Research Article

Recommending Ads from Trustworthy Relationships in Pervasive Environments

**Francisco Martinez-Pabon,¹ Juan Camilo Ospina-Quintero,²
Gustavo Ramirez-Gonzalez,¹ and Mario Munoz-Organero³**

¹Telematics Engineering Group (GIT), University of Cauca, Street 5 No. 4-70, Popayán 190003, Colombia

²Cluster CREATIC, Calle 17N No. 6-21, Popayán 190002, Colombia

³Telematic Applications and Services Group (GAST), Carlos III University of Madrid, Campus Leganés, Avenida Universidad 30, 28911 Madrid, Spain

Correspondence should be addressed to Francisco Martinez-Pabon; fomarti@unicauca.edu.co

Received 27 January 2016; Revised 18 May 2016; Accepted 7 June 2016

Academic Editor: Stavros Kotsopoulos

Copyright © 2016 Francisco Martinez-Pabon et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The use of pervasive computing technologies for advertising purposes is an interesting emergent field for large, medium, and small companies. Although recommender systems have been a traditional solution to decrease users' cognitive effort to find good and personalized items, the classic collaborative filtering needs to include contextual information to be more effective. The inclusion of users' social context information in the recommendation algorithm, specifically trust in other users, may be a mechanism for obtaining ads' influence from other users in their closest social circle. However, there is no consensus about the variables to use during the trust inference process, and its integration into a classic collaborative filtering recommender system deserves a deeper research. On the other hand, the pervasive advertising domain demands a recommender system evaluation from a novelty/precision perspective. The improvement of the precision/novelty balance is not only a matter related to the recommendation algorithm itself but also a better recommendations' display strategy. In this paper, we propose a novel approach for a collaborative filtering recommender system based on trust, which was tested throughout a digital signage prototype using a multiscreen scheme for recommendations delivery to evaluate our proposal using a novelty/precision approach.

1. Introduction

Currently, the use of pervasive computing technologies for advertising purposes is an interesting emergent field for large, medium, and small companies. One of the marketing areas, advertising, is defined as “any paid form of non-personal presentation and promotion of ideas, goods or services by an identified sponsor” [1]. Frequently, research has focused on advertisement delivery to potential clients through their personal devices (smartphones, tablets, or PCs), but public space is still attractive for advertisers, taking into account that 75% of purchase decisions are made at or near the sale place [2]. This field, known as digital signage, is related to digital content display using public screens [3]. However, there are challenges associated with pervasive advertising spaces implementation, which can be analyzed from the following

perspectives: First, the need for personalized content is an issue that can be addressed from recommender systems theory, which provides tools and techniques to provide suggestions in a huge collection of items for a specific user. Specifically, collaborative filtering has been one of the most popular techniques for recommendation purposes; it looks for correlations between users to recognize their affinity and then associates their items evaluations [4]. Nonetheless, the customization process also demands the inclusion of information about the user context. It may be challenging to include all variables that may be considered as “context” in the recommendation algorithm, but one variable is particularly interesting for the advertising domain: the user's “social context.” For years, the “word of mouth” has been a powerful technique for marketing purposes, so the inclusion of the user's social context information in the recommendation

algorithm (specifically, trust in other users) is a mechanism for obtaining ads' influence from other users in the closest social circle. The traditional collaborative filtering used in recommender systems calculates recommendations based on users' similarities in their ratings, but they are anonymous users. The inclusion of social context information may improve the ads recommendation process with persuasion purposes, considering that 67% of purchase decisions are influenced by the opinions of close others. Although some studies have developed proposals for the inclusion of information in users' social circle, they are frequently based on explicit mechanisms to calculate the trust between users, or they try to infer this same information from the ratings matrix itself. Other approaches define mechanisms to infer trust from social networks, but, in most cases, they use proprietary networks or the integration strategy into the recommendation algorithm is not completely defined and tested.

Another main concern is related to the evaluation focus for a recommender system in a pervasive advertising domain. Frequently, the evaluation of recommender systems has been concentrated on precision aspects, but the recommendation novelty may be an aspect even more relevant for persuasion purposes in advertising. Nonetheless, the improvement of the precision/novelty balance for the recommender system in these cases may be not only a matter of the RS algorithm itself but also a better display strategy issue. In digital signage environments, for example, most public display interaction initiatives do not consider multiscreen approaches, where the content may be distributed between different devices (e.g., public screens and smartphones). A screen content replication has been used instead; if this approach combines aggregation techniques for delivering recommendations to groups of people on the main screen with a more robust trust-based recommendation algorithm to deliver custom items on a personal screen (e.g., a smartphone), the balance between the precision and the novelty of the recommendations could improve. On the other hand, several approaches have been defined to measure novelty in different recommender system variants, but a suitable definition for the advertising domain deserves a deeper analysis, which is one of the contributions of this work.

According to the previous analysis, our proposal is focused on the research question of how to build a recommender system for pervasive advertising environments supported on a Smart TV-smartphone framework, and the following hypotheses are evaluated: (H1) the inclusion of multiscreen capabilities for the recommendations delivery improves the precision/novelty balance; (H2) the inclusion of trust information in the collaborative filtering algorithm helps to improve the precision/novelty balance.

The current work develops a new recommender system approach for the advertising domain and makes the following contributions: a trust inference algorithm from social network interaction information; a collaborative filtering recommendation algorithm variant to include trust information during the recommendation process; a digital signage case study definition, where we analyzed the effects of the trust inclusion and multiscreen display strategy over recommendation precision and novelty; and an evaluation scheme

derived from the work of Vargas and Castells [5] to measure the novelty in the pervasive advertising domain.

This paper is structured as follows: Section 2 summarizes some related works, Section 3 introduces the trust inference algorithm, Section 4 describes the strategy to include trust in classic collaborative filtering, Section 5 introduces the study case and the features of the implemented prototype, Section 6 describes the results of the experiments, and Section 7 outlines conclusions and future work.

2. Related Works

From pervasive advertising (specifically digital signage) as a field study, we will next summarize some of the most relevant related works, taking into account the main concepts for the current research.

Recommender systems have evolved from data mining and machine learning theories. They have been studied for years as a classic solution to decrease the cognitive effort when there are a huge collection of items the users may explore, independent of a specific domain. Some previous works have compiled the concepts, algorithms, and techniques that are more relevant from the recommender systems perspective [6–8]. Specifically, collaborative filtering has been one of the most used techniques for recommender systems. It looks for correlations between users to recognize their affinity and then to associate their item evaluations [4]. Some previous works have applied collaborative filtering techniques for advertising purposes [9], including some context variables, but from an explicit approach; the inclusion of trust during recommendation has been less explored for this specific domain.

By and large, the inclusion of trust in recommender systems may be studied from two approaches: explicit trust information from users or trust inferred from users' information. Regarding the first approach, Massa and Avesani [10] extended classic recommender system algorithms by including a trust matrix in addition to the ratings matrix to replace the traditional prediction mechanism with an algorithm that calculates the trust propagating in a network, and they used this estimation instead of similarity. Golbeck [11] developed FilmTrust, a website that uses trust inferred from a proprietary social network to offer movie recommendations; the work focuses on determining how to create interfaces to represent the connections between users based on the information they provided by themselves using an algorithm called TidalTrust. Other similar propagation approaches may be seen in [12, 13]. Several other works have developed proposals for trust inference, which are related to the second approach. O'Donovan and Smyth [14] extend the traditional user $\times$ item space to a user $\times$ item $\times$ context space, defining a trust matrix inferred from user ratings; a similar proposal was developed by Martín-Vicente et al. [15] but for expertise and reputation inferences. Other proposals go beyond and try to infer trust from social network information. For example, Chen and Fong [16] propose a recommendation framework for social networks based on collaborative filtering and trust. In this research, two methods to compute similarity are defined: the first one uses the similarity of profiles, and the second one is based on people's interactions. From these two values,

a unique value is computed and included in the recommendation algorithm; additionally, the authors propose a framework to use the social network Facebook. Bakshy et al. [17] demonstrated that the interactions between people in a social network are the strongest component of trust inference, and other research works are coherent with this trend, defining frameworks based on interaction information to infer trust. Gilbert and Karahalios [18] define a predictive model that maps social media data to tie strength on a dataset of over 2,000 social media ties. As a result, the top 15 predictive variables for trust inference were defined. Other proposals have used similar approaches based on interaction information, and we found several agreements between the reference variables used for trust inference purposes [19–21]. These works set an important starting point for the current research, and we will expand on some aspects of them in the following sections.

Although the previous works are important advances in trust inference with recommendation purposes, there are some noted gaps with regard to the context of the current research: the pervasive environment demands to capture as much information as possible in an implicit way, and the pervasive advertising domain is not an exception, so trust inference mechanism from explicit information captured from the users is not attractive for this domain. Although the trust inference from the ratings matrix itself or from an extended user $\times$ item $\times$ context matrix is interesting, it still infers information from anonymous users, and there is no assurance that these users belong to the closest social circle of the user. According to the previous characterization of the advertising domain, we are interested in inferring information from friends, which is a disadvantage of these techniques and classic collaborative filtering. Therefore, trust inference from social network information seems to be a more suitable solution for the pervasive advertising domain, but the use of proprietary social networks restricts the system scalability for recommendation purposes. On the other hand, there is no consensus about the best variables to infer trust from public social networks such as Facebook, and there is no clear strategy to include trust information in the recommendation algorithm. These aspects deserve deeper research.

Another important aspect is the evaluation approach to use for a trust-based recommender system in the pervasive advertising domain. One of the more demanding challenges in recommender systems theory is related to metrics fragmentation to evaluate different aspects of these types of systems. Herlocker et al. [22] developed an empirical analysis for different accuracy metrics, and Gunawardana and Shani [23] defined a guide for the design of offline experiments with evaluation purposes. Traditionally, precision has been the metric frequently used to evaluate the recommender system, but, according to the pervasive advertising domain features, the novelty of the recommendations may be more relevant with persuasion purposes. McNee et al. [24] propose new directions to evaluate recommender systems, including the novelty degree. Ge et al. [25] analyze the roles of coverage and novelty in the recommendation quality and introduce methods to measure them. Kawamae [26] proposes an algorithm to generate novel recommendations that focuses on

the search time that, in the absence of any recommendation, each user would need to find a desirable and novel item by himself, following the hypothesis that the degree of the user's surprise is proportional to the estimated search time. In one of the most interesting works related to novelty measurement, Vargas and Castells [5] developed a formal framework for the definition of novelty and diversity metrics that unifies and generalizes several state-of-the-art metrics. The novelty of a piece of information generally refers to how different it is with respect to “what has been previously seen,” by a specific user, as long as the diversity generally applies to a set of items, and it is related to how different the items are with respect to each other. This is related to novelty in that when a set is diverse, each item is novel with respect to the rest of the set [5]. This work is used as a starting point for the novelty metric definition in the current research, and its contribution will be described later in detail.

Finally, although the previous works define methods to measure the novelty degree in the recommendation, some of them also define approaches to influence novelty from the recommendation algorithm itself. However, as one of the main hypotheses of the current research, the balance between the precision and the novelty of recommendations may be affected by not only the recommendation algorithm itself but also a better recommendations' display strategy. For our specific study case, in the digital signage space design, we consider some works related to the delivery of recommendations to groups of people [27, 28]: the PolyLens system, a variant of MovieLens [29], MusicFX [30], and Intrigue [31]. Other approaches have defined mechanisms to improve the interaction between users and public displays ([32–34] or [35–38]), but a multiscreen approach from the recommender systems perspective and its effect on the balance between precision and novelty have not been explored.

3. Inferring Trust from Social Networks

The trust inference concept is strongly related to the homophily concept, a principle that postulates that people tend to form ties with other people who have similar characteristics [39]. In the most basic sense, this approach can lead to a relationship binary analysis to study a friend/no friend condition. However, this approach is not sufficient to try to infer trust between individuals, as is illustrated in the following example from social network analysis: suppose that user A and user B are friends, and user B and user C are also friends, but user A and user C are only acquaintances. A few days later, user C sends a friendship request to user A, and now they are connected as friends when A accepts the invitation. From the binary approach perspective, there are no differences between the relationship between users A and B and that between users A and C, although this is not accurate in practice. Trust inference is a concept that goes beyond a simple “friend/no friend” status analysis, and a new concept arises from the social network analysis to complement this approach: tie strength. The tie strength feature tries to calculate the strength of the relationship between two users when a tie exists between them. Therefore, a tie strength analysis will deliver

two possible main results: strong ties (real friends) and weak ties (acquaintances) [40].

Although profile information could provide some level of information about the similarities between people, previous research has found that interaction information is one of the most important sources to try to predict the tie strength. Several techniques may be applied to infer tie strength using diverse technologies: the reciprocity of calls between two mobile phone users, the number of tweets between Twitter users, or even the email exchange activity between individuals. Nonetheless, social networks provide a richer space to infer trust from several types of interactions and therefore build a more accurate picture of the trust map of a specific user. Specifically, the current research will use the social network Facebook as a reference. Other works have built proprietary social networks for their experiments, but the study of Facebook may enable a more scalable solution in several domains due to the network's popularity and use by millions of users.

As a starting point, the current work uses the model proposed in [18] where the authors introduce a predictive model that maps social media data to tie strength on a dataset of over 2,000 social media ties to distinguish between strong and weak ties. Table 1 shows a summary of the predictive model proposal for the variables in different categories.

This model defines the top 15 predictive variables; we compare these top 15 results with similar research to obtain a unified set of variables as the starting point for our study (see Figure 1).

According to the previous results, seven variables are selected as a starting point to build the trust inference algorithm between two users, A and B (as shown in the following part).

First set of variables to infer trust from Facebook are as follows:

- Exchanged inbox messages.
- Likes (from B to A).
- Tags to B.
- Tags from B.
- Cotags (posts where both users are tagged together).
- Comments (from B to A).
- Wall posts (from B to A).

Once we define the first set of variables for trust inference, we define a first approach for the algorithm. The main goal is to build an equation that combines the contributions of this set of variables to calculate a trust score that can later be integrated into a collaborative recommender filtering technique and keep it as simple as possible for performance purposes. Because trust is not necessarily symmetric, from the work of [41] and according to the simple Multiple

TABLE 1: Social media data predictive model.

Intensity	Wall words exchanged
	Wall posts
	Inbox messages exchanged
	Status updates
	Comments
	Likes
Intimacy	Tags
	Number of friends
	Days since last communication
	Wall and inbox intimacy words
	Appearances together in photos
	Distance between hometowns
Duration	Relationship status
	Days since first communication
Reciprocal services	Links exchanged by wall post
	Applications in common
Structural	Mutual friends
	Groups in common
	Similarity of interests
	Common networks
Emotional support	Number of positive words
	Number of negative words
	Gifts exchanged
Social distance	Age difference
	Occupations difference
	Educational difference
	Political and religious difference
Demographic	Gender
	Number of apps installed
	Number of inbox messages
	Number of wall posts
	Number of photo's comments

Source: [18].

Attribute Utility Theory (MAUT), we initially define the trust that user A has in user B as follows:

$$\begin{aligned}
 T_{A,B} = & \alpha \frac{\sum IM}{\text{totIM}} + \beta \frac{\sum Likes}{\text{totLikes}} + \gamma \frac{\sum TagsTo}{\text{totTagsTo}} \\
 & + \delta \frac{\sum TagsFrom}{\text{totTagsFrom}} + \epsilon \frac{\sum Cotags}{\text{totCotags}} + \theta \frac{\sum C}{\text{totC}} \\
 & + \emptyset \frac{\sum WP}{\text{totWP}},
 \end{aligned} \quad (1)$$

where $\alpha + \beta + \gamma + \delta + \epsilon + \theta + \emptyset = 1$, IM is inbox messages exchanged, C is comments, WP is wall posts, and tot is the prefix for "total," which is the total amount of items for a particular attribute, that is, the total amount of inbox messages exchanged by a certain user with all his friends. This normalization makes sense when we consider the interaction frequency between individuals as a clear indicator of trust. For example, suppose user A has the number of interactions with users B and C, as shown in Figure 2. At first glance, A could trust C more than B, but this is likely because C is a more active user than B, so the activity level of both

FIGURE 1: State of the art: trust inference from user interactions.

FIGURE 2: Users' interactions in a social network.

users must be taken into account. Suppose B has a total of 5 interactions with his friends, and C has 10 interactions. When normalization takes place, we see that almost 60% (3/5) of B interactions are with user A, as long as the interactions only represent the 40% (4/10) for user C, so the effect of the active user is strong. In conclusion, it is important to have a global vision of the network activity and not to focus on only the trust inference in the specific pairs' activity. Other works have used similar approaches for normalization [17, 21].

The final challenge is related to a correct estimation of the weights of each variable in the equation. This is not a trivial issue, and some proposals have used empirical estimations and subjective weights during experimentation [42]. Although it may be an acceptable approach, we looked for a mechanism to combine correlated variables; for example, a tag may trigger a comment or a like, so the numbers of these interactions are related. Accordingly, we choose to adapt the method suggested by Li [21], which applies the principal component analysis (PCA) statistical procedure. PCA uses an orthogonal transformation to convert a set of observations of possibly correlated variables into a set of uncorrelated values by finding a smaller set of linear combinations for the interaction variables. As a result, it also assigns weights to each independent component, which simplifies the process of calculating the weights according to the contribution of each interaction type. In summary, the process obtains the trust score for each friend of a user u , as follows.

(1) *Interaction Matrix Calculation.* For each friend j of user i ($j \in Fi$), an interaction vector is defined, $\vec{s}_{ij}$; therefore, each row represents the interaction between friend j and user i , and each column represents a type of interaction ($|Fi|$ rows and 7 columns, one for each interaction variable). Then, the matrix is normalized according to the previous requirements analysis.

(2) *Covariance Matrix Calculation.* This step looks for relationships between the set of variables. At the end, we select the eigenvectors that correspond to the largest eigenvalues to obtain a linear principal component combination, as follows:

$$T_j = \sum_{i=1}^m \frac{\lambda_i}{\sum \lambda_m} \times u_i \times Z_{ij},$$

$$Z_{ij}^k = \frac{s_{ij}^k - \bar{x}^k}{v^k},$$

$$\bar{x}^k = \frac{\sum x^k}{|Fi|},$$

$$v^k = \sqrt{\frac{\sum (s_{ij}^k - \bar{x}^k)^2}{|Fi| - 1}}.$$
(2)

(3) *Trust Score Calculation.* The trust score is calculated in two steps:

- (a) We obtain an initial score according to each principal component contribution (see (2)).
- (b) Then, we normalize this score on a -1 to 1 scale to ease the score integration with the recommendation algorithm, as it will be described later:

Trust score

$$= \frac{(\text{PCA}_{\text{score}} - \text{PCA}_{\text{min}}) \times (\text{Score}_{\text{max}} - \text{Score}_{\text{min}})}{\text{PCA}_{\text{max}} - \text{PCA}_{\text{min}}} + \text{Score}_{\text{min}}. \quad (3)$$

Once the method is defined to calculate a trust score between users from social network information, we validate this first set of variables to consider the perception of real users regarding the role of each variable in trust inference (see Appendix A).

From the results, we observe that inbox messages, comments, and wall posts contribute more to trust inference from

TABLE 2: Ground truth test results for trust inference.

Test	Variables	Weighting method	Average	Accuracy
1	Inbox messages, comments, wall posts	PCA	18.4	63%
2	Inbox messages, comments, wall posts, tags to	PCA	17.6	65%

the users' perspective. With these results as a starting point, we performed a ground truth test to calculate the accuracy of the trust inference algorithm. The top ten friends for each participant were computed using the first version, including the seven initial variables, and then the second version, updating the variables according to the users' perspective tests. Then, participants were asked to rank their calculated top friends from most trusted to least trusted (ground truth).

Finally, we compared the algorithm results for the two versions and the ground truth results using a simple subtraction between the order given by the algorithm and the order given by the users. The results from the second version of the algorithm outperformed those of the first version, so we conducted additional tests to verify whether the tag information inclusion was relevant (Table 2).

Table 2 shows that although there is no significant effect by excluding the tags variable, it slightly improves trust inference accuracy. Therefore, the final trust inference equation was defined as follows:

$$T_{A,B} = \alpha \frac{\sum IM}{\text{totIM}} + \gamma \frac{\sum \text{TagsTo}}{\text{totTagsTo}} + \theta \frac{\sum C}{\text{totC}} + \emptyset \frac{\sum WP}{\text{totWP}}. \quad (4)$$

4. Trust-Based Collaborative Filtering

Usually, recommender systems are used to accurately estimate the degree to which a particular user will like a particular item. Specifically, collaborative filtering is a recommendation algorithm that bases its predictions on the ratings or behavior of other users in the system; that is, it finds users whose previous ratings are similar to those of the current user and uses those ratings to predict what the current user will like [43]. Traditionally, collaborative filtering recommender systems use a similarity metric to find the user's neighbors, and, based on the preferences in that neighborhood, they compute a prediction for an item. The collaborative filtering algorithm is defined by the following aggregation function:

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_u) * W_{a,u}}{\sum_{u=1}^n W_{a,u}}, \quad (5)$$

where $p_{a,i}$ is the predicted rating for user a on item i , $\bar{r}_a$ represents the average rating for user a , n is the neighborhood, and $W_{a,u}$ is a similarity metric between users, frequently calculated using Pearson's correlation coefficient; other alternatives as Spearman's correlation coefficient are used too. In practice, the traditional collaborative recommender systems exhibit weaknesses related to the sparse nature of data (users typically rate only a small fraction of the available items), the cold

start problem (new users have not rated enough items to be linked to similar users), or, even more important, the algorithm philosophy itself to calculate the similarity with anonymous users. Swearingen and Sinha [44] and Sinha and Swearingen [45] demonstrate that people tend to rely more on recommendations from people they trust than anonymous users; it therefore is a strong motivation to try to incorporate trust in a classic collaborative filtering technique. In the literature, two strategies are commonly used to include trust in the recommendation algorithm: the trust-based weighted mean and trust-based collaborative filtering.

The first one redefines the recommendation strategy by computing a *trust-based weighted mean*; instead of computing the average rating for item i from the ratings $r_{u,i}$ from all system users u who are already familiar with i , it includes trust values $t_{a,u}$, which reflect the trust degree for raters u . Therefore, the highly trusted users ratings will have more weight:

$$p_{a,i} = \frac{\sum_{u \in n} t_{a,u} r_{u,i}}{\sum_{u \in n} t_{a,u}}. \quad (6)$$

Golbeck [46] proposed an algorithm called TidalTrust, which follows this strategy. According to Goldbeck's research findings, this strategy does not necessarily offer a clear benefit over the classic collaborative filtering, but it improves recommendations for users who disagree with the average rating for a specific item.

The second approach introduces an alternative for the weight calculation ($W_{a,u}$ in (5)), which tries to infer the weights throughout the relations of the target user in the trust network (using propagation or aggregation techniques). In (5), PCC weights are replaced by trust values $t_{a,u}$:

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_u) * T_{a,u}}{\sum_{u=1}^n T_{a,u}}. \quad (7)$$

Massa and Avesani [47] propose an example for this strategy. According to their findings, the strategy improved the behavior for controversial users, but it also improved the prediction accuracy for cold start users.

Although both approaches show improvement for controversial item rating predictions and cold start recommendations, they replace the correlation component that calculates the similarity between users, likely based on the assumption that trust and similarity are correlated, according to the analysis of Victor et al. [48]. However, we consider that it could not be an exact assumption as long as the trust between users does not necessarily indicate "similar tastes." Therefore, we adapted a trust-based collaborative filtering strategy, replacing the similarity correlation component of (5) with a new one, including a weighted contribution for similarity and trust:

$$p_{a,i} = \bar{r}_a + \frac{\sum_{u=1}^n (r_{u,i} - \bar{r}_u) * WTS_{a,u}}{\sum_{u=1}^n W_{a,u}}, \quad (8)$$

$$WTS_{a,u} = \alpha W_{a,u} + \beta T_{a,u}, \quad (9)$$

where $\alpha + \beta = 1$, $T_{a,u}$ represents the trust score between two users, calculated throughout the process illustrated in

FIGURE 3: Trust inclusion in the recommendation algorithm.

Section 3, and $W_{a,u}$ reflects the correlation between users, calculated by techniques as Pearson's correlation coefficient or Spearman correlation coefficient; in strict sense, Pearson's correlation coefficients have requirements related to the normal distribution of the data, so a nonparametric coefficient as Spearman may be used instead. As a matter of fact, Lathia et al. [49] reported, for example, that the precision of the recommender system is not affected significantly by the choice of the similarity measure, so it should not be a concern in the practice.

In simple terms, we calculate the correlation coefficient between users and then the trust score for neighborhood users (see Figure 3). An alternative could be to use a trust-based filtering technique where the trust values act as a filter, so only the most trustworthy neighbors participate in the recommendation process, as suggested by O'Donovan and Smyth [14]. However, it demands a highly connected group in the social network to assure an adequate number of users in the neighborhood setup, so this strategy could be more demanding to implement in practice.

In practice, α and β are weights that enable the algorithm calibration to give higher priority to the trust or similarity contribution. We used this feature for different tests during the experimentation using different weights for trust and similarity.

5. Trust-Based Recommendations on Digital Signage Environments: An Implementation Approach

Advertising has played an important role in commerce since it originated. Recently, a new paradigm known as pervasive advertising, which refers to the use of pervasive computing technologies for advertising purposes, has arisen as a promising bet for modern advertisers and consumers. Although most pervasive advertising approaches have been directed to

mobile devices (smartphones or tablets), public spaces are also very interesting for the industry, taking into account that the 75% of purchase decisions are made at or near the purchase place [2]. This field, known as digital signage, is related to digital content display using public screens [3]. We implemented a digital signage prototype as a study case to test the new recommendation algorithm based on trust and similarity contributions. This scenario implies the analysis of interesting requirements for the purpose of the current research: first, ads are personalized and adapted to the context; second, ads are addressed to a group of people instead of individuals; finally, the precision of recommendations is a metric frequently evaluated in RS research, but the extent of novelty is also important in the advertising domain as part of the persuasion. Therefore, a balance between precision and novelty is a desirable feature.

Regarding the first requirement, for all variables that may be considered as "context," likely the most relevant variable for the advertising domain is the social context. For years, "word of mouth" has been a powerful technique for marketing purposes, so the inclusion of users' social context information in the recommendation algorithm (specifically, trust in other users) is a mechanism for obtaining ads' influence on other users in the closest social circle, instead of anonymous users, as in traditional collaborative filtering. The inclusion of trust information may improve the ad recommendation process with persuasion purposes, knowing that 67% of purchase decisions are influenced by the opinions of close others.

Regarding the second requirement, the recommendation process for a group of people implies a new set of challenges, which have been addressed in several studies through the use of special aggregation techniques during the recommendation process; specifically, the techniques described by Masthoff [28] were used as reference. However, the recommendation process improvement in these cases may be not only a matter of the RS algorithm itself but also a better display strategy issue. Most public display interaction initiatives do

FIGURE 4: Prototype basic architecture.

not consider multiscreen approaches, where the content is distributed between screens in a complementary way; screen content replication has been used instead. This thought is related to the third factor too because a better display strategy may also contribute positively to the novelty perception: by definition, group recommendations displayed on a public screen will be less personalized than recommendations displayed in a personal device (e.g., smartphone or tablet), so they could be more novel for users. This hypothesis is an important contribution of the current research, and it will be analyzed in the following section for the current scenario.

According to the previous description, we implemented a novel electronic alternative for a traditional static ads board, in which people post ads using paper posters; these boards are frequently found in small shops or on academic campuses. The proposed implementation replaces the old board with a new cooperative Smart TV-smartphone model, where both devices' screens offer ads to users under different but complementary approaches: ads recommendations for group profiles are on public TV screens, and ads recommendations for individual profiles are on smartphones screens. Moreover, the interaction capabilities between the two devices change the static behavior of the traditional board; the basic architecture for the prototype is shown in Figure 4.

In summary, Android applications were developed for mobile phones and Google Smart TV set top box. These applications used a middleware API for interaction management developed for us. The users in front of the public display log in to the system using a "Login with Facebook" feature. The middleware details are out of the scope of this paper. Please refer to [50] for more details.

Briefly, some of the main functionalities of the implemented prototype are as follows:

- (i) Ad recommendations for a group of users watching the TV screen: the group of users was limited to four people, basically for usability reasons, taking into account the screen size (42 inches). The recommender calculates the best ad list for the people interacting with the public screen using aggregation techniques;

FIGURE 5: Public display ads screen.

FIGURE 6: Mobile device list screen.

these ads were organized on a list of six ads for each one (see Figure 5).

- (ii) Ad recommendations according to individual preferences on the smartphone screen: the recommender calculates the best ad list for each person in front of the public display, and it shows the ads on the mobile device screen (see Figure 6). It uses the recommendation algorithm described in the previous section. The user can obtain detailed information for a particular advertisement and add ads to his or her favorites.
- (iii) Basic interaction between mobile device application and public display: the user can go over the ads on the public display using a control pad from the mobile application (Figure 7); using a tap gesture, the user can obtain detailed information for a particular ad in a public display and watch it on his or her mobile device screen. Each user is identified by a specific color.
- (iv) Explicit and implicit ad ratings: users can rate the ads on a public screen or the mobile device list using the "Like" and "No Like" options available in the mobile application user interface. Thanks to the middleware capabilities, it is also possible to generate implicit ratings according to the users' actions (e.g., request more information about an ad, ignore ads from the screen, and add the ad to favorites).
- (v) Posting ads: users post ads to the public screen, writing the ad information from the mobile device.

FIGURE 7: User pad control and ads post.

The user can use the smartphone's camera or photo gallery to upload a product picture (Figure 7).

6. Experimental Results

We design an experimental framework based on the work of Herlocker et al. [22] to define important considerations during the tests. Specifically, the purpose of the experiments was to evaluate the following hypotheses:

- (H1) The inclusion of multiscreen cooperation mechanisms improves the precision/novelty balance during the recommendation process.
- (H2) The inclusion of trust information in the collaborative filtering algorithm improves the precision/novelty balance during the recommendation process.

6.1. Domain Considerations. According to the previous description, we tested our recommendation proposal in a pervasive advertising domain, specifically using a digital signage prototype. Thinking about pervasive advertising features, some important aspects about the domain were considered: the main task for the RS will be to find “some good items”; not all good items are required for advertising purposes. Utility maximization for a user is related not only to good recommendation accuracy but also to some extent of novelty. Recommendation novelty may have higher priority over the first one for advertising purposes. In this sense, although the false positive rate may be considered spam advertising, some of these ads may be novel, so a balance between false positive and novel ads is desirable. Additionally, the false negative rate is particularly relevant for advertisers, as long as they want the items they consider relevant to become recommended ads. Finally, according to the analysis of the advertising domain presented in Introduction of this paper, context information in this case was represented by the inclusion of “trust” in the recommendation algorithm, according to the procedures described in the previous sections.

FIGURE 8: Trust graph for test group.

6.2. Dataset Considerations. Because our recommender system proposal is based on trust, we required two datasets: one with ratings and the other one with trust scores. One of the main challenges was the lack of a suitable ratings' dataset for the advertising domain, so we decided to build one. We built a web application where students at the University of Cauca could post and rate ads; then, a set of users participated in interactive sessions using the implemented prototype described previously to improve the dataset information. The first interactions were useful to test the collaborative filtering and the aggregation techniques for the advertising domain. The results were published in [51, 52]. At the end, we completed a dataset with 127 ads, 176 users, and 10.128 ratings for the tests.

For the dataset including trust information, the main challenge was to find a homogeneous group of people where each member had at least one connection in Facebook with another member of the group. For the offline tests, we decided to generate this information in a hard-code way based on the users of the rating dataset because our main interest was to see the effect of trust inclusion in the recommendation process; ground truth tests were performed previously for the trust inference algorithm itself, as described in Section 3. At the end, we obtained a dataset with 30852 trust connections, as shown in Figure 8. The graph shows the connections between users, so each pair of connected users shares a trust value between 0 and 1. The green connections represent higher values of trust as long as red connections represent lower trust values. Regarding the users in the graph, the color and size are given by the node degree; green ones have a higher degree, and red ones have a lower degree.

For the online tests, the task was more challenging because of the requirement of connections between people, as described previously. Several volunteers groups were tested to meet this criterion. Finally, twenty volunteers from Fundación InnoGen, in Popayán, met these requirements, and they agreed to participate in the interactive session during the experiment. We used the Graph API Explorer from Facebook to obtain information from the social network with

FIGURE 9: Fundación InnoGen social graph.

the authorization of each user. All data were anonymous, so only IDs were used for all users, and only the number of interactions was processed without intervention over the message content. Once the interaction information was complete, we computed the trust score between users using the process described in Section 3.

Figure 9 shows the graph representing the communities inside the group. According to the graph, 3 communities were detected: the green community, which is the largest and has the nodes with higher degrees, and the yellow one and the red one, which have nodes with lower degrees. This information allowed us to set up the groups during the interactive experience with the recommender system implemented in the prototype. Due to the high degrees in the green community, some of its members participated more than once during the experiment.

6.3. Precision and Novelty Metrics. As a rule, precision has been the most popular metric for a recommender system evaluation, and it is defined as the ratio of relevant items selected to number of items selected or the probability that a selected item is relevant:

$$P = \frac{N_r}{N}. \quad (10)$$

However, for the purposes of the current research and according to the advertising domain considerations described previously, it is more interesting to evaluate the recommender system from a precision/novelty perspective. Although several approaches measure the recommendations novelty in the literature, Vargass and Castells [5] proposed an interesting

FIGURE 10: Novelty model basic principles.

framework for the definition of novelty and diversity metrics that unifies several state-of-the-art metrics. Specifically, this framework supports metrics taking into account the ranking and relevance of the recommended items.

These properties are important for the current research because the interactive nature of our proposal for the specific domain (pervasive advertising), ranking, and relevance take into account how users interact with recommendations (top items receive more attention in each screen list) and user subjectivity (how relevant the item may be for the user).

According to the authors, the novelty of a piece of information generally refers to “how different it is with respect to what has been previously seen” by a user. Nonetheless, in the advertising domain, the effectiveness of ads will be measured not only by “how different they are from previous ads” but also by “how relevant actually they could be for the users,” so the novelty metric should include the influence of these properties during the recommendation process.

In simple terms, the framework is based on three fundamental relationships between users and items (Figure 10): (i) *discovery*, that is, an item is seen by (or is familiar to) a user; (ii) *choice*, that is, an item is used, picked, selected, or consumed by a user; and (iii) *relevance*, that is, an item is liked by, useful to, or enjoyed by a user. To simplify the model, the authors assume that relevant items are always chosen if they are seen, irrelevant items are never chosen, and items are discovered independently of their relevance.

In terms of probability distribution, these relations are expressed as

$$p(\text{choose}) \sim p(\text{seen}) p(\text{rel}). \quad (11)$$

Given a ranked list R of items recommended to user u , the novelty can be expressed as

$$m(R | \theta) = C \sum_{i \in R} p(\text{choose} | i, u, R) \text{nov}(i | \theta), \quad (12)$$

where C is a normalizing constant and θ is a generic context variable to consider different perspectives in the novelty definition. $p(\text{choose} | i, u, R)$ reflects the browsing model grounded on item choice and $\text{nov}(i | \theta)$ an item novelty model. For the purposes of the current research, we used a popularity based item novelty where high novelty values correspond to long-tail items that few users have interacted with and low novelty values correspond to popular top items, including ranking and relevance factors. According to the domain features, this approach makes sense not only for

users but also for advertisers: frequently, advertisers would like to promote new products, and they will likely be long-tail items at first for the recommender system. Therefore, a metric of how well the recommender system behaves in these cases could be useful for advertisers. It is expected that this behavior improves with the inclusion of a multiscreen approach according to our hypothesis.

In summary, we describe the equations we defined for the novelty metric according to the item popularity based model:

$$\begin{aligned} \text{nov}(R | u) \\ = C \sum_{i_k \in R} \text{disc}(k) p(\text{rel} | i_k, u) (1 - p(\text{seen} | i_k)). \end{aligned} \quad (13)$$

In this case, the novelty metric can be read as the expected number of seen relevant recommended items that were not previously seen. The equation includes a ranking component $\text{disc}(k)$ that defines a logarithmic decrease according to the item position in the list:

$$\text{disc}(k) = \frac{1}{\log_2(k+1)}. \quad (14)$$

It also includes a relevance component $p(\text{rel} | i_k, u)$ that can be modeled as a heuristic mapping between rating values and probability of relevance, according to the following function:

$$\begin{aligned} p(\text{rel} | i, u) &\sim \frac{2^{g(u,i)} - 1}{2^{g_{\max}}}, \\ g(u, i) &= \max(0, r(u, i) - \tau). \end{aligned} \quad (15)$$

g is a utility function derived from the ratings, where τ represents the indifference rating value according to Breese et al. [53].

6.4. Offline Tests. First, we tested the trust-based recommender system using the dataset built from scratch, including ratings and trust scores, as described in Section 6.2. The tests were related to hypothesis (H2) and were focused on precision and novelty calculations.

To observe the effect of trust on the precision and novelty of the recommendations, we conducted several tests for different values for the contribution of the trust and similarity components of (9). The tests ranged from a classic collaborative filtering ($\alpha = 1, \beta = 0$) to pure trust-based collaborative filtering ($\alpha = 0, \beta = 1$). Table 3 summarizes the test values for similarity and trust weights.

Figure 11 shows the precision results for the different tests. It is interesting to observe the poor performance of the recommendation algorithm when a pure trust-based collaborative filtering takes place; this finding is coherent with previous findings of other researchers [11, 54]. Therefore, a similarity component should be included in the recommendation algorithm to improve the precision results when a trust score influences the recommendation process.

Although it seems that the trust component does not have a meaningful influence on the precision of the recommendation algorithm compared to a classic collaborative filtering

TABLE 3: Similarity and trust weights for the different tests.

Test number	Trust	Similarity
1	100%	0%
2	90%	10%
3	80%	20%
4	70%	30%
5	60%	40%
6	50%	50%
7	40%	60%
8	30%	70%
9	20%	80%
10	10%	90%
11	0%	100%

FIGURE 11: Recommender system precision results for different similarity-trust values.

approach, it is worth studying the algorithm behavior from a novelty perspective. In this case, we selected 50 random users from the dataset, and we calculated the novelty value, including rank and relevance factors, as described in Section 6.3. According to the previous findings, we conducted two tests for a classic collaborative filtering recommender system (test 11; $\alpha = 1, \beta = 0$) and a trust-based recommender system (test 6; $\alpha = 0.5, \beta = 0.5$). Table 4 summarizes the results.

The mean value suggests that the trust-based recommender system performed better than a classic collaborative filtering from a novelty perspective. We conducted a t -test to determine whether this difference is statistically significant; previously, considering the sample size we ran a Kolmogorov-Smirnov test to check the normal distribution of data getting a p value of 0.19 for the data associated with the classic algorithm and a p value of 0.09 for the data associated with the algorithm including trust, so the data follow a normal distribution in both cases. The t -test delivered a p value < 0.001 , supporting hypothesis (H2). This finding indicates that trust inclusion in the recommendation algorithm improves the precision/novelty balance. However, it is important to contrast these results with the real users' perceptions. This process will be described in the next section.

TABLE 4: Novelty results for the recommender system including and excluding trust.

Test	Mean	N	Standard deviation
Traditional recommender system	0.18	50	0.058
Trust-based recommender system	0.23	50	0.070

TABLE 5: Session configuration for online tests.

Group	Smart TV	Smartphone	
	Aggregation technique	RS algorithm	
Odd	Least Misery (LM)	S1	$\alpha = 1, \beta = 0$
		S2	$\alpha = 0.5, \beta = 0.5$
Even	Most Pleasure (MP)	S1	$\alpha = 1, \beta = 0$
		S2	$\alpha = 0.5, \beta = 0.5$

6.5. Online Tests. The tests with real users used the infrastructure of the prototype described in Section 5 for the pervasive advertising domain, specifically in a digital signage environment. The purpose of this set of tests was related to hypotheses (H1) and (H2). As described in Section 6.2, one of the main challenges was to find a homogeneous group with enough connections to enable a correct trust inference from the social network. Twenty volunteers from the Fundación InnovaGen group met these requirements to perform the test in two sessions: During the first one, the recommender system used a classic collaborative filtering approach ($\alpha = 1, \beta = 0$) to deliver individual recommendations to personal devices (smartphones) as long as, during the second session, the recommender system included a trust and similarity component with equal weights ($\alpha = 0.5, \beta = 0.5$).

A total of 8 groups of 3 people were configured to participate in the experiment according to the social connections graph information. Some of them repeated the experiment because of their relationships with other members of the group. Regarding the group recommendations displayed on the Smart TV, we alternated two aggregation techniques between the groups to test the effect of each one; previous experiments showed a better performance for the chosen techniques [51]. We used the Least Misery aggregation technique for the odd groups (1, 3, 5, and 7); this technique makes a list of ratings with the minimum individual ratings, and the items are recommended based on the rating on that list. Higher rankings indicate less misery, so the group is as happy as its least happy member. We used the Most Pleasure technique for the even groups (2, 4, 6, and 8); in this case, a list of ratings is made with the maximum of the individual ratings. The items are recommended based on the rating on that list, and the higher ranking indicates more pleasure. Table 5 summarizes the tested algorithm variants for each participant group.

During each session, the users interacted with the system for five to ten minutes. As part of the interaction, the users

TABLE 6: Mapping between user actions and recommender system scale.

User interaction	RS scale
Explicit like	5
Add to favorites	4
Request add information	3
Ignored item from screen	2
Explicit dislike	1

FIGURE 12: Precision gap ($\{\alpha = 1, \beta = 0\} - \{\alpha = 0.5, \beta = 0.5\}$).

performed several actions, such as browsing the ads (on the Smart TV and smartphone screen), rating ads (on the Smart TV and smartphone screen), and detailing a group ad displayed in the Smart TV in the smartphone screen or adding an ad to their favorites list. For the explicit ad-rating task, we used a binary scale (like, no like) because it makes more sense than the classic 1 to 5 stars in the user context. This interaction introduces a novel mechanism to capture implicit ratings for the system throughout a mapping between the user actions and a classic scale from 1 to 5, used during the experiment (Table 6).

At the end of each session, we also captured the users' perception through a short survey in the smartphone of each user to complete our analysis from a qualitative perspective (see Appendix B).

During the experiment, we used Smart TV-smartphone middleware capabilities to capture logs in JSON format for the activity of the whole sessions, and, then, we processed the data. Next, we will introduce the most important results of this experiment.

6.5.1. Trust Influence over Recommendations Precision. Figure 12 shows the gap between the precision values during the experiment sessions for each group ($\{\alpha = 1, \beta = 0\} - \{\alpha = 0.5, \beta = 0.5\}$). The results are coherent with the trend observed during the offline experiments. This difference is always positive; that is, a higher weight for similarity component (α) will increase the precision value. The observed variation for the precision value was between 0.03 and 0.09 (p value < 0.001). According to the results, a positive effect over novelty was expected and it will be analyzed in the next section.

TABLE 7: Trust influence over recommendations novelty.

Session	Mean	Std. dev.	Std. error mean
(S1) $\alpha = 1, \beta = 0$	0.6247	0.0814	0.0192
(S2) $\alpha = 0.5, \beta = 0.5$	0.7072	0.0799	0.0188
<i>p</i> value			0.011

TABLE 8: Multiscreen influence over recommendations novelty.

Recommendation strategy	Mean	Std. dev.	Max.	Min.
Trust-based collaborative filtering (smartphone)	0.3922	0.1499	0.2456	0.7982
Aggregation technique (smart TV)	0.4562	0.1189	0.2997	0.6840
<i>p</i> value				0.02

6.5.2. Trust Influence over Recommendation Novelty. Table 7 shows a statistical descriptive analysis of the effect of trust introduction in the recommendation algorithm on novelty. The novelty metric included the rank and relevance components, and it was calculated for each session to compare them.

According to the results, novelty is positively affected when the trust component is included in the recommendation algorithm, which is coherent with the results observed during the offline tests. Therefore, there is good evidence to support hypothesis (H2).

6.5.3. Multiscreen Influence over Recommendations Novelty. Table 8 shows a statistical descriptive analysis of the effect of the multiscreen approach on novelty. In this case, we calculate the novelty value for the group recommendations in the Smart TV and the personalized recommendations in the smartphone introducing the trust component ($\alpha = 0.5$, $\beta = 0.5$). The novelty metric includes rank and relevance components.

According to the results, the novelty value is larger for the recommendations delivered in the Smart TV for the group profile. Therefore, there is good evidence to support hypothesis (H1). However, we conducted a deeper analysis taking into account the trust inclusion and the real users' perception. Figure 13 shows the results for the number of cases where the group novelty is higher than the individual novelty including and excluding the trust component for different variants of the novelty metric. The trend is coherent with the previous results because the novelty for group recommendations is higher than the novelty for personalized recommendations in more than 50% of cases, with or without the inclusion of the trust component. However, there is an interesting finding: when the trust component is included, the number of cases where the group recommendations novelty is higher is reduced. A proper justification of this behavior may be analyzed in the user novelty perception results shown in Figure 14. According to the survey results, it is interesting to observe how the trust component inclusion in the recommendation algorithm improves the novelty perception for the personalized recommendations delivered in the smartphone.

FIGURE 13: Cases where group recommendation novelty is greater than personalized recommendation novelty.

FIGURE 14: Users' novelty perception when trust component is included.

It explains the previous behavior, and it is coherent with the findings described in Section 6.5.2. In conclusion, although the trust component has a positive effect on the novelty value for personalized recommendations, the multiscreen approach favors the novelty perception regarding group recommendations. This evidence also supports hypotheses (H1) and (H2).

6.5.4. Aggregation Technique Influence over Recommendation Novelty. Initially, we considered the survey results to compare the user perception with the quantitative analysis during the interaction. Figure 15 shows the novelty trend for the users' perception using a specific aggregation technique during both sessions. The results are coherent with the findings described in the previous section for the multiscreen effect on the recommendation novelty for both techniques.

FIGURE 15: Users' novelty perception per aggregation technique.

FIGURE 16: Novelty per aggregation technique (group novelty value > individual novelty value).

However, the MP technique seems to exhibit better behavior than the LM technique, specifically during the second session (trust 50%), but because the statistical evidence we found was not enough ($p = 0.558$), we went a step further with the quantitative analysis.

Table 9 shows a statistical descriptive analysis of the effect of the aggregation technique change on the group recommendation novelty. The novelty metric included the rank and relevance components, and it was calculated for the sessions where the Least Misery or Most Pleasure techniques were used.

According to the results, there is evidence that the aggregation technique affects the novelty value for the group recommendations, with a slight positive effect for the Most Pleasure technique.

TABLE 9: Aggregation technique influence over recommendation novelty.

Aggregation technique	Mean	Std. dev.	Std. error mean
LM	0.4046	0.1133	0.0327
MP	0.5076	0.1046	0.0302
<i>p</i> value			0.03

Figure 16 shows the trend when the two aggregation techniques are compared taking into account the number of cases where the group recommendation novelty value overcomes the personalized recommendation novelty value. As expected, the MP technique exhibits a better behavior than LM including the ranking and relevance factors in the novelty

metric, but the effect is less evident when the trust component is enabled in the recommendation algorithm because the novelty value increases for personalized recommendation, according to the analysis of Section 6.5.2.

7. Conclusions

In this work, we developed a proposal for a collaborative filtering recommender system based on trust. The trust component was included in the recommendation algorithm throughout two phases: (i) an algorithm calculated a trust score from social network interaction information (Facebook was used for this purpose); specifically, we defined four variables from the state-of-the-art and ground truth tests for the algorithm: inbox messages, tags to user, comments, and wall posts. (ii) Then, the recommendation algorithm calculated the recommendations based on the similarity and trust components between users. The algorithm was designed to modify the similarity and trust contributions, which eased the algorithm calibration for specific practical requirements.

We studied the algorithm behavior in the pervasive advertising domain, specifically in a digital signage prototype. This domain study let us think about the recommender system not only from the algorithms perspective itself but also from the recommendations display strategy. A multiscreen advertising prototype was designed and implemented to evaluate the proposed recommender framework using custom datasets and real users' perception. Traditionally, recommender systems are evaluated from a precision perspective, but the recommendations novelty is a relevant aspect, specifically for the advertising domain. We used the model proposed in [5] to define a novelty metric based on an item popularity approach that takes into account ranking and relevance factors during the recommendation process; it does make sense because the novelty is related not to recommending unknown items but to recommending novel and potentially useful items.

We evaluated our framework from a precision/novelty perspective with interesting findings: we found during the offline and online tests that the trust component introduction improved the novelty for the recommendations. However, a pure trust-based algorithm affected the precision adversely, so it is advisable to do a combination of similarity and trust components to keep a better precision/novelty balance. We also found that a multiscreen approach using aggregation techniques to generate group recommendations improved the precision/novelty balance for the whole system; we obtained higher novelty value for Smart TV recommendations in more than 50% of the cases during the online experiments. Additionally, we found that this trend remained unalterable when the trust component was introduced; nonetheless, the number of cases where the Smart TV recommendations' novelty value was higher decreased because the users perceived an increase in the novelty of the smartphone recommendations, due to the trust component effect in the recommendation algorithm.

Finally, we demonstrated that the aggregation technique influenced the novelty for group recommendations; in this case, the Most Pleasure technique exhibited better behavior than the Least Misery technique, but the effect was less

evident when the trust component was introduced due to the novelty increase in the personalized recommendations displayed in the smartphone.

8. Future Work

Next, we will address the limitations found during the study that can be used as a starting point for future work. First, the offline tests were limited by the absence of suitable datasets for the advertising domain but also Facebook datasets with the right information to infer trust according to our algorithm features. The dataset we built for ad information may be used as a starting point for gathering more information and for building a robust dataset with test purposes. It could be complemented with Facebook dataset information that enables a trust score calculation. We used a simulation for the trust score during the offline tests.

In practice, the operation of our proposed algorithm requires taking into account important considerations: first, the trust inference depends on the availability of sufficient interaction information. For now, it implies having a homogeneous group with enough social activity between participants, which may be challenging in digital signage environments where people may join a group in an ad hoc way. For online experiments, we conducted a previous study of the social activity between the members of the target group to get accurate results, but it is a feature to improve in practice. Additionally, the frequent changes to the Facebook API impose several restrictions to implement trust inference algorithms from the social network information in real time. Our approach was to precalculate the trust scores for the known group and store them in a database to avoid a permanent connection to Facebook, but, again, it could be problematic for ad hoc digital signage environments.

Finally, we restricted the number of participants to four people during the experimentation of our multiscreen recommender system approach because of usability issues. The number of participants depends on several factors, such as the main screen size, the hardware capabilities to support the middleware interaction information flow, and the recommender system operation itself. Although we reached some conclusions as a starting point in our previous work [50], these aspects should be carefully analyzed during a practical implementation of the system.

Appendix

A. Users Perception about Trust Inference Variables

We conducted a survey with 57 anonymous volunteers from the SmartSoft Play company (we used an intranet web page and all of them were familiar with technology), and we asked them to select which variables better represented trust in another person on Facebook. We simplified the subset to 5 variables. (All tag-related variables were represented simply as tags.) The results are summarized in Figure 17 for a total of 145 votes for the 5 variables.

FIGURE 17: Trust inference set of variables validation.

B. Users Perception Survey about Recommender System Precision and Novelty

It is shown in the following part.

Users Perception Survey

Session

- (1) What screen offered you recommendations closer to your personal preferences?
 - (i) Smart TV
 - (ii) Smartphone
- (2) What screen offered you more novel recommendations (likely unknown but interesting for you)?
 - (i) Smart TV
 - (ii) Smartphone.

Competing Interests

The authors declare that they have no competing interests.

Acknowledgments

This work was supported by the University of Cauca through Project VRI 3593 “SMARTA: Modelo para el despliegue de publicidad en entornos de computación ubicua soportado en un esquema de cooperación Smart TV-Smartphone.” Juan Camilo Ospina is funded by the Clúster CreaTIC and Colciencias young researcher program through the project “ScorPICUS: Sistema de recomendaciones para entornos de publicidad ubicua apoyado en información contextual y redes sociales.” Francisco Martinez is funded by Colciencias Doctoral Scholarship no. 567. Part of this work was conducted at Carlos III University of Madrid, Spain, where Francisco Martinez and Juan Camilo Ospina were visiting scholars in 2014 and 2015, respectively. Special thanks are due to Fundación InnoGen, SmartSoft Play, and University of Cauca volunteers for their valuable support during the experiments.

References

- [1] J. Müller, F. Alt, and D. Michelis, *Pervasive Advertising*, Human-Computer Interaction Series, Springer, London, UK, 2011.
- [2] C. Bauer and S. Spiekermann, “Conceptualizing context for pervasive advertising,” in *Pervasive Advertising*, J. Müller, F. Alt, and D. Michelis, Eds., pp. 159–183, Springer, London, UK, 2011.
- [3] U. Stalder, “Digital out-of-home media: means and effects of digital media in public space,” in *Pervasive Advertising*, J. Müller, F. Alt, and D. Michelis, Eds., pp. 31–56, Springer, London, 2011.
- [4] T. Mahmood and F. Ricci, “Improving recommender systems with adaptive conversational strategies,” in *Proceedings of the 20th ACM Conference on Hypertext and Hypermedia (HT '09)*, pp. 73–82, ACM, Turin, Italy, July 2009.
- [5] S. Vargas and P. Castells, “Rank and relevance in novelty and diversity metrics for recommender systems,” in *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys '11)*, pp. 109–116, Chicago, Ill, USA, October 2011.
- [6] J. L. Herlocker, J. A. Konstan, A. Borchers, and J. Riedl, “An algorithmic framework for performing collaborative filtering,” in *Proceedings of the 22nd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 230–237, Berkeley, Calif, USA, August 1999.
- [7] Q. Li and B. M. Kim, “An approach for combining content-based and collaborative filters,” in *Proceedings of the 6th International Workshop on Information Retrieval with Asian Languages (AsianIR '03)*, pp. 17–24, Stroudsburg, Pa, USA, 2003.
- [8] G. Adomavicius and A. Tuzhilin, “Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 17, no. 6, pp. 734–749, 2005.
- [9] B. De Carolis, “Adapting news and advertisements to groups,” in *Pervasive Advertising*, J. Müller, F. Alt, and D. Michelis, Eds., Human-Computer Interaction Series, chapter 11, pp. 227–246, Springer, London, UK, 2011.
- [10] P. Massa and P. Avesani, “Trust-aware recommender systems,” in *Proceedings of the ACM Conference on Recommender Systems (RecSys '07)*, Minneapolis, Minn, USA, October 2007.
- [11] J. Golbeck, “Generating predictive movie recommendations from trust in social networks,” in *Trust Management: 4th International Conference, iTrust 2006, Pisa, Italy, May 16–19, 2006. Proceedings*, vol. 3986 of *Lecture Notes in Computer Science*, pp. 93–104, Springer, Berlin, Germany, 2006.
- [12] P. Avesani, P. Massa, and R. Tiella, “Moleskiing. it: a trust-aware recommender system for ski mountaineering,” *International Journal for Infonomics*, vol. 20, 2005.
- [13] P. Victor, C. Cornelis, M. De Cock, and A. M. Teredesai, “Key figure impact in trust-enhanced recommender systems,” *AI Communications*, vol. 21, no. 2-3, pp. 127–143, 2008.
- [14] J. O'Donovan and B. Smyth, “Trust in recommender systems,” in *Proceedings of the 10th International Conference on Intelligent User Interfaces (IUI '05)*, pp. 167–174, San Diego, Calif, USA, January 2005.
- [15] M. I. Martín-Vicente, A. Gil-Solla, M. Ramos-Cabrer, Y. Blanco-Fernández, and M. López-Nores, “Semantic inference of user's reputation and expertise to improve collaborative recommendations,” *Expert Systems with Applications*, vol. 39, no. 9, pp. 8248–8258, 2012.
- [16] W. Chen and S. Fong, “Social network collaborative filtering framework and online trust factors: a case study on facebook,” in *Proceedings of the 5th International Conference on Digital*

- Information Management (ICDIM '10)*, pp. 266–273, IEEE, Thunder Bay, Canada, July 2010.
- [17] E. Bakshy, D. Eckles, R. Yan, and I. Rosenn, “Social influence in social advertising: evidence from field experiments,” in *Proceedings of the 13th ACM Conference on Electronic Commerce (EC '12)*, pp. 146–161, ACM, Valencia, Spain, June 2012.
 - [18] E. Gilbert and K. Karahalios, “Predicting tie strength with social media,” in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '09)*, pp. 211–220, Boston, Mass, USA, April 2009.
 - [19] I. Kahanda and J. Neville, “Using transactional information to predict link strength in online social networks,” in *Proceedings of the 3rd International AIAA Conference on Weblogs and Social Media (ICWSM '09)*, San Jose, Calif, USA, 2009.
 - [20] K. Panovich, R. Miller, and D. Karger, “Tie strength in question & answer on social network sites,” in *Proceedings of the ACM 2012 Conference on Computer Supported Cooperative Work (CSCW '12)*, pp. 1057–1066, Seattle, Wash, USA, February 2012.
 - [21] X. Li, *Interpersonal trust measurements from social interactions in Facebook [Ph.D. thesis]*, Montana State University, Bozeman College of Engineering, Bozeman, Mont, USA, 2014.
 - [22] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, “Evaluating collaborative filtering recommender systems,” *ACM Transactions on Information Systems*, vol. 22, no. 1, pp. 5–53, 2004.
 - [23] A. Gunawardana and G. Shani, “A survey of accuracy evaluation metrics of recommendation tasks,” *The Journal of Machine Learning Research*, vol. 10, pp. 2935–2962, 2009.
 - [24] S. M. McNee, J. Riedl, and J. A. Konstan, “Being accurate is not enough: how accuracy metrics have hurt recommender systems,” in *Proceedings of the Conference on Human Factors in Computing Systems (CHI EA '06)*, pp. 1097–1101, Montreal, Canada, April 2006.
 - [25] M. Ge, C. Delgado-Battenfeld, and D. Jannach, “Beyond accuracy: evaluating recommender systems by coverage and serendipity,” in *Proceedings of the 4th ACM Conference on Recommender Systems (RecSys '10)*, pp. 257–260, Barcelona, Spain, September 2010.
 - [26] N. Kawamae, “Serendipitous recommendations via innovators,” in *Proceedings of the 33rd Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '10)*, pp. 218–225, Geneva, Switzerland, July 2010.
 - [27] A. Jameson, “More than the sum of its members: challenges for group recommender systems,” in *Proceedings of the Working Conference on Advanced Visual Interfaces*, pp. 48–54, Gallipoli, Italy, May 2004.
 - [28] J. Masthoff, “Group modeling: selecting a sequence of television items to suit a group of viewers,” *User Modelling and User-Adapted Interaction*, vol. 14, no. 1, pp. 37–85, 2004.
 - [29] M. O'Connor, D. Cosley, J. A. Konstan, and J. Riedl, “PolyLens: a recommender system for groups of users,” in *Proceedings of the 7th European Conference on Computer Supported Cooperative Work (ECSCW '01)*, W. Prinz, M. Jarke, Y. Rogers, K. Schmidt, and V. Wulf, Eds., pp. 199–218, Springer, Rotterdam, Netherlands, 2002.
 - [30] J. F. McCarthy and T. D. Anagnost, “MUSICFX: an arbiter of group preferences for computer supported collaborative workouts,” in *Proceedings of the 7th ACM Conference on Computer Supported Cooperative Work (CSCW '98)*, pp. 363–372, November 1998.
 - [31] K. Kabassi, “Personalizing recommendations for tourists,” *Telematics and Informatics*, vol. 27, no. 1, pp. 51–66, 2010.
 - [32] P. Holleis, G. Broll, and S. Böhm, “Advertising with NFC,” in *Proceedings of the Workshop on Pervasive Advertising and Shopping, in Conjunction with the 8th International Conference on Pervasive Computing (Pervasive '10)*, Helsinki, Finland, 2010.
 - [33] R. José, N. Otero, S. Izadi, and R. Harper, “Instant places: using bluetooth for situated interaction in public displays,” *IEEE Pervasive Computing*, vol. 7, no. 4, pp. 52–57, 2008.
 - [34] H. Mahato, D. Kern, P. Holleis, and A. Schmidt, “Implicit personalization of public environments using bluetooth,” in *Proceedings of the CHI '08 Extended Abstracts on Human Factors in Computing Systems (CHI EA '08)*, pp. 3093–3098, Florence, Italy, April 2008.
 - [35] N. Otero and R. José, “Worth and human values at the centre of designing situated digital public displays,” *International Journal of Advanced Pervasive and Ubiquitous Computing*, vol. 1, no. 4, pp. 1–34, 2009.
 - [36] P. Santos, F. R. Ribeiro, and J. Metrôlho, “Using pervasive computing technologies to deliver personal and public ads in public spaces,” in *Proceedings of the 7th Iberian Conference on Information Systems and Technologies (CISTI '12)*, pp. 1–6, Madrid, Spain, June 2012.
 - [37] D. Michelis and H. Send, “Engaging passers-by with interactive screens—a marketing perspective,” in *Jahrestagung der Gesellschaft für Informatik (GI '09)*, pp. 3875–3881, 2009.
 - [38] J. C. Cardoso and R. José, “A framework for context-aware adaptation in public displays,” in *Proceedings of the Federated International Workshops and Posters on On the Move to Meaningful Internet Systems: ADI, CAMS, EI2N, ISDE, IWSSA, MONET, OnToContent, ODIS, ORM, OTM Academy, SWWS, SEMELS, Beyond SAWSDL, and COMBEK 2009*, pp. 118–127, Vilamoura, Portugal, 2009.
 - [39] M. McPherson, L. Smith-Lovin, and J. M. Cook, “Birds of a feather: homophily in social networks,” *Annual Review of Sociology*, vol. 27, no. 1, pp. 415–444, 2001.
 - [40] J. J. Jones, J. E. Settle, R. M. Bond, C. J. Fariss, C. Marlow, and J. H. Fowler, “Inferring tie strength from online directed behavior,” *PLoS ONE*, vol. 8, no. 1, Article ID e52168, 2013.
 - [41] W. Chen and S. Fong, “Social network collaborative filtering framework and online trust factors: a case study on facebook,” in *Proceedings of the 5th International Conference on Digital Information Management (ICDIM '10)*, pp. 266–273, Ontario, Canada, July 2010.
 - [42] V. Podobnik, D. Striga, A. Jandras, and I. Lovrek, “How to calculate trust between social network users?” in *Proceedings of the 20th International Conference on Software, Telecommunications and Computer Networks (SoftCOM '12)*, pp. 1–6, Split, Croatia, September 2012.
 - [43] M. D. Ekstrand, J. T. Riedl, and J. A. Konstan, “Collaborative filtering recommender systems,” *Foundations and Trends in Human-Computer Interaction*, vol. 4, no. 2, pp. 81–173, 2010.
 - [44] K. Swearingen and R. Sinha, “Beyond algorithms: an HCI perspective on recommender systems,” in *Proceedings of the ACM SIGIR 2001 Workshop on Recommender Systems*, vol. 13, pp. 1–11, 2001.
 - [45] R. R. Sinha and K. Swearingen, “Comparing recommendations made by online systems and friends,” in *Proceedings of the DELOS-NSF Workshop on Personalization and Recommender Systems in Digital Libraries*, vol. 106, Dublin, Ireland, 2001.
 - [46] J. Golbeck, “Generating predictive movie recommendations from trust in social networks,” in *Trust Management*, K. Stølen,

- W. H. Winsborough, F. Martinelli, and F. Massacci, Eds., vol. 3986 of *Lecture Notes in Computer Science*, pp. 93–104, Springer, Berlin, Germany, 2006.
- [47] P. Massa and P. Avesani, “Trust-aware recommender systems,” in *Proceedings of the 2007 ACM Conference on Recommender Systems*, pp. 17–24, ACM, New York, NY, USA, 2007.
 - [48] P. Victor, M. De Cock, and C. Cornelis, “Trust and recommendations,” in *Recommender Systems Handbook*, F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, Eds., pp. 645–675, Springer, Berlin, Germany, 2011.
 - [49] N. Lathia, S. Hailes, and L. Capra, “The effect of correlation coefficients on communities of recommenders,” in *Proceedings of the 23rd Annual ACM Symposium on Applied Computing (SAC ’08)*, pp. 2000–2005, Cear, Brazil, March 2008.
 - [50] F. Martinez-Pabon, J. Caicedo-Guerrero, J. J. Ibarra-Samboni, G. Ramirez-Gonzalez, and D. Hernndez-Leo, “Smart TV-smartphone multiscreen interactive middleware for public displays,” *The Scientific World Journal*, vol. 2015, Article ID 534949, 14 pages, 2015.
 - [51] F. Martinez-Pabon, J. C. Ospina-Quintero, V. Garzon-Marin, A. Chantre-Astaiza, M. Muoz-Organero, and G. Ramirez-Gonzalez, “Enriching public displays ads recommendations using an individual: group cooperation model,” in *Proceedings of the 3rd ACM International Symposium on Pervasive Displays (PerDis ’14)*, pp. 186–187, ACM, Copenhagen, Denmark, June 2014.
 - [52] F. Martinez-Pabon, J. Caicedo-Guerrero, J. J. Ibarra-Samboni, G. Ramirez-Gonzalez, M. Muoz-Organero, and A. Chantre-Astaiza, “Smart TV—smartphone cooperation model on digital signage environments: an implementation approach,” in *Proceedings of the 4th International Conference on Mobile Services, Resources, and Users (MOBILITY ’14)*, pp. 25–31, Paris, France, July 2014.
 - [53] J. S. Breese, D. Heckerman, and C. Kadie, “Empirical analysis of predictive algorithms for collaborative filtering,” in *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pp. 43–52, Madison, Wis, USA, July 1998.
 - [54] J. O’Donovan and B. Smyth, “Trust in recommender systems,” in *Proceedings of the 10th International Conference on Intelligent User Interfaces (IUI ’05)*, vol. 5, pp. 167–174, San Diego, Calif, USA, January 2005.