Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 98 (2016) 455 – 460

Procedia Computer Science

International Workshop on Data Mining on IoT Systems (DaMIS16)

A numerical approach for assigning a reputation to users of an IoT framework

Salvatore Cuomo[a,*], Pasquale De Michele[a], Ardelio Galletti[b], Giovanni Ponti[c]

[a]*Department of Mathematics and Applications, University of Naples Federico II, Strada Vicinale Cupa Cintia 21, 80126, Naples, Italy*
[b]*Department of Science and Technology, University of Naples Parthenope, Centro Direzionale Isola C4, 80143, Naples Italy*
[c]*DTE-ICT-HPC, ENEA Research Center, Piazzale Enrico Fermi, Portici, Naples, Italy*

Abstract

Nowadays, in the Internet of Things (IoT) society, the massive use of technological devices available to the people makes possible to collect a lot of data describing tastes, choices and behaviours related to the users of services and tools. These information can be rearranged and interpreted in order to obtain a rating (i.e., evaluation) of the subjects (i.e., users) interacting with specific objects (i.e., items). Generally, reputation systems are widely used to provide ratings to products, services, companies, digital contents and people. Here, we focus on this issue, adopting a Collaborative Reputation System (CRS) to evaluate the visitors' behaviour in a real cultural event. The results obtained, compared with those obtained by other methods (i.e., classification), have confirmed the reliability and the usefulness of CRSes for deeply understand dynamics related to visiting styles.

Peer-review under responsibility of the Program Chairs
Keywords: Internet of Things, Collaborative Reputation Systems, Cultural heritage, Classification;

1. Introduction

A challenge in the Internet of Things (IoT) society is to manage the huge amount of information coming from the interaction of users with technological devices. Nowadays, approaches based on reputation systems are analysed to assign ratings to products, services, companies, digital contents and people [1]. More in detail, the key goals of a reputation system are, firstly, to establish the trust in on-line transactions and, secondly, to detect frauds or abnormal behaviours. Real examples of systems needing a feedback mechanism to establish the trust are on-line auction houses, on-line selling companies, peer-to-peer networks, opinion websites, e-mail spam filters, search engines [2,3,4]. Since the reputation is computed from a set of assessments given by the evaluators themselves, these algorithms are commonly referred to as Collaborative Reputation Systems (CRSes). More specifically, in a CRS, a group of *items* is evaluated repeatedly over time by a group of *raters* (i.e., evaluators). Notice that the size of both items and raters may change over time. The system can achieve an overall reputation for all the involved items on the basis of evaluations given at some time step by some users, in a way that such reputation could be fruitfully adopted by other users in taking

* Corresponding author. Tel.: +39-081-675624.
E-mail address: salvatore.cuomo@unina.it

1877-0509
Peer-review under responsibility of the Program Chairs
doi:10.1016/j.procs.2016.09.073

decisions about items.

In this paper we are interested to adopt CRSes to evaluate visitors' behaviours cultural heritage scenario. More in detail, we want to deduce a reputation value for each object (i.e., artworks) in a cultural space, analysing the behaviours assumed by the users (i.e., visitors) of a cultural event. In other words, the *visitors* and *artworks* can be assimilated to raters and items, respectively. With these aims, we have resorted to the CRS model described in[5], which furnishes a generalization to the class of iterative procedures, proposed in some recent works on[1,6,7] and are known as Iterative Filtering (IF) methods. As case of study, we have considered a real art exhibition consisting of 253 sculptures (i.e., cultural objects), divided into 7 thematic sections and named "*The Beauty or the Truth*"[1]. This exhibition has shown, for the first time in Italy, the Neapolitan sculpture of the late nineteenth and early twentieth centuries, through the major sculptors of the time. The sculptures have been exposed in the monumental complex of *San Domenico Maggiore*, located in the historical centre of Naples. The experimental results highlight that, after a suitable parameter tuning in the CRS procedure, compared with those described in[8], confirm the reliability and the usefulness of this approach.

The rest of the paper is organized as follows. Section 2 recall main definition about CRSes. Section 3 is devoted to the description of the specific application context. In Section 4, we outline our main idea, i.e. we show how the information about each visitor could be used to obtain a suitable input for CRS. Finally, conclusions are reported in Section 5.

2. Preliminaries on CRSes

In this section, we recall basic concepts and details about collaborative reputation systems. The approach and notation are the same introduced in[5] A CRS is an iterative procedure that can be described through a discrete set of scoring states $k \in \mathbf{N}_0$, where $k = 0$ is the initial state and each following state $k \geq 1$ corresponds to the overall evaluation of a set of $m = m(k)$ items by means of a group of $n = n(k)$ raters. A system, with $m(k) \leq m$ and $n(k) \leq n$, for any step k, is called finite-dimensional and can always be represented by a sequence of $n \times m$ matrix triples

$$\{A(k), E(k), T(k)\}, \quad k = 0, 1, 2, \ldots$$

where $A(k) = (a_{ij}(k))$, $E(k) = (e_{ij}(k))$ and $T(k) = (t_{ij}(k))$ are the *adjacency*, *evaluation* and *trust* matrices of the system, respectively. More in detail, $a_{ij}(k)$ is 1 if rater i scores the item j at step k (0 otherwise), $e_{ij}(k)$ is the given evaluation and $t_{ij}(k)$ is a measure of the trust, or degree of reliability, of that evaluation. Trust values are used to weight the evaluations $e_{ij}(k)$ in order to get a score, or *reputation*, $r_j(k)$ for each item j, i.e. it results

$$r_j(k) = \sum_{i=1}^{n} w_{ij}(k) a_{ij}(k) e_{ij}(k) \,, \tag{1}$$

where the weights $w_{ij}(k)$ are computed coherently with the trust degree for the raters, i.e. it is

$$w_{ij}(k) = \frac{t_{ij}(k)}{\sum_{i=1}^{n} t_{ij}(k) a_{ij}(k)} \,. \tag{2}$$

Moreover, the trust values are also used to define a reputation $\rho_i(k)$ for each rater i, that is

$$\rho_i(k) = \frac{1}{m_i(k)} \sum_{j=1}^{m} t_{ij}(k) a_{ij}(k) \qquad \text{with} \qquad m_i(k) = \sum_{j=1}^{m} a_{ij}(k). \tag{3}$$

In a CRS the evolution of this trust over time is specified by means of a so-called CRS filter function Φ. In general, the filter function provides the value of the trust matrix, given the matrix of the scores at the same time step and the

[1] http://www.ilbellooilvero.it

trust matrix at the previous time step. One can express the above dependency by the iterative recurrence

$$T(k) = \Phi(A(k), E(k), T(k-1)) \quad k \geq 1 \,, \tag{4}$$

where $T(k) = T_0$ in a starting trust matrix, i.e. an $n \times m$ set of initial conditions. Notice that the dependence of $T(k)$ on $T(k-1)$ (and on T_0) implies that a CRS can use an initial set of trusted users (in T_0) to filter out unfair users at following steps. Eq. (4) states that, at any scoring step $k \in \mathbf{N}$, a reputation of the raters can be computed through their evaluations at the current step and their degrees of reliability at the previous step. This poses the problem of determining functions Φ that meaningfully convey the transmission of trust. A way to construct several CRSes of that kind (that are able to propagate trustiness and exhibit effective filtering properties, at least for a small number of scoring steps) is to consider the class of filter functions arising from the iterative procedures[6,1,7]. Following[6], we refer to such procedures as *Iterative Filtering* (IF) methods. In fact, these methods can be viewed as single-step CRSes, i.e. with just one scoring step, and no dependence on k. For more details about IF procedures see[5].

In this work we are interested in using an *IF-based CRS with memory*, which are systems whose filter function value uses the outcome of an IF method. More in detail, these methods are realized by using the trust modelling

$$\widetilde{T}(k) = T(k)\,\mathrm{diag}(\mathbf{1} - z) + \widetilde{T}(k-1)\,\mathrm{diag}(z) \qquad (k \in \mathbf{N}) \,, \tag{5}$$

where $T(k)$ is the IF outcome at step k, $\widetilde{T}(k-1)$ is the trust matrix at the previous step, and z is a vector of weights whose meaning is deeply discussed in[5]. Experimental results provided in Section 4, will show that IF-based CRSes with memory can be suitably used in the context of a cultural heritage scenario.

3. Definition of vising styles

For the art exhibit "*The Beauty and the Truth*", we have collected log files related to 253 visitors. The analysis of their behaviours within the cultural space has enabled us to define a classification of the visiting styles. In the literature, several research papers focus this objective. The starting point of our research was the work in[9], where authors proposed a classification method based on a comparison between behaviours of museum visitors and four "typical" animals (i.e., ant, fish, butterfly and grasshopper). Moreover, we have resorted to the work presented in[10], where, recalling the above mentioned approach, authors have introduced a methodology based on two unsupervised learning approaches for validating empirically their model of visiting styles. Finally, in[11,8,12], we proposed a classification technique able to discover how visitors interact with a complex IoT framework, redefining the visiting styles' definition. For completeness, we report a brief description below. A visitor is: (i) an *ant* (**A**) if it tends to follow a specific path in the exhibit and intensively enjoys the furnished technology; (ii) a *butterfly* (**B**) if it does not follow a specific path but rather is guided by the physical orientation of the exhibits and stops frequently to look for more media contents; (iii) a *fish* (**F**) if it moves around in the center of the room and usually avoids looking at media content details; (iv) a *grasshopper* (**G**) if it seems to have a specific preference for some preselected artworks and spends a lot of time observing the related media contents. Notice that, the four visiting styles are characterized by three different parameters: a_i, τ_i and v_i. More in detail, for the i-th visitor, we denote by:

- a_i, the percentage of viewed artworks;
- tau_i, the average time spent by interacting with the viewed artworks;
- v_i, a value in [0, 1] that measures the quality of the visit, in terms of the sequence of crossed sections (i.e., path).

The classification of the visiting styles is obtained following the scheme summarized in Table 1. In particular, a parameter is assumed *low* if its value is lower than a suitably given threshold, *high* otherwise. By using thresholds $\bar{a} = 0.1$, $\bar{\tau} = 0.5$ and $\bar{v} = 0.58$ for the parameters a_i, τ_i and v_i, respectively, and by following the rules indicated in Table 1, we obtained the visiting styles' classification shown in Figure 1. Details about the tuning of the thresholds are deeply discussed in[13].

Table 1. Characterization of the visiting styles.

Visiting Style	a_i	τ_i	v_i
A	high	-	high
B	high	-	low
F	low	low	-
G	low	high	-

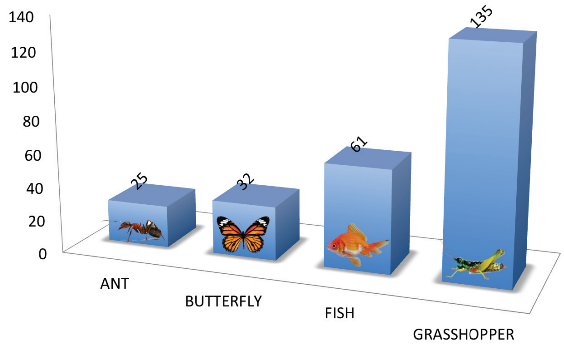

Fig. 1. Classification results.

4. Experiments

Here, we use the CRS methodology by considering the artworks as items and the visitors as raters. The reputation values assigned by the system will be inherently used to confirm the classification results discussed in Section 3 and shown in Figure 1. To this aim we use the following parameters:

- a_i and v_i (see Section 3);
- τ_{ij}, the time spent by the i-th visitor for the j-th artwork. Times τ_{ij} (for $i, j = 1, \dots, 253$) are not actual but just normalized in $[0, 1]$ (by dividing by the maximum of the actual times);
- A_{ij}, the adjacency matrix value. $A_{ij} = 1$ if $\tau_{ij} > 0$, $A_{ij} = 0$ if $\tau_{ij} = 0$.

To use the IF-based CRS with memory we assume that 6 visitors, with high path value (that is 6 **A**s) are trusted raters of an IF-based CRSes with memory with $K = 9$ steps. The simulation is conceived as follows. In the first step the system just assigns high reputations to the 6 trusted visitors. At the next step ($k = 2$), we introduce in the systems all remaining visitors and make the CRSes continue to evolve for last 8 steps. In these steps no more evaluations are added and raters who voted an item at step 2 keep giving the same evaluation on that item at each following step. In terms of adjacency matrix of the CRS, this is realized by setting $A(k) = A$, for $k = 2, \dots, K$, with A defined as before, and $A(1)$ with all ones in the rows related to the trusted raters, and zeros elsewhere. The evaluation matrix does not change along steps and is defined by the following formula:

$$E_{ij} = \begin{cases} v_i \cdot A_{ij} & \text{if } \; v_i \geq \bar{v}, \; a_i \geq \bar{a} \\ 0.5 \cdot v_i \cdot A_{ij} & \text{if } \; v_i < \bar{v}, \; a_i \geq \bar{a} \\ 0 & \text{elsewhere} \end{cases} \tag{6}$$

where the thresholds $\bar{v} = 0.58$ and $\bar{v} = 0.1$ are chosen, respectively, in order to more separate the path values of **A**s and **B**s, and to assign the value 0 to **F**s and **G**s. Moreover, in order to to weight the memory, i.e. the dependence on trusted raters, following[5] we set we set weights z_j so that, at each new iteration, about 80% of the old trust matrix is kept. The results of this test are showed in Figure 2).

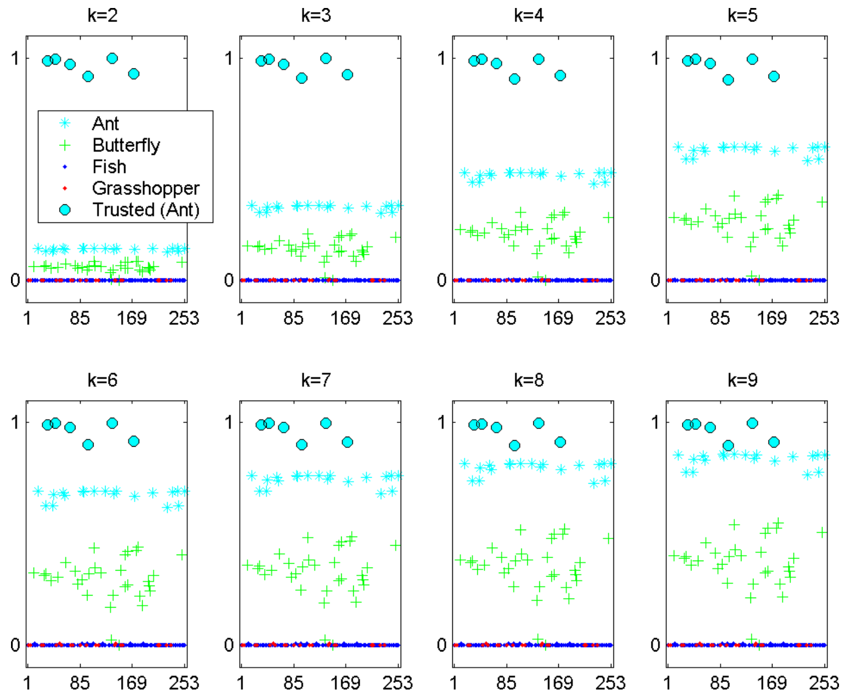

Fig. 2. Reputation values of visitors for various steps. Mark symbol "o": trusted raters (**A**s). Mark symbol "*": other **A**s (introduced at step 2). Mark symbol "+": **B**s (introduced at step 2). Mark symbol ".": **G**s and **F**s (introduced at step 2).

In this case, at each step, the reputation values seem to cover three not overlapped ranges:

- in the top range we get 6 trust raters (marked with "o") and all the other visitors classified as **A**s. Then, the insertion of the memory has the effect of attributing high reputations to the visitors with the same visiting style of the trusted. In other words, since the evaluation parameter we use is the path substantially, this test assigns maximum reputations to the **A**s that have the highest path values;
- in the medium range we get almost all the **B**s. Then, the system is able to differentiate the behaviour of **B**s from the behaviour of **G**s and **F**s.
- in the bottom range we get the remaining visitors which share very low reputation values. Apart from the exception of 2 **B**s, they are all **G**s and **F**s.

Then, this test proves that the system in able to clearly distinguish the reputations by assigning reputation scores that partition the visitors in three populations: **A**s, **B**s, and **G**s with **F**s. Moreover, the convergence of such a system can be observed also looking at the behaviours of reputations when the steps increase: indeed, trusted visitors seem to attract more and more other visitors with similar behaviour. To obtain a further distinction between **G**s and **F**s, we perform a sort of second stage of the test in which visitor with high reputation, all **A**s and almost all **B**s, are removed from the input dataset. Finally, Figure 3 highlights that the system assigns, in mean, high reputations at **G**s, while low values are typical for **F**s. However, in this case there is not a net separation between values and so a suitable threshold for the reputation should be introduced to classify these last visitors.

5. Conclusions

In this paper we have highlighted that the usage of well-known mathematical models and computational approaches, in order to manage data is applicable and usable in the IoT research field. At this aim, we have adopted a CRS methodology to evaluate visiting styles' dynamics in the cultural heritage context. The results obtained are comparable

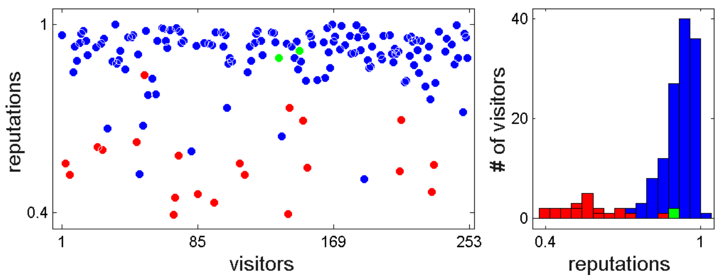

Fig. 3. Assigning reputation values to the remaining visitors. **F**s are in red, **G**s are in blue, the 2 **B**s are in green

with other methods known in literature. In future works we will extend this approach also to assign a rating to the artworks.

Acknowledgment

Authors thank the Department of Mathematics and Applications "R. Caccioppoli", for supporting the organization of the workshop Data Mining on IoT Systems (DaMIS16) in the 7th International Conference on Emerging Ubiquitous Systems and Pervasive Networks EUSPN 2016, and DATABENC, the High Technology District for Cultural Heritage management of Regione Campania (Italy), under CHIS project.

References

1. de Kerchove, C., Dooren, P.V.. Reputation systems and optimization. *Siam News* 2008;**41**(2):1–3.
2. Yang, Y., Sun, Y.L., Kay, S., Yang, Q.. Defending online reputation systems against collaborative unfair raters through signal modeling and trust. In: *Proceedings of the 2009 ACM Symposium on Applied Computing (SAC), Honolulu, Hawaii, USA, March 9-12, 2009*. 2009, p. 1308–1315. URL: `http://doi.acm.org/10.1145/1529282.1529575`. doi:10.1145/1529282.1529575.
3. Taylor, B.. Sender reputation in a large webmail service. In: *CEAS 2006 - The Third Conference on Email and Anti-Spam, July 27-28, 2006, Mountain View, California, USA*. 2006, URL: `http://www.ceas.cc/2006/listabs.html#19.pdf`.
4. Walsh, K., Sirer, E.G.. Experience with an object reputation system for peer-to-peer filesharing (awarded best paper). In: *3rd Symposium on Networked Systems Design and Implementation (NSDI 2006), May 8-10, 2007, San Jose, California, USA, Proceedings*. 2006, URL: `http://www.usenix.org/events/nsdi06/tech/walsh.html`.
5. Galletti, A., Giunta, G., Schmid, G.. A mathematical model of collaborative reputation systems. *International Journal of Computer Mathematics* 2012;**89**(17):2315–2332. doi:10.1080/00207160.2012.715641.
6. de Kerchove, C., Dooren, P.V.. Iterative filtering in reputation systems. *SIAM Journal on Matrix Analysis and Applications* 2010;**31**(4):1812–1834. doi:10.1137/090748196.
7. Laureti, P., Moret, L., Zhang, Y.C., Yu, Y.K.. Information filtering via iterative refinement. *EPL (Europhysics Letters)* 2006;**75**(6):1006. URL: `http://stacks.iop.org/0295-5075/75/i=6/a=1006`.
8. Cuomo, S., De Michele, P., Galletti, A., Ponti, G.. Visiting styles in an art exhibition supported by a digital fruition system. In: *11th International Conference on Signal-Image Technology and Internet-Based Systems, SITIS 2015, Bangkok, Thailand, November 23-27, 2015*. 2015, p. 775–781. URL: `http://dx.doi.org/10.1109/SITIS.2015.87`. doi:10.1109/SITIS.2015.87.
9. Veron, E., Levasseur, M., Barbier-Bouvet, J.. *Ethnographie de l'exposition*. Etudes et recherche. Centre Georges Pompidou; 1991. URL: `https://books.google.it/books?id=OT-hHAAACAAJ`.
10. Zancanaro, M., Kuflik, T., Boger, Z., Goren-Bar, D., Goldwasser, D.. Analyzing museum visitors' behavior patterns. In: *User Modeling 2007, 11th International Conference, UM 2007, Corfu, Greece, June 25-29, 2007, Proceedings*. 2007, p. 238–246.
11. Cuomo, S., De Michele, P., Galletti, A., Pane, F., Ponti, G.. Visitor dynamics in a cultural heritage scenario. In: *DATA 2015 - Proceedings of 4th International Conference on Data Management Technologies and Applications, Colmar, Alsace, France, 20-22 July, 2015*. 2015, p. 337–343.
12. Cuomo, S., De Michele, P., Galletti, A., Ponti, G.. *Data Management Technologies and Applications: 4th International Conference, DATA 2015, Colmar, France, July 20-22, 2015, Revised Selected Papers*; vol. 584 of *Communications in Computer and Information Science*; chap. Classify Visitor Behaviours in a Cultural Heritage Exhibition. Springer International Publishing. ISBN 978-3-319-30161-7; 2016, p. 17–28. URL: `http://link.springer.com/chapter/10.1007/978-3-319-30162-4_2`. doi:10.1007/978-3-319-30162-4.
13. Cuomo, S., De Michele, P., Galletti, A., Ponti, G.. *Intelligent Interactive Multimedia Systems and Services 2016*; vol. 55 of *Smart Innovation, Systems and Technologies*; chap. Influence of Some Parameters on Visiting Style Classification in a Cultural Heritage Case Study. Springer International Publishing. ISBN 978-3-319-39345-2; 2016, p. 567–576. URL: `http://dx.doi.org/10.1007/978-3-319-39345-2_50`; iIMSS - IS07: Internet of Things: Architecture, Technologies and Applications Invited Session of KES 2016.