

Exploring Recommendations in Internet of Things

Lina Yao
The Uni. of Adelaide, Australia
lina@cs.adelaide.edu.au

Quan Z. Sheng
The Uni. of Adelaide, Australia
qsheng@cs.adelaide.edu.au

Anne H.H. Ngu
Texas State Uni., USA
angu@txstate.edu

Helen Ashman
UniSA, Australia
helen.ashman@unisa.edu.au

Xue Li
Uni. of Queensland, Australia
xueli@itee.uq.edu.au

ABSTRACT

With recent advances in radio-frequency identification (RFID), wireless sensor networks, and Web-based services, physical things are becoming an integral part of the emerging ubiquitous Web. In this paper, we focus on the things recommendation problem in Internet of Things (IoT). In particular, we propose a unified probabilistic based framework by fusing information across relationships between users (i.e., users' social network) and things (i.e., things correlations) to make more accurate recommendations. The proposed approach not only inherits the advantages of the matrix factorization, but also exploits the merits of social relationships and thing-thing correlations. We validate our approach based on an Internet of Things platform and the experimental results demonstrate its feasibility and effectiveness.

Categories and Subject Descriptors

H.3.5 [Information Storage and Retrieval]: Online Information Services; E.2 [Data Structures]: Graphs and Networks

Keywords

Internet of Things, Recommendation, Social Networks

1. INTRODUCTION

Recent advances in identification technologies such as radio frequency identification (RFID), wireless sensor networks, and nanotechnology make computing power available in smaller and smaller physical things. Indeed, our world is slowly evolving into an environment where everyday things such as buildings, commodities are readable, recognizable, addressable, and controllable through the Internet [1, 9]. While it is well understood that such a ubiquitous environment offers the capability of integrating both the physical world and the virtual one, which creates tremendous business opportunities such as efficient supply chains, it also presents significant challenges. With many things connected and interacting over the Internet, there is an urgent need to provide effective mechanism for search, recommendation, annotation and classification across the diverse set of things to reveal interesting patterns among them.

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.
SIGIR '14, July 6–11, 2014, Gold Coast, Queensland, Australia.
Copyright 2014 ACM 978-1-4503-2257-7/14/07 ...\$15.00.
http://dx.doi.org/10.1145/2600428.2609458.

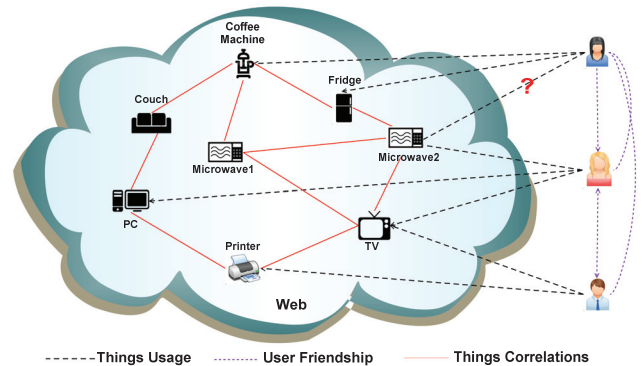


Figure 1: The connections consist of three types of links, including users' friendships within users' social network, things correlations with correlation graph of things and things usage links within user-thing interactions. Our task is to predict the dyadic relationship between users and things, i.e., recommending certain things to users.

Things recommendation is a crucial step for promoting and taking full advantage of Internet of Things (IoT), where it benefits the individuals, businesses and society on a daily basis in terms of two main aspects. On the one hand, it can deliver relevant things to users based on users' preferences and interests. On the other hand, it can also serve to optimize the time and cost of using IoT in a particular situation. Physical things in reality have multiple unique attributes. For example, they have states (e.g., in use or not in use; expired or not expired). When a certain thing is in use, it can not be used simultaneously by another user. Under this circumstance, a recommender system can refer the user to a list of things which have same or similar functionalities. For example, if microwave 1 is in use, microwave 2 will be recommended to a user who would like to warm her food, as illustrated in Figure 1.

We argue that both users' relationships and things correlations should be taken into account to make accurate recommendations. Users' relations (i.e., friendships) can have significant impact on things usage patterns. Research in [4] shows that relations between users play a substantial role in human decision making. For instance, people usually turn to a friend's advice about a commodity (i.e., hair straightener) or a restaurant before they go for them. Sometimes, this influence from the friend circle is even more substantial than high ratings given by other people. As mentioned above, things are functionality-oriented and things with similar functionalities hold strong relationships. We also argue that physical things have more distinctive structures and connections in terms of their func-

functionalities in real life (i.e., usefulness), as well as non-functionalities (i.e., availability). For example, different things provide different functionalities (e.g., microwave and printer), and will be of interest to different groups of people. Pairwise things with strong correlations indicate either they have similar functionalities (i.e., microwave and roaster) or they have more likelihood to be used together. For instance, a water tap and a chop board are both in use when we prepare meals, since most of the time we need to wash cooking ingredients (e.g., vegetables) before chopping them. In either case, related things usage patterns will be reflected.

Things recommendation in IoT can be formulated as predicting the dyadic relationships (e.g., things usage) between people and things. This prediction is based on three types of dyadic relationships, namely *people to things*, *people to people*, and *things to things*, as shown in Figure 1. A straightforward method is to utilize a series of discrete characteristics of things in order to recommend additional items with similar properties. However, this approach is not feasible in IoT since it is not easy to directly craft things' profiles and present in a meaningful feature space. Despite recent efforts in semantic techniques [1], it is still hard to characterize things due to their diverse and heterogeneous nature in terms of functionalities, access methods or descriptions. Moreover, utilizing a series of discrete characteristics does not take into account the influence of users' social networks and mutual correlations of things, nor the historical information on user interactions with things.

Collaborative Filtering (CF) recommendation has been widely used by exploiting past behaviors between users and items (i.e., rating scores on items, click-through data or browsing history etc.) without profiling users or items [7]. Several recent approaches based on CF also take into account the users' social networks [6, 8] to achieve more accurate performance. However, few work considers the impact of contextual information on mutual correlations between things.

In this paper, we propose a probabilistic matrix factorization based model to address things recommendation problem in IoT. We fuse information from users' social networks and things correlation networks, by learning shared latent factors stemming from the probabilistic matrix factorization on three matrices, namely users' relations, things correlations, and observable things usage interactions. Our main contributions are summarized as follows:

- We propose a unified CF model based on probabilistic matrix factorization, by fusing users' social networks and things correlations simultaneously. We connect these two graphs via shared latent factors and the information in users' relations. Things' correlations can be uncovered and propagated to serve user-thing usage pattern prediction.
- We validate and evaluate the model using a dataset collected from our IoT testbed. The experimental results demonstrate the feasibility and effectiveness of our approach.

The rest of this paper is organized as follows. In Section 2, we present our model on things recommendation. We show our experimental evaluation in Section 3. In Section 4, we overview the related work and provide some concluding remarks.

2. THE MODEL

Figure 2 shows our model that fuses social network, things correlations and user-thing interactions, and incorporates three relationships: user-user connections $s_{ii'}$, thing-thing correlations $t_{jj'}$ and user-thing interactions (thing usage) y_{ij} with shared factors u_i and v_j . We describe how to encode these three relationship matrices in our model from Section 2.1 to Section 2.3.

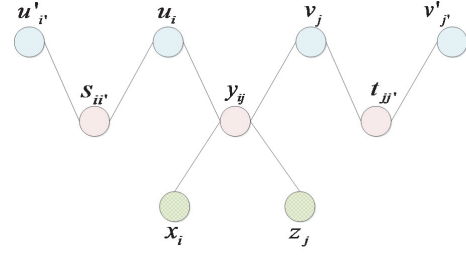


Figure 2: Our model: x_i and z_j are the explicit features (e.g., textual descriptions) of users and things respectively; $s_{ii'}$ and $t_{jj'}$ denote users friendship and things correlations, while y_{ij} denote interactions between users and things; $u_i, u_{i'}, v_j, v_{j'}$ are the latent factors induced from the three relationships.

2.1 Encoding User Friendships

We construct a directed weighted graph $\mathbf{G}_u = (V_u, E_v)$, whose vertex set V_u corresponds to users set $\{u_1, \dots, u_m\}$, edges set E_v represents the friendships between users and the range of their associated weight are in $[0, 1]$. Bigger weights represent stronger ties between users. W_u indicates the user similarity influenced by the social links between users, reflecting the *homophily* (i.e., similar users may have similar interests). We use the cosine similarity to calculate $s_{ii'}$ as follows:

$$s_{ii'} = \frac{e^{\alpha \cos(b(i), b(i'))}}{\sum_{k \in \Omega(i)} e^{\alpha \cos(b(i), b(k))}} \quad (1)$$

where $\cos(b(i), b(i')) = \frac{b(i) \cdot b(i')}{\|b(i)\| \|b(i')\|}$, $\Omega(i)$ is the set of the user i 's friends (i.e., $j \in \Omega(i)$), $b(i)$ is the binary vector of things used by user i , $\|\cdot\|$ is the L-2 norm of vector $b(\cdot)$, and α is a parameter that reflects the preference for transitioning to a user who interacts with the same things.

After we obtain the users friendship matrix from \mathbf{G}_u , we factorize users' friendship matrix to derive a high-quality, low dimensional feature representation to user-based latent features vectors $u_i \in \mathbb{R}^{1 \times m}$ and factor-based latent feature vectors $u_{i'} \in \mathbb{R}^{1 \times m}$ on analyzing the social network graph \mathbf{G}_u . The conditional probability of $s_{ii'}$ over the observed social network is determined by:

$$s_{ii'} \sim Pr(s_{ii'} | u_i^T u_{i'}; \sigma_s) \quad (2)$$

where $u_i \sim \mathcal{N}(0, \sigma_u^2)$, $u_{i'} \sim \mathcal{N}(0, \sigma_{u'}^2)$

Similar to the Web link adjacency [10], if a user i has lots of links to other users, the trust value of $s_{ii'}$ should be decreased. Whereas if a user i is trusted by many other users, the trust value of $s_{ii'}$ should be increased, since the user can be considered as local authority. $s_{ii'}$ should be adjusted as:

$$s_{ii'}^* = \sqrt{\frac{d^{in}(i')}{d^{out}(i) + d^{in}(i')}} \times s_{ii'} \quad (3)$$

where $d^{out}(i)$ represents the outdegree of node i , while $d^{in}(i')$ indicates the indegree of i' . Equation 2 can be reformulated as:

$$s_{ii'}^* \sim Pr(s_{ii'}^* | u_i^T u_{i'}; \sigma_s) \quad (4)$$

2.2 Encoding Things Correlations

A correlation graph of things is a directed weighted graph, denoted by $\mathbf{G}_t = (T, E)$. For each thing $t_i \in T$, let T_i^k denote the

top- k set of correlated things to t_i . $E = \{e(x, i) | \forall t_i \in T, t_x \in T_i^k\}$, where $e(x, i)$ is an edge from t_x to t_i . Each edge is associated with a weight w_{t_x, t_i} with the correlation value R_{t_x, t_i} .

Things correlation value R_{t_x, t_i} is not handy to obtain due to some unique challenges. To tackle this problem, we have developed a graph-based model to predict correlations [9], which is briefly described here. We derive correlations among things by mining the history of things usage events. In particular, we construct a spatio-temporal graph \mathbf{G}_m and a social graph \mathbf{G}_u to model things usage contextual information and interactive relationships between users and things. The spatio-temporal graph captures the spatial and temporal information in things usage events, i.e., where and when a certain thing is accessed. In constructing this graph, we integrate the spatial and temporal information to capture periodic patterns between locations and timestamps for improved performance. Then, we perform random walks with restart [3] on both graphs to obtain pairwise relevance R_m and R_u respectively, and sum them up to get overall pair correlation $R = (R_m + R_u)/2$, which indicates thing-thing correlations.

Similar to modeling social networks, things correlation matrix is decomposed to thing-based latent features vectors $v_j \in \mathbb{R}^{1 \times n}$ and factor-based latent feature vectors $v'_{j'} \in \mathbb{R}^{1 \times n}$ on analyzing the things correlation graph \mathbf{G}_t . The conditional probability of $t_{jj'}$ relies on things latent factors and can be denoted as:

$$t_{jj'} \sim Pr(t_{jj'} | v_j^T v'_{j'}; \sigma_r) \quad (5)$$

where $v_j \sim \mathcal{N}(0, \sigma_v^2), v'_{j'} \sim \mathcal{N}(0, \sigma_{v'}^2)$

2.3 Encoding User-Thing Interactions

User-thing interactions y_{ij} are embodied by the usage frequency of thing i by user j in a certain time frame. We can map the usage frequency to interval $[0, 1]$ by using function $f(x) = (x - y_{min}) / (y_{max} - y_{min})$ without loss of generality, where y_{max} and y_{min} are the maximum and minimum usage values respectively. The dyadic relationship between a user and a thing does not only depend on their latent factor $U^T V$, whose vulnerability is that it makes use of past interactions and can not handle brand new things well, i.e., cold start problem. To tackle this issue, we use the explicit features directly by profiling users observable features $x_i \in \mathbb{R}^c$ (i.e., age, gender, location etc.) and things observable features $z_j \in \mathbb{R}^d$ (i.e., textual description of things functionalities). Here, c and d are the dimensionality of users observable features and things observable features respectively. The dyadic relationship (thing usage value) depends on not only the inner product of latent factors of users and things, but also their observable features. Things usage value y_{ij} can be defined as the following conditional probability:

$$y_{ij} \sim Pr(y_{ij} | u_i, v_j, x_i, z_j, \sigma_y^2) \quad (6)$$

We adopt the bilinear product to specify the similarity between user observable features and thing observable features [2]. The pairwise similarity between x_i and z_j can be denoted as:

$$r_{ij} = \mathbf{w}^T (x_i \otimes z_j) \quad (7)$$

where \mathbf{w} is a column vector of entries $\{w_{mn}\}$, $x_i \otimes z_j$ denotes the Kronecker product of x_i and z_j . Equation 7 can be rewritten as:

$$r_{ij} = x_i^T W z_j \quad (8)$$

where matrix $W \in \mathbb{R}^{m \times n}$ is a weight coefficients capturing pairwise associations between user i 's explicit feature vector and thing j 's explicit feature vector. The thing usage value depends on both the inner product of user and thing latent factors and the bilinear

product of user observable features and thing observable features. Equation 6 can be reformulated as:

$$y_{ij} \sim Pr(y_{ij} | u_i^T v_j + r_{ij}, \sigma_y^2) \quad (9)$$

2.4 Model Learning

Given a training dataset for $\mathbb{O} = \{\mathbb{O}_y, \mathbb{O}_s, \mathbb{O}_t\}$, the joint posterior probability of model parameters Σ can be obtained through Bayes' theorem:

$$Pr(\Sigma | \mathbb{O}) \propto Pr(\mathbb{O} | \Sigma) Pr(\Sigma) \quad (10)$$

Maximizing Equation 10 can be converted to minimizing the negative logarithm of $Pr(\mathbb{O} | \Sigma) Pr(\Sigma)$ via:

$$\begin{aligned} \min_{\Sigma} L(\Sigma) = & \min \lambda_y \sum_{ij \in \mathbb{O}_y} \ell(y_{ij}, u_i^T v_j + r_{ij}) \\ & + \lambda_s \sum_{ii' \in \mathbb{O}_s} \ell(s_{ii'}, u_i^T u_{i'}) + \lambda_t \sum_{jj' \in \mathbb{O}_t} \ell(t_{jj'}, v_j^T v_{j'}) \\ & + \lambda_w \|\mathbf{w}\|^2 + \lambda_u \|u\|^2 + \lambda_{u'} \|u'\|^2 + \lambda_v \|v\|^2 + \lambda_{v'} \|v'\|^2 \end{aligned} \quad (11)$$

where $\ell \cdot$ is a loss function (we adopt the most widely used ℓ_2 loss), and $\lambda = \{\lambda_y, \lambda_s, \lambda_t, \lambda_w, \lambda_u, \lambda_{u'}, \lambda_v, \lambda_{v'}\}$ are trade-off parameters.

A gradient descent process can be implemented to solve the parameters. Given a training dataset $\{y_{ij}\}$, the objective function in Equation 11 can be found by performing gradient descent in u_i, v_j and w_{mn} .

$$\begin{aligned} u_i & \rightarrow u_i - \delta \times \left(\frac{\partial \ell}{\partial u_i} (y_{ij}, u_i^T v_j + r_{ij}) v_j + u_i^T u_i \right) \\ v_j & \rightarrow v_j - \delta \times \left(\frac{\partial \ell}{\partial v_j} (y_{ij}, u_i^T v_j + r_{ij}) u_i + v_j^T v_j \right) \\ w & \rightarrow w - \delta \times \left(\frac{\partial \ell}{\partial w} (y_{ij}, u_i^T v_j + r_{ij}) u_i v_j^T + w^T w \right) \\ u'_{i'} & \rightarrow u'_{i'} - \delta \times \left(\frac{\partial \ell}{\partial u'_{i'}} (s_{ii'}, u_i^T u'_{i'}) u_i + u'_{i'}^T u'_{i'} \right) \\ v'_{j'} & \rightarrow v'_{j'} - \delta \times \left(\frac{\partial \ell}{\partial v'_{j'}} (t_{jj'}, v_j^T v'_{j'}) v_j + v'_{j'}^T v'_{j'} \right) \end{aligned} \quad (12)$$

where δ is the learning rate. After we obtain the optimal parameters Σ^* , we can use them to predict the given testing data $\{\tilde{i}, \tilde{j}, y_{\tilde{i}, \tilde{j}}\}$

$$y_{\tilde{i}, \tilde{j}} = u_{\tilde{i}}^T v_{\tilde{j}} + \sum_{mn}^{cd} x_{\tilde{i}n} z_{\tilde{j}m} w_{mn}^* \quad (13)$$

3. EXPERIMENTS

We have conducted extensive experiments to study the proposed approach. Due to space constraints, we report one experiment, which compares our approach with other state-of-the-art methods.

3.1 Dataset and Performance Metrics

Since Internet of Things (IoT) is still relatively new, it is hard to find large-scale data for our experiments. We set up a testing environment that consists of several different physical places including one author's home (e.g., kitchen, bedroom, living area). 127 different things (e.g., printers, desktop computers, backpacks, TV, coffee machine, first-aid kit) are tagged using RFID tags or sensors. We also manually labeled these things with 397 labels, which serve as

the ground-truth dataset in our experiments. Ten volunteers participated in the data collection by interacting with RFID-tagged things for a period of four months, generating 20,179 records on the interactions of the tagged things. Table 1 shows the statistics of six categorical things of the dataset. Scale refers to the usage frequency between the most frequent used things and the least frequent used things. For example, in *Cooking*, one of the least used thing is the blender (48), while one of the most used thing is the fridge (2336, the door open/close times of the fridge). All scale values are processed to make them in [0,1] as discussed in Section 2.3.

Table 1: Dataset Statistics

No.	Category	# Things	Scale
1	Entertainment	28	47 - 1028
2	Office	20	511 - 1790
3	Cooking	25	48 - 2336
4	Transportation	11	31 - 870
5	Medicine/Medical	10	22 - 89
6	House Appliances	33	37 - 2680

We adopt Mean Absolute Error (MAE) to measure the accuracy of our approach. MAE calculates the average of absolute difference between predicted usage values and actual values as the following:

$$MAE = \frac{\sum_{ij} |y_{ij} - \tilde{y}_{ij}|}{n} \quad (14)$$

where y_{ij} is the actual thing's usage value between user i and thing j , \tilde{y}_{ij} is the predicted value, and n is the number of the predicted values. The lower the MAE, the better the performance.

3.2 Performance Comparison

We compare the prediction accuracy of our approach based on fusing social networks and things correlations (FST) with some state-of-the-art approaches based on probabilistic factor analysis: Probabilistic Matrix Factorization (PMF) [7], SoRec [6] and SVD++ [5]. This experiment evaluated our approach, in particular its capability in handling the cold start problem, which refers to providing accurate prediction when some users only use few things or even have no usage historical records at all. In order to verify the capability of our approach on predicting usage value of things that have not been used, we randomly selected and marked off $p\%$ of data ($p=10, 20$ and 50) from our dataset as training data and different number of latent factors (5 and 10) to test all the methods. The experimental results are shown in Table 2.

Table 2: MAE Comparison with Other Approaches

Training Data	10%		20%		50%	
# of Factors	5	10	5	10	5	10
PMF	0.8635	0.8544	0.8245	0.8153	0.7662	0.7495
SVD++	0.8425	0.8311	0.8004	0.7982	0.7410	0.7273
SoRec	0.8103	0.7978	0.7872	0.7723	0.7317	0.7168
FST	0.7303	0.7246	0.7112	0.6948	0.6631	0.6406

From the table, it is clear that our approach outperforms other methods on different training ratios and different number of factors, especially when the training data is small. PMF is a pure probabilistic factor model. Relying heavily on user-thing usage matrix, it can not deal with the circumstance where little interactions information is available. SoRec works better than PMF and SVD++ because of its aggregation of user-user internal information (social links). Our approach not only incorporates users and things internal information, but also defines the explicit features (i.e., content) for users (e.g., users profile) and things (e.g., description of things functionalities), which makes our approach performing better when there is a cold start problem. The experimental result further demonstrates the effectiveness on improving the recommendation accuracy by incorporating things correlations.

4. RELATED WORK AND CONCLUSION

Recommending the most appropriate things to users is a fundamental concern in the emergent Internet of Things research. Most recent research work focus on exploiting the information of users' connections for recommendation. Jamali and Ester [3] propose a social recommendation framework based on probabilistic matrix factorization via employing user social networks and user-item ratings. Yang et al. [8] design a joint friendship and interest propagation model, where the user-item interest network and the user-user friendship network are jointly modeled through latent user and item factors. Zhou et al. [11] propose a kernel-based probabilistic matrix factorization, which incorporating external information into the matrix factorization process via assuming latent factors are in Gaussian distribution. Ma et al. [6] propose a matrix factorization method to exploit the social network information.

Compared with these research efforts, we construct the things correlation graph capturing global connections and similarities between things, and integrate the graph with users' social relation graph to learn the latent factors simultaneously. We also integrate the content and other additional information of users and things to cope with the cold-start problem in recommendations. The experiments demonstrate the feasibility and effectiveness of our model.

In real situation, physical things are more dynamic compared with traditional Web resources. Some of these dynamic features are availability, statefulness and its changing attributes (i.e., geographical information, status). In the future work, we plan to improve our model that can adaptively propagate up-to-date information from things correlations network and make more accurate recommendations. We also plan to explore further the impact caused by the unique feature of non-duplicability in Internet of Things.

5. REFERENCES

- [1] P. Barnaghi et al. Semantics for the Internet of Things: Early Progress and Back to the Future. *Intl. Journal on Semantic Web and Information Systems*, 8(1):1–21, 2012.
- [2] W. Chu and S.-T. Park. Personalized Recommendation on Dynamic Content Using Predictive Bilinear Models. In *Proc. of the 18th WWW*, 2009.
- [3] M. Jamali and M. Ester. TrustWalker: a Random Walk Model for Combining Trust-based and Item-based Recommendation. In *Proc. of the 15th ACM SIGKDD*, 2009.
- [4] T. Kameda et al. Centrality in Sociocognitive Networks and Social Influence: An Illustration in a Group Decision Making Context. *Journal of Personality and Social Psychology*, 73(2):296, 1997.
- [5] Y. Koren. Factorization Meets the Neighborhood: a Multifaceted Collaborative Filtering Model. In *Proc. of the 14th ACM SIGKDD*, 2008.
- [6] H. Ma, H. Yang, M. Lyu, and I. King. Sorec: Social Recommendation Using Probabilistic Matrix Factorization. In *Proc. of the 17th ACM CIKM*, 2008.
- [7] R. Salakhutdinov and A. Mnih. Probabilistic Matrix Factorization. In *Proc. of the 21st NIPS*, 2007.
- [8] S.-H. Yang et al. Like Like Alike: Joint Friendship and Interest Propagation in Social Networks. In *Proc. of the 20th WWW*, 2011.
- [9] L. Yao et al. A Model for Discovering Correlations of Ubiquitous Things. In *Proc. of the 13th ICDM*, 2013.
- [10] D. Zhou et al. Semi-supervised Learning on Directed Graphs. In *Proc. of the 18th NIPS*, 2004.
- [11] T. Zhou et al. Kernelized Probabilistic Matrix Factorization: Exploiting Graphs and Side Information. In *Proc. of SDM*, 2012.