Tracking and Analyzing the "Second Life" of TV Content: a Media and Social-driven Framework

L. Vignaroli[1], C. Schifanella[1], K. S. Candan[2], R. Pensa[3], M.L. Sapino[3]
[1]RAI – Centre for Research and Technological Innovation, Turin, Italy
{luca.vignaroli, claudio.schifanella}@rai.it

[2]Arizona State University
candan@asu.edu

[3]Department of Compute Science, University of Turin
{pensa, mlsapino}@di.unito.it

Abstract

People on the Web talk about television. TV users' social activities implicitly connect the concepts referred to by videos, news, comments, and posts. The strength of such connections may change as the perception of users on the Web changes over time. With the goal of leveraging users' social activities to better understand how TV programs are perceived by the TV public and how the users' interests evolve in time, in this work, a framework that allows to manage, explore and analyze the heterogeneous and dynamic data coming from different information sources which play a role in what we call the "second life" of TV content will be exposed

Keywords: social network analysis, social TV, second screen applications

1. Introduction

In recent years, the way users watch television has radically changed. With the introduction of digital television and the growing number of generic and thematic channels the final user tends to use new forms of navigation in the television content space. To help users' navigation, broadcasters provide new enriched metadata services such as EPGs (Electronic Program Guides) which accurately describe scheduled programs. More in general, traditional broadcasters are moving to become "digital media companies" making their content available not only on the traditional TV channels, but, in a multi-platform perspective, the introduction of OTT (Over The Top) services for these new companies is a must to bring its content offering to the needs and expectation of the audience. Also the home environment is changing since many smart users watch television while using a notebook, smartphone or a tablet as secondary screen more or less related to the broadcast programmes. At the same time, social networks allow the final user to be immersed in a collaborative environment and to talk about television. TV users' social activities implicitly make connections between concepts by means of videos, news, comments, and posts. The strength of such connections may change as the perception of users on the Web changes over time. Moreover, user-generated contents (UGC) are revolutionizing all phases of the content production value chain, in particular it can be observed that a very large number of UGCs include significant portions of content already broadcast by the TV networks. In this context a number of Social TV applications are emerging, providing to the final user tools for social interaction while watching television or media content related to a particular TV program. If properly leveraged, these collaborative social environments can be seen as rich information data sources, indirectly returning to broadcasters and content producers some form of implicit feedback from the final users.

In general, television content evolves in time and its life undergoes a number of different steps. Firstly, a content is typically produced and put on air by (e.g.) the broadcaster. In addition, a copy of it, enriched with its description together with a collection of related metadata is (statically) stored in the TV archive to be reused if needed, and the broadcaster puts the description in his EPG. Big broadcasters also make their TV contents available in the Internet site or as OTT service. Secondly, after the on-air time, the broadcaster is interested in estimating to which degree users are satisfied with the broadcasted content by means of quantifiable data, such as TV audience measurements, or the number of user views in the online resource.

This data collection concludes the "on-air phase" of the content, whose life most likely spans much beyond that point.

In fact, successful TV programs will be probably commented on Twitter and Facebook and published (either entirely, or more often in part) online by users, for example on YouTube. The online posting of (a fragment of) a TV program starts a second phase for it, which can be called "on-line phase".

During this on-line phase the content will be watched, tagged, liked, commented and shared again and again by users in the network. The television content turns to be a "magnet" for users in the network attracting other users, and it becomes a "Social Object".

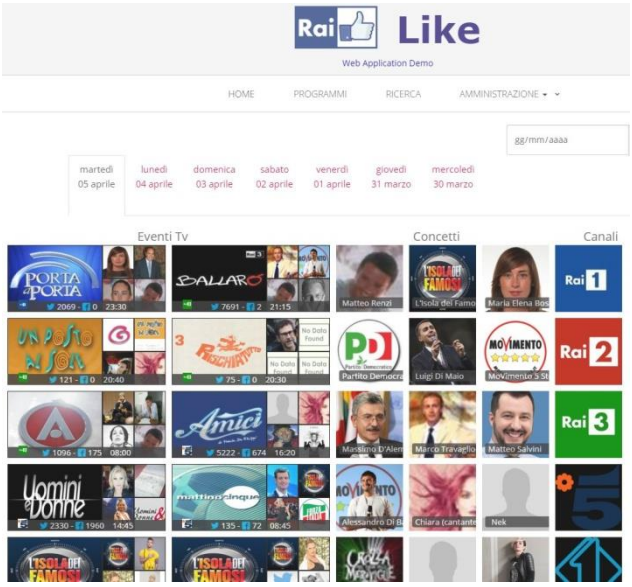

Figure 1: home page of the application demo

Typically, YouTube is the first place where people come to look for a television content which they missed and despite the copyright issues this might raise, this is a fact. Instead of contrasting this tendency, content producers can try to leverage it to their benefit. In fact, users that post TV content in YouTube are implicitly disseminating, describing and publishing it for free.

As an example of the "extended" life of television content, it can be observed that it is very easy to find in YouTube segments of TV programs that have been uploaded long time ago, even years ago, and are still nowadays very often watched, commented and liked. Without video-sharing platforms, the TV content would otherwise be just stored in the archives of the broadcaster and it would be inaccessible to users. As time passes and the users' social context changes, the way any specific television content is perceived also evolves. For example, a content can attract a new community of users interested in it, or it might change its own meaning because of a new fact happened in the world. If timely discovered these phenomena could be leveraged by the broadcaster and some of the contents already available in the archives could be considered for the production of new programmes based on new interests of the public.

In this context, one of the main objectives of the project is to capture how can the TV content evolve and detecting which phenomena can emerge from the contents' evolution.

2. The second life of TV content

As a result, the Rai Research Centre has been developed a framework for enabling the integration of heterogeneous data coming from the knowledge sources (broadcasters' archives, EPGs, collected audience data, social networks, etc.) which play a role in the "second life" of TV content, starting from its production phase, going through the on-air phase, and continuing with the on-line phase. The system is designed to provide powerful tools helping to highlight the tight interactions between the Web world and the TV world.

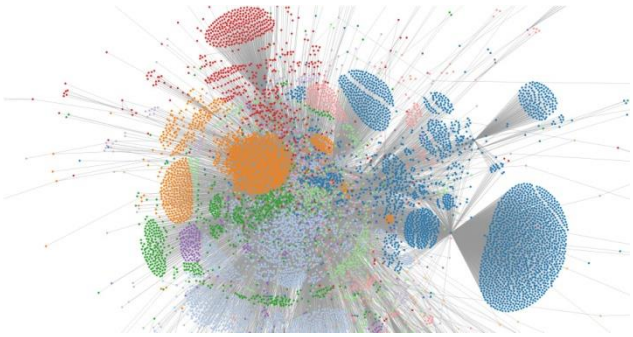

Figure 2: Analyzing interaction graph communities

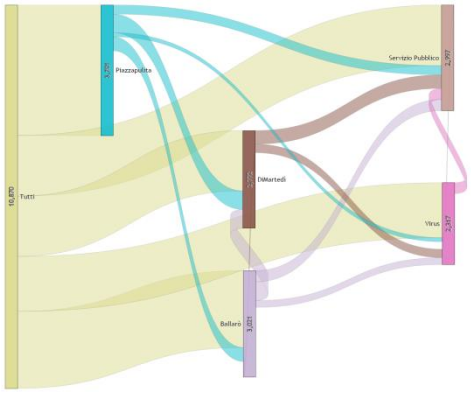

Figure 3: Migration of users between different TV programs

From a more technical point of view the framework enables the integration of various information sources into a unique "knowledge base", modelled as a knowledge graph. Integrating domain and general purpose ontologies, as well as social interactions among users and social media, the knowledge base can be queried and analyzed as a whole, enabling the discovery of new and interesting cross-domain patterns.

The integration framework consists of three main layers: a source processing layer, a knowledge graph layer and a knowledge query and analysis layer.

The "source processing layer" has the role of collecting all the data which will be conveyed in the model. It accesses a number of predefined web/social/media sources (e.g., broadcasters official web sites, social networks, TV channels, etc) and processes them in order to extract those information units which will be represented as nodes in the knowledge graph, as well as information that support the existence of relationships among them.

The "knowledge layer" is the core of the system and represents the result of public actions of users in social environments. Furthermore, relationships are introduced between subjects and social objects and between social objects and concepts. Other relationships involve entities of the same type and these are called structural dependencies. Moreover, social objects evolve in time, hence, as a special case of representation relationship, we consider the temporal representation of a social object against a special type of concept called time objects.

Finally, the "knowledge query and analysis layer" consists in a set of components for querying, browsing and analyzing the knowledge graph. A query module extracts sub-graphs from the knowledge layer based on user's requests and constraints. An analysis module provides a set of analysis tools to obtain insights of the data.

The development of the system had to face some challenges in order to make manageable the analysis over the huge amount of data which is being collected from the various knowledge sources. The optimisation of algorithms and the definition of the data model has been done in close cooperation with the Department of Computer Science at the University of Torino.

The interactive demo (Figure 1), thanks to powerful and flexible analysis tools (Figure 2,3) of the social network data flow, allows to make available to Rai a number of strategic services such as user behaviour profiling, brand reputation, community detection and recommendation systems for contents and advertisements.

3. Bibliographical References

A. Antonini, L. Vignaroli, C. Schifanella, R. G. Pensa, and M. L. Sapino, "Mesoontv: a media and social-driven ontology-based TV knowledge management system," in Proceedings of ACM HT '13. ACM, 2013, pp. 208–213.

A. Antonini, R. G. Pensa, M. L. Sapino, C. Schifanella, R. T. Prioletti, and L. Vignaroli, "Tracking and analyzing TV content on the web through social and ontological knowledge," in Proceedings of EuroITV '13. ACM, 2013, pp. 13–22.

X. Han, W. Wei, C. Miao, J. Mei, and H. Song, "Context-aware personal information retrieval from multiple social networks," IEEE Comp. Int. Mag., vol. 9, no. 2, pp. 18–28, 2014. [Online]. Available: http://dx.doi.org/10.1109/MCI.2014.2307222.