For *cold-start* evaluation, in each fold, we hold out a random 10% subset of the users for testing and train on the remaining 90% of users. In the test set, we further hold out 30% of the test user data. In the *non cold-start* evaluation, we test on the same data as in the *cold-start* evaluation, but we add in the 30% of held-out test user data to the *cold-start* training set thus allowing the *non cold-start* setting to train on some of the test user data. In Fig 4, we clearly see that the accuracy⁷ of the SAF predictor for cold-start is significantly better than the baseline Constant predictor. Furthermore, the accuracy of the cold-start predictor is actually comparable to the non

⁷The slight decrease in accuracy for non cold-start case compared to Fig 3 is due to the decreased amount of test user data present in the training set for this set of experiments.

cold-start predictor, indicating that SAF exhibits strong cold-start performance.

3.4 Interaction Analysis

In this section we analyse the informativeness of Interaction Social Affinity Features (ISAFs), namely user interactions according to their modality, type, and direction, as described in Sec 2.

A general method for measuring the amount of information that a feature $X_k^{u,i}$ provides w.r.t. predicting a user preference $likes(u,i)$ (in this case, just *true* or *false*) is to calculate its conditional entropy:

$$\begin{aligned}
 & H(likes(u,i)|X_k^{u,i} = true) \\
 & = - \sum_{y \in \{true, false\}} p(likes(u,i) = y|X_k^{u,i} = true) \\
 & \quad \cdot \ln(p(likes(u,i) = y|X_k^{u,i} = true))
 \end{aligned}$$

Lower conditional entropies generally indicate more informative features. We use conditional entropy $H(likes(u,i)|X_k^{u,i} = true)$

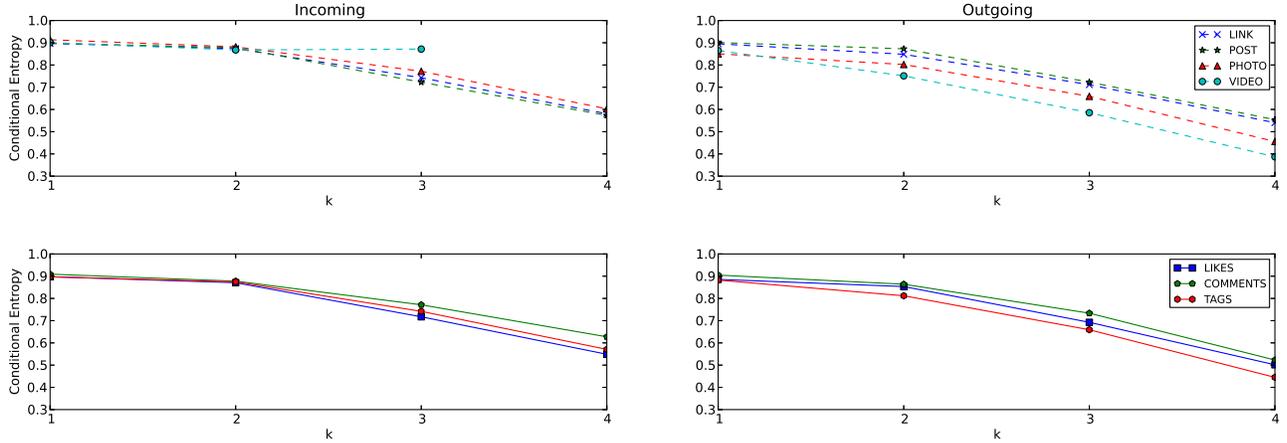


Figure 5: Conditional Entropy of modalities/activities for incoming/outgoing interactions vs. item liked by at least k friends. Increasing k generally has a stronger influence on informativeness than other features of interaction SAGs.

Table 5: Conditional entropy of various interactions (lower conditional entropies are more informative). We observe that interactions on videos are more informative than other modalities (link, post, photo), tagging is marginally more informative than commenting and liking, and outgoing interactions are slightly more informative than incoming ones. Breaking down the analysis by modality-direction and action-direction reveals finer-grained distinctions.

| Modality (X) | $H(Y X = true)$ |
|------------------|-----------------|
| video | 0.850 |
| link | 0.915 |
| post | 0.918 |
| photo | 0.926 |

| Action Type (X) | $H(Y X = true)$ |
|---------------------|-----------------|
| tags | 0.920 |
| comments | 0.921 |
| likes | 0.924 |

| Direction (X) | $H(Y X = true)$ |
|-------------------|-----------------|
| outgoing | 0.928 |
| incoming | 0.935 |

| Modality-Direction (X) | $H(Y X = true)$ |
|----------------------------|-----------------|
| tags-outgoing | 0.885 |
| likes-outgoing | 0.885 |
| tags-incoming | 0.900 |
| likes-incoming | 0.902 |
| comments-outgoing | 0.908 |
| comments-incoming | 0.912 |

| Action-Direction (X) | $H(Y X = true)$ |
|--------------------------|-----------------|
| photo-outgoing | 0.857 |
| video-outgoing | 0.863 |
| link-outgoing | 0.895 |
| link-incoming | 0.896 |
| post-incoming | 0.902 |
| post-outgoing | 0.906 |
| video-incoming | 0.915 |
| photo-incoming | 0.921 |

rather than mutual information $I(\text{likes}(u, i); X_k^{u,i})$, as we found that mutual information is highly correlated with (and dominated by) the frequency of the feature $X_k^{u,i} = true$ in the dataset.

First we analyse various interactions to better understand what interaction-defined SAGs have a high affinity with a user’s preferences. To this end, we make a few observations from the conditional entropy analysis of Table 5:

- Interaction on *videos* indicates a stronger preferential affinity between users than other modalities (links, posts and photos). We conjecture this is because videos are time-consuming to view and hence users mainly watch the videos of those users whose preferences they share.
- Tagging has a slightly better conditional entropy than commenting and liking, potentially since tagging often results from direct social interaction (appearing in a photo or video together) indicating common interests.
- A user is more likely to share preferences with someone who she initiates the interaction with (outgoing) vs. with someone who initiates the interaction with her (incoming). E.g., we note that while outgoing photo and video interactions are *most* informative in the last table of Table 5, it appears that incoming photo and video interactions are *least* informative.

In Fig 5 we plot the conditional entropy of modality and action for incoming/outgoing interactions constrained to links liked by at least k friends in the SAG (measuring the implicit or explicit *exposure* of a user to their friends’ preferences via a SAG). We note that preference affinity with any SAG increases as more people in the SAG like the item. E.g., while incoming interactions were not as predictive as outgoing interactions for the same k , we note that higher k on incoming *can be more predictive* than lower k for outgoing. Similar principles hold for modality and action vs. k — a larger k is generally more predictive than the individual variation among modality and action at a fixed k , an exception being the modality-outgoing analysis.

3.5 Activity Analysis

Now we analyse the informativeness of Activity Social Affinity Features (ASAFs) by looking at the correlation between the size and type of groups, pages and favourites.

Fig 6 provides scatter plots of conditional entropies and logistic regression weights vs. activity group size.⁸ Both plots show that small activity groups *can* be highly predictive (low conditional entropy or weights that deviate extremely from zero) whereas large groups are *rarely* predictive.

In Fig 7 we plot the average conditional entropy of the top 10% of features cumulative up to the size of the activity group given on the x-axis. This graph distinctly shows that the small sizes of groups, pages and favourites have low average conditional entropy that transitions sharply to a higher average at about 50 for groups and 10⁵ for pages/favourites.

We also analyse predictiveness of favourites by categories obtained from the Facebook API in Fig 8. While half of category instances in movie, books, or movies with “long-tail” (less popular, specialised) content may not be highly predictive (judging by median informativeness), these categories do contain some highly predictive instances (as evidenced by the top two quartiles). On the contrary, highly generic categories (e.g. interests) and those with fewer choices (e.g. sports or fav-teams) tend to be less predictive overall. These observations of Fig 8 are also reiterated by the examples provided in Table 6 where uninformative favourites tend to have a broad appeal whereas informative favourites generally appear much more specialised. This reinforces the insight of SAF that it is important to learn which SAGs are predictive.

One might ask how the number of activities a user joins affects recommendation performance. In Fig 9, we see that on average, more user activities generally leads to higher accuracy. As an alternative analysis, Fig 2 shows performance vs. the number of *active features*, i.e., for any (user,item) recommendation in a row of Fig 2, the active features are those that are true (1). Here we see that excessive item popularity among activities hurts the discriminative power of SAF to make good recommendations.

4. RELATED WORK

This work relates to many others in inferring user preferences on social and information networks. We structure the discussion into three parts: the first is concerned with the nature and observations on user traits, interactions and diffusion mechanisms; the second is concerned with correlating these user traits and interactions to user preferences and interests; the third is concerned with methods that use these observations for predicting user interest or recommending content on social networks.

The first group of related work studies the nature of user profile, interactions, and diffusion. Profile information and demographics is correlated with user behaviour patterns. Chang *et al* [8] showed that the tendency to initiate a Facebook friendship differs quite widely across ethnic groups, while Backstorm *et al* [3] have additionally showed that female and male users have opposite tendencies for dispersing attention for within-gender and across-gender communication. Two particular measurement studies on Facebook attention [3, 31] have inspired our work. Although the average number of friends for a Facebook user is close to the human psychological limit, known as the Dunbar number [14], the findings concur that a user’s attention (i.e., interactions) are divided among a much smaller subset of Facebook friends. [3] studied two types of attention: communication interaction and viewing attention (e.g. looking at profiles or photos). Users’ communication attention is focused on small numbers of friends, but viewing attention is dispersed across

⁸Here the size of a *group*, *page* and *favourite* is the number of total users in the activity group. For *pages* and *favorites* this is the total number of Facebook users, whether or not they are in the App users’ ego network, while for *groups* only the number of users in the App users’ ego network is visible to our App.

all friends. This finding supports our approach of looking at many types of user interactions across all of a user’s contact network, as a user’s interest is driven by where he/she focuses attention.

The mechanisms of diffusion invite interesting mathematical and empirical investigations. The Galton-Watson epidemics model suits the basic setup of social message diffusion, and can explain real-world information cascade such as email chain-letters when adjusted with selection bias [13]. For social diffusions in a one-to-many setting, however, the epidemics model has been less accurate. Ver Steeg *et al* [29] found that online message cascades (on Digg social reader) are often smaller than prescribed by the epidemics model, seemingly due to the diminishing returns of repeated exposure. Romero *et al* [25], in an independent study, confirmed the effect of diminishing returns with Twitter hashtag cascades, and further found that cascade dynamics differ across broad topic categories such as politics, culture, or sports. Our observations on user preference on items liked by a number of Facebook friends suggest a large cumulative number of friend preferences is more predictive, although further investigation is needed to pinpoint the effect of diminishing returns on repeated exposures.

The nature of social diffusion seem to be not only democratic [2, 4], but also broadening for users [5]. While influential users are important for cascade generation [4], large active groups of users are needed to contribute for the cascade to sustain [2]. Moreover, word-of-mouth diffusion can only be harnessed reliably by targeting large numbers of potential influencers, confirmed by observations on Twitter [4] and online ads [30]. In a study facilitated by A/B testing on Facebook links, [5] found that while people are more likely to share the information they were exposed to by their strong ties rather than their weak ties, the bulk of information we consume and share comes from people with different perspectives (weak ties). SAF aims to leverage many of these insights for social recommendation by viewing affinity groups as diffusion channels. Yet, information diffusion and recommendation are distinct problems — while we observed best *recommendation* performance using ASAGs ranging across all Facebook users, the vast majority of *information diffusion* happens within one step from the source node [12].

The second group of related work tries to correlate from user interactions to preferences and tie strength. Saez-Trumper *et al* [26] found that incoming and outgoing activities are highly correlated on broadcast platforms such as Facebook and Twitter, and such correlation does not hold in one-to-one mode of communication such as email. Multiple studies have found that online interactions tend to correlate more with interests than with user profile. Singla *et al* [28] found that users who frequently interact (via MSN chat) tend to share (web search) interests. Anderson *et al* [1] concluded that the level of user activities correlate with the positive ratings that they give each other, i.e., it is less about what they say (content of posts) but more about who they interacted with. Such findings echo those by Brandtzig [7] that real-world interactions (e.g., appearing in the same photo evidenced by tagging) further strengthens friendship on Facebook, while virtual interactions reveal interests. Furthermore, ratings of real-world friendship strength and trust [11] seems to be better predicted from the intimacy, intensity, and duration of interactions, than from social distance and network structure. Our work is not only inspired by these observations, we also quantify the strength of correlations of user interest with a large variety of user affinities – namely, activities, and group preferences in different categories.

The third group of related work is concerned with using social network and behavior information for recommendation. Matrix factorization is one of the prevailing approaches for recommender sys-

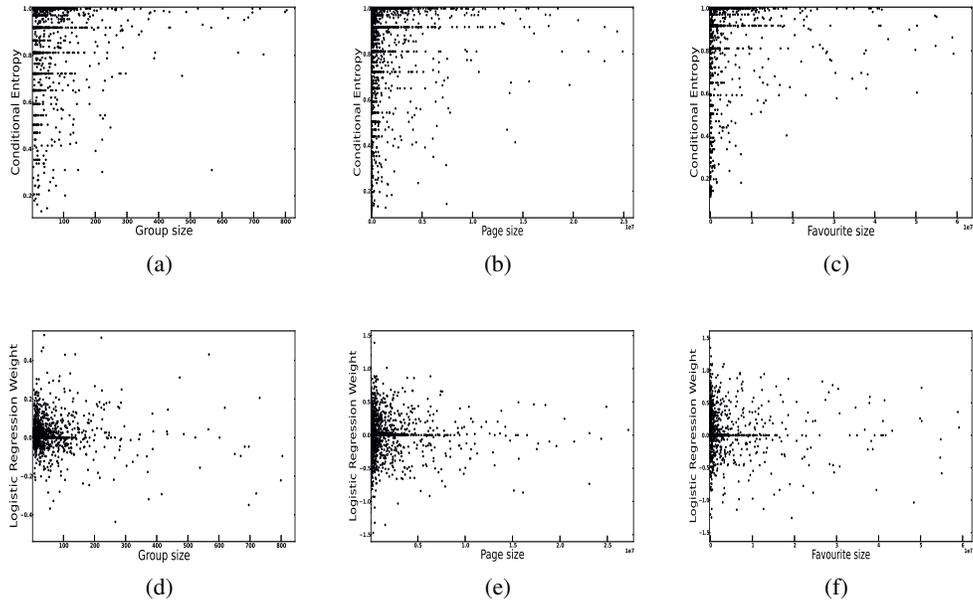


Figure 6: Conditional entropy vs. size (a-c); logistic regression feature weights vs size (d-f). In (a-c) we observe that the large membership ASAGs are rarely informative while the most informative SAGs tend to have low memberships. Similarly in (d-f) we see that the most predictive features with the largest weights (positive or negative) are concentrated toward small ASAGs.

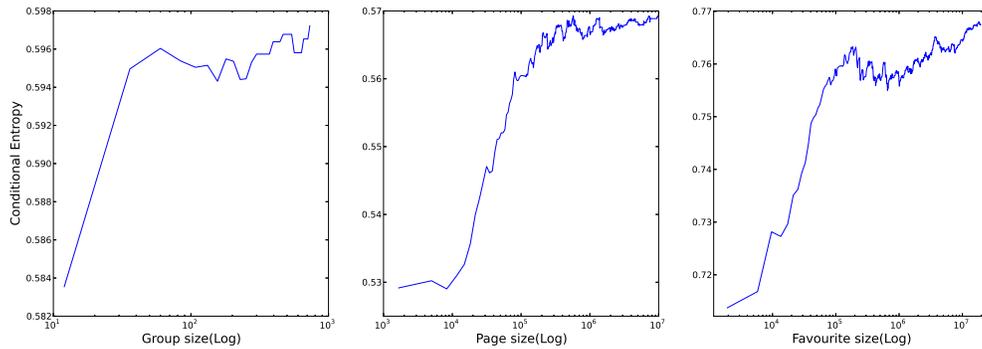


Figure 7: Average conditional entropy of top 10% groups, pages and favourite features *cumulative* over the size. Here we see that as we add in larger membership ASAGs, the average informativeness decreases substantially (entropy increases).

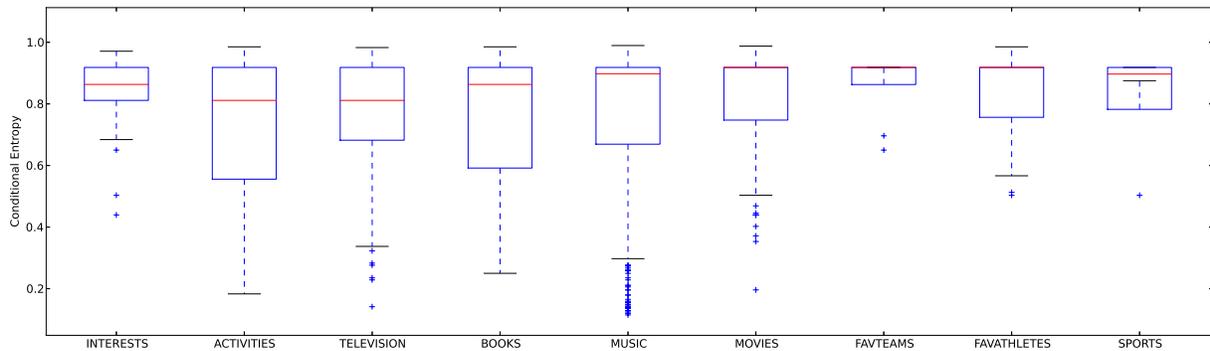


Figure 8: Conditional entropy for top 1000 favourites breakdown by categories. While at least half of ASAG categories with many options like music are not informative (judging by median values), some of the most informative ASAGs are music. This reiterates the point that it is crucial to *learn* which ASAGs are informative rather than aggregating average information.

| Median Informative Favourites by Category | | | | |
|---|--------------------------|-----------------------|----------------------------------|------------|
| Books | Movies | Music | Television | Interests |
| Harry Potter series | Forrest Gump | John Lennon | Futurama | Travel |
| A Song of Ice and Fire | Pretty Woman | U2 | Star Trek | Music |
| Discworld | Napoleon Dynamite | AC/DC | The Trap Door | Literature |
| Hitchhiker's Guide To The Galaxy | Harry Potter | The Smashing Pumpkins | Drawn Together | Painting |
| The Hobbit | Toy Story 3 | Goye | Sherlock(Official) | Running |
| The Magician's Guild | The Godfather | The Rolling Stones | Hitchhiker's Guide to the Galaxy | Sports |
| Ranger's Apprentice | Mulan | All Axess | Buffy The Vampire Slayer | Films |
| Cosmos | How to Train Your Dragon | Steve Aoki | South Park | Genetics |

| Most Informative Favourites by Category | | | | |
|---|----------------------------|--------------------|-------------------------------|------------------|
| Books | Movies | Music | Television | Interests |
| Calvin and Hobbes | Billy Madison | Avascular Necrosis | Metalocalypse | Computers |
| Tomorrow when the War Began | Team America: World Police | Tortured | Beast Wars | Texas HoldEm |
| I really like ceilings | Pan's Labyrinth | Elysian | Hey Arnold! | Programming |
| Angels and demons | Pirates of the Caribbean | Anno Domini | Sherlock | Economics |
| Magician | Aladdin | Darker Half | Hey Hey It's Saturday | Martial arts |
| Digital Fortress | Starship Troopers | Hellbringer | Neil Buchanan and Art Attack! | Graphic design |
| The Bible | Happy Gilmore | Johnny Roadkill | Breaking Bad | Cooking |
| Interview with the Vampire | Timon and Pumbaa | Aeon of Horus | Red vs. Blue | Klingon language |

Table 6: (top) Examples of 8 items per Favourite category near the *median* conditional entropy (*median informativeness*). (bottom) Examples of top 8 items with the lowest conditional entropy (*most informative*). A general trend is that more informative favourite category ASAGs tend to be more specialised in appeal, e.g. "Avascular Necrosis" is an informative music group favourite — its members tend to share common preferences — while "John Lennon" and "U2" have a broader audience with more diverse preferences. Interestingly, "Sherlock" appears in both most and median informative table but the median informative is an official page with wide range of fans, whereas the most informative is a duplicate fan page with few fans.

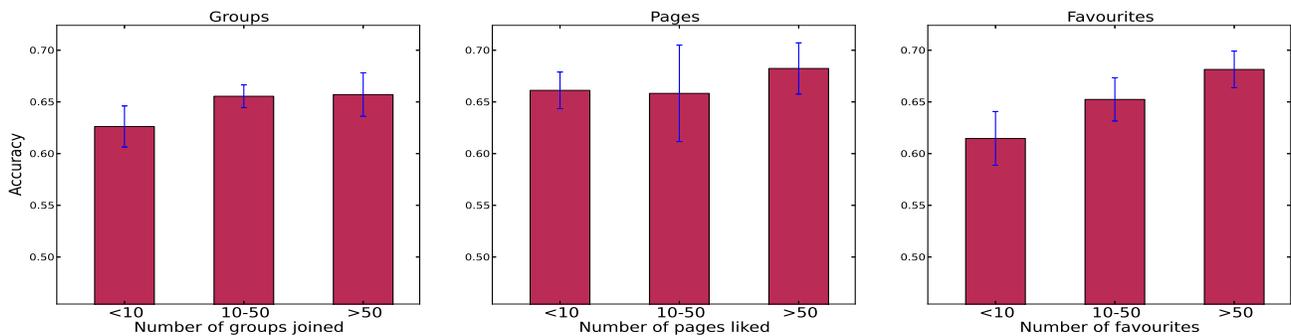


Figure 9: Accuracy of the SAF increases as users become more active in social network by joining more groups/pages/favourites. It does not appear too many activities hurts — SAF learns to discriminate when activities are predictive.

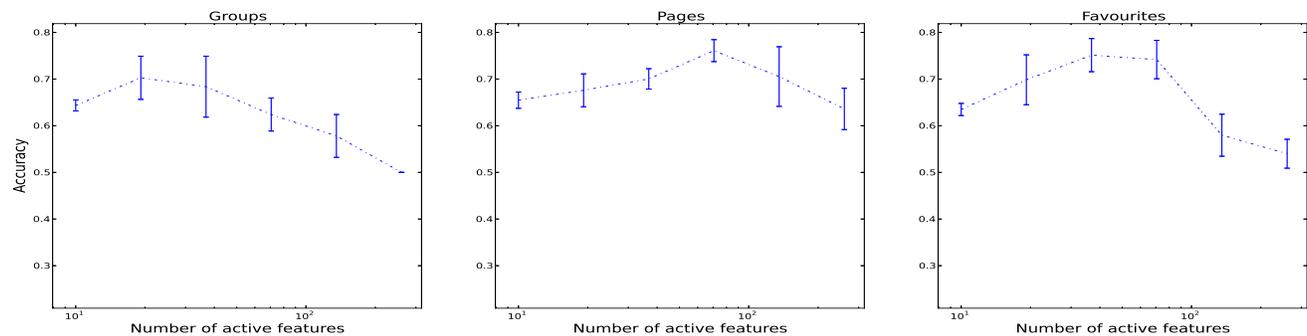


Figure 10: Accuracy increases as the number of active features increases, but then, after reaching a certain limit, it starts to decrease, i.e., excessive item popularity among activities hurts the discriminative power of SAF to make good recommendations.

tems [17, 21]. Recent advances include extending matrix factorization to user social relation in regularization [22, 19], to take into account multiple relations [24, 16], and to model social context [15]. In particular, there are different designs for using social information to regularize objective functions [32], a trust ensemble [20], a low-rank factorization of the social interactions matrix [21], or social-spectral regularization that takes into account user and item features [23]. These systems have shown very promising performance across a range of problems, but their all collapse social affinity (fine-grained interactions and group affinity) into one or a very low-dimensional representation. The point of departure of this work is to explore the predictive power of fine-grained social information.

5. FUTURE WORK

For a summary of the key contributions and insights from this work, we refer the reader back to the points outlined at the end of Sec 1. Future research directions can examine the nature of social groups via additional metrics — e.g., properties of the social network within SAGs or the activity level of SAGs. Other work might explore non-binary representations of ASAFs and ISAFs or SAF-influenced extensions of (social) collaborative filtering approaches like nearest neighbor or matrix factorization to better exploit informative fine-grained social interaction and activity features.

6. ACKNOWLEDGEMENTS

NICTA is funded by the Australian Government as represented by the Department of Broadband, Communications and the Digital Economy and the Australian Research Council through the ICT Centre of Excellence program. This work was partially funded by a Google Research Award and by the US Air Force Research Laboratory, under agreement number FA2386-12-1-4041. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright notation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Air Force Research Laboratory or the U.S. Government.

7. REFERENCES

- [1] A. Anderson, D. Huttenlocher, J. Kleinberg, and J. Leskovec. Effects of User Similarity in Social Media. In *WSDM*, 2012.
- [2] S. Asur, B. A. Huberman, G. Szabo, and C. Wang. Trends in social media: Persistence and decay. In *ICWSM*, 2011.
- [3] L. Backstrom, E. Bakshy, J. Kleinberg, T. Lento, and I. Rosenn. Center of attention: How facebook users allocate attention across friends. In *ICWSM*, 2011.
- [4] E. Bakshy, J. M. Hofman, W. A. Mason, and D. J. Watts. Everyone’s an influencer: quantifying influence on twitter. In *WSDM*, 2011.
- [5] E. Bakshy, I. Rosenn, C. Marlow, and L. Adamic. The role of social networks in information diffusion. *Facebook Report*, <http://www.scribd.com/facebook>, 2012.
- [6] R. M. Bell and Y. Koren. Scalable collaborative filtering with jointly derived neighborhood interpolation weights. In *ICDM*, 2007.
- [7] P. B. Brandtzaeg and O. Nov. Facebook use and social capital — a longitudinal study. In *ICWSM*, 2011.
- [8] J. Chang, I. Rosenn, L. Backstrom, and C. Marlow. epluribus: Ethnicity on social networks. In *ICWSM*, 2010.
- [9] P. Cui, F. Wang, S. Liu, M. Ou, and S. Yang. Who should share what? item-level social influence prediction for users and posts ranking. In *SIGIR*, 2011.
- [10] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin. Liblinear: A library for large linear classification. *JMLR*, pages 1871–1874, 2008.
- [11] E. Gilbert and K. Karahalios. Predicting tie strength with social media. In *Proc. CHI*. ACM, 2009.
- [12] S. Goel, D. J. Watts, and D. G. Goldstein. The structure of online diffusion networks. In *EC*, 2012.
- [13] B. Golub and M. O. Jackson. Using selection bias to explain the observed structure of internet diffusions. *Proc. Nat. Academy Sci.*, 107(24), 2010.
- [14] R. Hill and R. Dunbar. Social network size in humans. *Human Nature*, 14(1):53–72, 2003.
- [15] M. Jiang, P. Cui, R. Liu, Q. Yang, F. Wang, W. Zhu, and S. Yang. Social contextual recommendation. In *CIKM*, 2012.
- [16] M. Jiang, P. Cui, F. Wang, Q. Yang, W. Zhu, and S. Yang. Social recommendation across multiple relational domains. In *CIKM*, 2012.
- [17] Y. Koren, R. Bell, and C. Volinsky. Matrix factorization techniques for recommender systems. *Computer*, 42, 2009.
- [18] K. Lang. NewsWeeder: Learning to filter netnews. In *ICML*, 1995.
- [19] W.-J. Li and D.-Y. Yeung. Relation regularized matrix factorization. In *IJCAI*, 2009.
- [20] Ma, King, and Lyu. Learning to recommend with social trust ensemble. In *SIGIR*, 2009.
- [21] H. Ma, H. Yang, M. R. Lyu, and I. King. Sorec: Social recommendation using probabilistic matrix factorization. In *CIKM*, 2008.
- [22] H. Ma, D. Zhou, C. Liu, M. R. Lyu, and I. King. Recommender systems with social regularization. In *WSDM*, 2011.
- [23] J. Noel, S. Sanner, K.-N. Tran, P. Christen, L. Xie, E. V. Bonilla, E. Abbasnejad, and N. Della Penna. New objective functions for social collaborative filtering. In *WWW*, 2012.
- [24] S. Rendle, L. B. Marinho, A. Nanopoulos, and L. Schmidt-Thieme. Learning optimal ranking with tensor factorization for tag recommendation. In *KDD*, 2009.
- [25] D. M. Romero, B. Meeder, and J. Kleinberg. Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter. In *WWW*, 2011.
- [26] D. Saez-Trumper, D. Nettleton, and R. Baeza-Yates. High correlation between incoming and outgoing activity: A distinctive property of online social networks? In *ICWSM*, 2011.
- [27] R. Salakhutdinov and A. Mnih. Probabilistic matrix factorization. In *NIPS*, 2008.
- [28] P. Singla and M. Richardson. Yes, there is a correlation: - from social networks to personal behavior on the web. In *WWW*, 2008.
- [29] G. Ver Steeg, R. Ghosh, and K. Lerman. What stops social epidemics? In *ICWSM*, 2011.
- [30] D. J. Watts and P. S. Dodds. Influentials, networks, and public opinion formation. *Journal of Consumer Research*, 2007.
- [31] C. Wilson, B. Boe, A. Sala, K. Puttaswamy, and B. Zhao. User interactions in social networks and their implications. In *EuroSys*, 2009.
- [32] S.-H. Yang, B. Long, A. Smola, N. Sadagopan, Z. Zheng, and H. Zha. Like like alike: Joint friendship and interest propagation in social networks. In *WWW*, 2011.