



AN ARCHITECTURE FOR EVOLVING THE ELECTRONIC PROGRAMME GUIDE FOR ONLINE VIEWING

Emad Abdulrazzaq Jasim Al-Mohammed

School of Computing, Science and Engineering

University of Salford, Manchester, UK

Submitted in Partial Fulfilment of the Requirements for the Degree of

Doctor of Philosophy, 2017

LIST OF CONTENT

| | |
|--|-----|
| LIST OF CONTENT | I |
| LIST OF PUBLICATIONS | IV |
| LIST OF TABLES | V |
| LIST OF FIGURES | VI |
| ACKNOWLEDGEMENTS | XI |
| ABBREVIATIONS | XII |
| LIST OF SYMBOLS | XIV |
| ABSTRACT | XV |
| Chapter One..... | 2 |
| 1.1 Introduction..... | 2 |
| 1.2 Research Problem | 7 |
| 1.3 Research Aim..... | 9 |
| 1.4 Research Questions..... | 9 |
| 1.5 Research Methodology | 9 |
| 1.6 Contribution to Knowledge | 10 |
| 1.7 Thesis Layout..... | 11 |
| Chapter Two | 14 |
| 2.1 Introduction..... | 14 |
| 2.2 Electronic Programme Guide Systems | 14 |
| 2.2.1 Electronic Programme Guide for IPTV System..... | 18 |
| 2.2.2 Personalized Electronic Programme Guide..... | 22 |
| 2.3 IPTV System..... | 25 |
| 2.4 Searching Systems | 28 |
| 2.4.1 World Wide Web Searching Systems | 29 |
| 2.4.2 Multimedia Searching Systems..... | 31 |
| 2.4.3 Integrating the Internet with Multimedia Searching Systems | 34 |
| 2.4.4 Smart TV Searching Systems..... | 35 |
| 2.5 Recommendation Systems | 38 |
| 2.5.1 Traditional Recommendation Approaches..... | 38 |
| 2.5.2 Personalized Recommendation | 45 |
| 2.5.3 Group Recommendation | 49 |

| | |
|--|-----|
| 2.5.4 Context-awareness Recommendation | 51 |
| 2.6 Social Networks | 55 |
| 2.6.1 Social Networks-based Recommendation Systems | 57 |
| 2.6.2 Hybridization the Recommendation Operation with Social Network | 61 |
| 2.6.3 Integrating Facebook into Recommendation Operation | 63 |
| 2.7 Summary | 67 |
| Chapter Three | 74 |
| 3.1 Introduction..... | 74 |
| 3.2 Content Searching System | 75 |
| 3.3 Recommendation System | 75 |
| 3.4 Context-awareness Function..... | 77 |
| 3.5 Service Requirements | 77 |
| Chapter Four | 80 |
| 4.1 Introduction..... | 80 |
| 4.2 General System Architecture | 80 |
| 4.3 Content Searching Module | 81 |
| 4.3.1 Extracting Metadata | 85 |
| 4.4 Recommendation Module..... | 90 |
| 4.4.1 User Viewing History..... | 90 |
| 4.4.2 User Clusters | 94 |
| 4.4.3 Social Networks Recommendations | 99 |
| 4.5 Context-aware Module | 101 |
| 4.6 Content Database Server..... | 104 |
| 4.7 Summary | 105 |
| Chapter Five | 108 |
| 5.1 Introduction..... | 108 |
| 5.2 Content Search Module | 108 |
| 5.3 Recommendation Module..... | 117 |
| 5.3.1 Viewing History Based (Content-Based) Recommendation..... | 118 |
| 5.3.2 Users Clusters..... | 121 |
| 5.3.3 Social Networks Contribution..... | 124 |
| 5.3.4 Filtering Operation | 128 |
| 5.4 Context-aware Module | 133 |

| | |
|--|-----|
| 5.5 Content Database Server..... | 137 |
| 5.6 Online Electronic Programme Guide..... | 146 |
| 5.7 Summary..... | 149 |
| Chapter Six..... | 151 |
| 6.1 Introduction..... | 151 |
| 6.2 User Trial Scenario..... | 151 |
| 6.3 Evaluating the Function of Multiple Online Content Sources..... | 155 |
| 6.4 Evaluating the Recommendation Function..... | 159 |
| 6.4.1 Test Scenario..... | 159 |
| 6.4.2 Evaluation Metrics..... | 161 |
| 6.4.3 First Step of User Trial (Cold-start Problem)..... | 162 |
| 6.4.4 Second, Third and Fourth Step of Users' Trial (Viewing History)..... | 166 |
| 6.5 Filtering Operation..... | 172 |
| 6.6 Direct Search..... | 175 |
| 6.7 Summary..... | 179 |
| Chapter Seven..... | 182 |
| 7.1 Conclusions..... | 182 |
| 7.2 Future Work..... | 185 |
| REFERENCES..... | 187 |
| Appendices..... | 204 |

LIST OF PUBLICATIONS

- Al-Mohammed, E., & Linge, N. (2016, December). A generic, personalized electronic program guide system for accessing multiple online TV providers. In *11th International Conference for Internet Technology and Secured Transactions (ICITST), 2016* (pp. 291-296). IEEE.
- Al-Mohammed, E., & Linge, N. (2016, December). An Architecture for Personalized Electronic Program Guide to Access Online Television Content. *International Journal of Chaotic Computing (IJCC)*, vol. 4, issue. 2 (pp. 87-95).
- Al-Mohammed, E., & Linge, N. (2017, June). Cloud based electronic programme guide system with integrated social network recommendations. *International Journal of e-Education, e-Business, e-Management and e-Learning (IJEEEE)*, vol. 7, No. 2 (pp. 100-110).

LIST OF TABLES

Chapter One

| | |
|--|---|
| Table 1.1 Average Time Spent per day on Different Online Activities (hrs:mins) | 6 |
| Table 1.2 Internet Video Consuming (Petabyte per month) until 2019 | 6 |

Chapter Two

| | |
|---|----|
| Table 2.1 List of sources that connected to Atlas | 16 |
| Table 2.2 Summary of Literature Review | 71 |

Chapter Four

| | |
|---|----|
| Table 4.1 Programme Metadata Features | 84 |
| Table 4.2 Programme's Features Provided by Several Online Content Providers | 88 |
| Table 4.3 List of Used Categories | 95 |
| Table 4.4 Online Content Sources' Categories | 96 |

Chapter Five

| | |
|---|-----|
| Table 5.1 Types of Videos Resolution That Employed in the Prototype | 130 |
| Table 5.2 Device Type versus Resolution Table | 130 |
| Table 5.3 Available Network Bandwidth versus Resolution Table | 131 |
| Table 5.4 Timelines of the System Operations | 146 |

LIST OF FIGURES

Chapter One

| | |
|--|----|
| Figure 1.1 Proportion of Adults who Using On-Demand and Online Service in the UK ... | 4 |
| Figure 1.2 Proportion of Time Spent Watching, attributed to activities, by Age Group | 5 |
| Figure 1.3 General Diagram of EPG Provided by Sky | 8 |
| Figure 1.4 Stages of Research Methodology | 12 |

Chapter Two

| | |
|--|----|
| Figure 2.1 Part of Content Metadata showing the Basic Description | 15 |
| Figure 2.2 User Notifications Service in the Developed EPG | 21 |
| Figure 2.3 Architecture of Personal TV Guide System | 23 |
| Figure 2.4 Community-Based Programme Recommendation for EPG System | 25 |
| Figure 2.5 Main Functional Parts of the IPTV System | 26 |
| Figure 2.6 Overall Structure of personalized mobile IPTV system | 28 |
| Figure 2.7 Intelligent Information Retrieval Agent | 31 |
| Figure 2.8 EPG Searching System | 33 |
| Figure 2.9 Supplementary User Interface of Internet Searching system on TV | 35 |
| Figure 2.10 Search Processing in the Smart TV system | 37 |
| Figure 2.11 Hybrid Recommendation Approach | 44 |
| Figure 2.12 Typical Personalized Recommendation System | 46 |
| Figure 2.13 Overview of Show Time System | 48 |
| Figure 2.14 Merging User Profiles Strategy | 50 |

| | |
|--|-----|
| Figure 2.15 Multiple Viewers Graphical User Interface | 51 |
| Figure 2.16 The Possible Ways for Incorporating the Contextual Information into the Recommendation Systems | 53 |
| Figure 2.17 Context-aware Recommendation System | 55 |
| Figure 2.18 Block Diagram Describes the Idea of Social Network Recommendation System | 59 |
| Figure 2.19 IPTV Social Network Recommendation System | 60 |
| Figure 2.20 Incorporating Social contribution and collaborative filtering | 62 |
| Figure 2.21 Selecting and Rating Preferred Movies within the Facebook Application ... | 64 |
| Figure 2.22 Interface Showing Viewing Statistics | 66 |
| Chapter Four | |
| Figure 4.1 Overview of the Proposed System | 80 |
| Figure 4.2 Content Search Module | 82 |
| Figure 4.3 XML-Based Metadata Example | 86 |
| Figure 4.4 YouTube Metadata | 87 |
| Figure 4.5 Recommendation Module | 91 |
| Figure 4.6 Users' Clustering Process | 97 |
| Figure 4.7 Social Network Contribution | 100 |
| Figure 4.8 Context-aware Module | 102 |
| Figure 4.9 Overall System Configuration | 106 |
| Chapter Five | |
| Figure 5.1 Retrieved Programme Raw Metadata From The Movie Database Source | 109 |
| Figure 5.2 Retrieved Programme Metadata After Analysing and Translation Operation | 109 |

| | |
|---|-----|
| Figure 5.3 Part of The Movie Database Interface PHP Code for Extracting the Categories | 110 |
| Figure 5.4 Part of YouTube Interface PHP Code | 112 |
| Figure 5.5 Part of BBC-iPlayer Interface Code | 113 |
| Figure 5.6 Programme Matrix | 114 |
| Figure 5.7 Part of Matlab Code Represents the Combiner and Parser | 115 |
| Figure 5.8 Search Keywords before and after Search Agent Leader | 116 |
| Figure 5.9 Part of Search Agent Leader Code to Read and Write the Search Keywords in Text File | 117 |
| Figure 5.10 Part of Recommendation Process Matlab Code Represents the Similarity Calculations Based on Jaccard Coefficient Measure | 119 |
| Figure 5.11 Part of User Matrix which Represents the Recommended List Based on User Viewing History | 120 |
| Figure 5.12 Part of Matlab Code That Used for User Clustering | 122 |
| Figure 5.13 Part of User Matrix which Represents the Recommended List of Programmes based on the Users Clusters Method | 123 |
| Figure 5.14 Facebook Application | 124 |
| Figure 5.15 Part of Facebook Application that used to retrieve the Information about the 'liked', 'watched' and 'want to watch' Movies and TV Shows for each user | 125 |
| Figure 5.16 Data Retrieved from Facebook (in JSON) about Watched Movies and TV Shows for Particular User | 126 |
| Figure 5.17 Information retrieved from Facebook after Translation Operation | 127 |
| Figure 5.18 Joining the Application in Facebook | 127 |
| Figure 5.19 Facebook Recommended List of Programmes after Analysing and Ranking Operation | 129 |

| | |
|--|-----|
| Figure 5.20 Part of Matlab Code that is used for filtering the Recommended Programmes based on Available Network Bandwidth | 132 |
| Figure 5.21 YouTube Example for encoding MPEG-DASH stream showing Media Presentation Description (MPD) | 133 |
| Figure 5.22 Programme's Features Columns in User Viewing History Matrix | 134 |
| Figure 5.23 User Viewing History | 135 |
| Figure 5.24 Part of the Matlab Code that used to save the Watched Programme in User Viewing History | 136 |
| Figure 5.25 User Matrix for Several Users | 138 |
| Figure 5.26 Content Matrix for Several Users | 139 |
| Figure 5.27 Programmes that Located in One of the Cells (Columns) of Content Matrix | 140 |
| Figure 5.28 Part of Matlab Code that is responsible of Checking the User Content Matrix for any relevant Programmes to Social Network's Recommended List of Programmes . | 141 |
| Figure 5.29 Data Flow Diagram when the User joins the System for the First Time | 142 |
| Figure 5.30 Data Flow Diagram when the System creating the Recommended Lists of Programmes based on all of the Recommendation Schemes | 144 |
| Figure 5.31 Snapshot of the Designed EPG | 148 |
| Chapter Six | |
| Figure 6.1 The Users who joined the Facebook Application | 152 |
| Figure 6.2 The EPG Application (EPGTest) within the List of Application in a Facebook User's Profile | 153 |
| Figure 6.3 Facebook Application Request for Information Retrieving Permission | 154 |
| Figure 6.4 List of Programmes retrieved from Multiple Online Content Sources | 158 |
| Figure 6.5 User Interface Includes the Recommended Lists of Programmes | 160 |

| | |
|--|-----|
| Figure 6.6 Average Precision versus Number of Recommendations for Step One of Users' Trial | 163 |
| Figure 6.7 Mean Absolute Error versus Number of Recommendations for Step One of Users' Trial | 164 |
| Figure 6.8 Average Precision before and after Recommendation Performance Improvement | 165 |
| Figure 6.9 Average Precision versus Number of Recommendations for Second Step of Users' Trial | 166 |
| Figure 6.10 Average Precision versus Number of Recommendations for Third Step of Users' Trial | 167 |
| Figure 6.11 Average Precision versus Number of Recommendations for Fourth Step of Users' Trial | 167 |
| Figure 6.12 Mean Absolute Error versus Number of Recommendations for Second Step of Users' Trial | 168 |
| Figure 6.13 Mean Absolute Error versus Number of Recommendations for Third Step of Users' Trial | 168 |
| Figure 6.14 Mean Absolute Error versus Number of Recommendations for Fourth Step of Users' Trial | 169 |
| Figure 6.15 Average Precision Value as a Function of User Trial's Step | 171 |
| Figure 6.16 Average Mean Absolute Error Value as a Function of User Trial's Step | 171 |
| Figure 6.17 Filtered Recommended List for Three Different Bandwidth Values | 174 |
| Figure 6.18 Filtered Recommended List based on User's Device | 174 |
| Figure 6.19 Entering the Search Keyword in the Direct Search Operation | 175 |
| Figure 6.20 Resulted List of Programmes from the Direct Search Operation | 176 |
| Figure 6.21 The Details of Selected Programme in Direct Search Operation | 177 |
| Figure 6.22 Watching a Selected Programme from the displayed List | 178 |

ACKNOWLEDGEMENTS

This thesis would not have been possible without the help and support of many people. I would like to gratefully acknowledge the contributions of those named below and those whom I may have unfortunately forgotten.

My deep gratitude goes first to Professor Nigel Linge for his kind support, patient guidance and valuable suggestions that helped to produce this thesis in its present form. His unwavering enthusiasm for knowledge kept me constantly engaged with my research and his personal generosity helped make my time at University of Salford enjoyable.

I'm indebted to the support of Ministry of higher education and scientific research of Iraq. Sincere thanks to the staff of Iraqi Cultural attaché in United Kingdom. And I would like to thank the school of Computing, Science and Engineering of the University of Salford for providing the support and facilities to undertake this research.

Sincere thanks to all my family members and especially my mother, my wife and son for supporting and encouraging me to pursue this degree. Without their encouragements, I would not have finished the PhD degree.

ABBREVIATIONS

| | |
|-------|--|
| API | Application Programming Interface |
| CAGR | Compound Annual Growth Rate |
| CF | Collaborative Filtering |
| DDL | Description Definition Language |
| DSLAM | Digital Subscriber Line Access Multiplexer |
| EPG | Electronic programme Guide |
| GUI | Graphic User Interface |
| HD | High Definition |
| HTTP | Hypertext Transfer Protocol |
| ID | Identity |
| IMDB | Internet Movies Database |
| IP | Internet Protocol |
| IPTV | Internet protocol Television |
| JSON | Java Script Object Notation |
| MAE | Mean Absolute Error |
| MPEG | Moving Picture Expert Group |
| OTT | Over The Top |
| PC | Personal Computer |
| PVR | Personal Video Recorder |
| QoS | Quality of Service |

| | |
|------|----------------------------|
| RWR | Random Walk with Restarts |
| SN | Social Networks |
| STB | Set-Top-Box |
| STD | Standard Definition |
| TV | Television |
| UCC | User Created Content |
| URL | Uniform Resource Locator |
| VoD | Video on Demand |
| VOIP | Voice Over IP |
| XML | Extensible Markup Language |

LIST OF SYMBOLS

| | |
|-------|---|
| A_i | Vector Component |
| B_i | Vector Component |
| C_j | jth Clusters |
| i_l | Data Points |
| p_i | Predicted Rating |
| r_i | Real Rating |
| w_j | Cluster Centres |
| A | Object Vector |
| B | Object Vector |
| k | Number of Clusters |
| K | Number of Keywords in Vector A |
| L | Number of Keywords in Vector B |
| SL | Similarity Level |
| E | Error Function of Quality of Clustering |
| N | Number of Predictions |
| Sim | Similarity |
| l | Number of Points |
| n | Number of Objects Component |

ABSTRACT

Watching television and video content is changing towards online viewing due to the proliferation of content providers and the prevalence of high speed broadband. This trend is coupled to an acceleration in the move to watching content using non-traditional viewing devices such as laptops, tablets and smart phones. This, in turn, poses a problem for the viewer in that it is becoming increasingly difficult to locate those programmes of interest across such a broad range of providers.

In this thesis, an architecture of a generic cloud-based Electronic Programme Guide (EPG) system has been developed to meet this challenge. The key feature of this architecture is the way in which it can access content from all of the available online content providers and be personalized depending on the viewer's preferences and interests, viewing device, internet connection speed and their social network interactions. Fundamental to its operation is the translation of programme metadata adopted by each provider into a unified format that is used within the core system. This approach ensures that the architecture is extensible, being able to accommodate any new online content provider through the addition of a small tailored search agent module. The EPG system takes the programme as its core focus and provides a single list of recommendations to each user regardless of their origins.

A prototype has been developed in order to validate the proposed system and evaluate its operation. Results have been obtained through a series of user trials to assess the system's effectiveness in being able to extract content from several sources and to produce a list of recommendations which match the user's preferences and context.

Results show that the EPG is able to offer users a single interface to online television and video content providers and that its integration with social networks ensures that the recommendation process is able to match or exceed the published results from comparable, but more constrained, systems.

CHAPTER ONE

INTRODUCTION

Chapter One

1.1 Introduction

Nowadays, watching television has become a daily part of our digital lives. However, television viewing has seen a switch from the traditional linear viewing of broadcast content to time-shifted or on-demand online viewing. The online consumption of television and video content through the Internet has grown significantly through the use of computers, mobile phones, tablets and other devices. The term online video is general and refers to all video content available over the Internet. It is more general than IPTV or Internet TV. IPTV is the delivery of programming by video stream encoded as a series of IP packets and it is end-to-end service which is distributed by a service provider. Although IP (Internet Protocol) is included in its name, it just refers to the method of sending the encoded video stream over a secure and managed network. It differs from Internet TV which is generally available content distributed over the Internet which can be accessed wherever a broadband connection exists. To understand the difference between those two systems, it is better to compare them based on some of distribution and services features such as geographical limitations, service access, image quality, content and charges. IPTV can only deliver its services over networks that are belonging to telecom operators. Therefore, it is only available where these companies have installed networks. On the other hand, Internet TV is based on the Internet; therefore, it is available wherever a broadband Internet connection exists. Additionally, IPTV requires a decoder box (set-top-box) and contract with the service provider to access its services while Internet TV follows the Internet access rules and the content are free. Moreover, the image quality could be guaranteed by IPTV providers while for Internet TV it depends on the quality of connection. Furthermore, the content over Internet TV is either user-generated content or traditional television via the Internet while the IPTV content is completely controlled by the companies that offer the service and provides broadcast television content and movies supplied by the established media companies.

On the other hand, the online video is more general than Internet TV. Online video content providers include major broadcasting networks, film studios, Over The Top (OTT) service providers and television channels together with user-generated content. Moreover, online TV can include broadcast pay-TV or TV everywhere that provides the subscribers with TV

channels besides the Video on Demand (VoD) content. However, VoD refers to a feature provided by the online content providers and companies to enable the users to select and watch video content from a server at any time. This content could be watched on television or computer's screen and it is available for users who have the authority to access the providers' server. It is always pre-paid service or accompanied with another service such as IPTV or pay-TV.

Viewing content online offers many advantages such as an increase in the variety of content providers, accessing video-on-demand at any time and a more interactive and personalized experience due to the ease of sharing content through online channels. In 2016, the Pew project reported that 72% of Americans used some type of shared or on-demand online service (Smith, 2016). On the same topic, a separate report stated that the number of users that watched or downloaded video online was 71% of Internet users (Bondad-Brown, Rice, & Pearce, 2012). Moreover, according to a study by Hub Entertainment Research, time-shifted viewing accounts for 53% of the total viewing amongst those who have broadband Internet and spend at least five hours watching television per day. This ratio rises to 61% for those who are between 16-34 years old (Taylor, 2015).

Research published by Ofcom (the Government-approved regulatory and competition authority for the broadcasting, telecommunications and postal industries of the United Kingdom) in 2016 also supports this trend. Their research showed that 74% of UK adults have viewed an on-demand or online service in the past 12 months as shown in figure 1.1. Furthermore, the figure shows that there is an increase in viewing the on-demand and online service in all age levels for both genders comparing with the same set of data gathered in 2014 (2014's data has shown to the right side of the figure). Moreover, the research showed that 50% of young people in the UK are no longer watching live TV (Ofcom, 2016, Dawes, 2015).

Online OTT content providers such as Netflix, Amazon and YouTube now have a massive number of subscribers. As an example, the average number of unique YouTube users per month in the US is more than 167 million users. These providers can often offer more choices to their subscribers for watching movies and TV shows/series that are not available on cable, satellite or terrestrial TV. These services are attractive to viewers for which the scheduled live broadcast time is inconvenient and thereby facilitate the option for time-

shifted and on-demand viewing. An example, there are more than 57 million people across 50 countries who have a Netflix subscription. Moreover, more than 2 billion hours of TV shows and movies are streamed by Netflix to their subscribers every month (Bing, 2015).

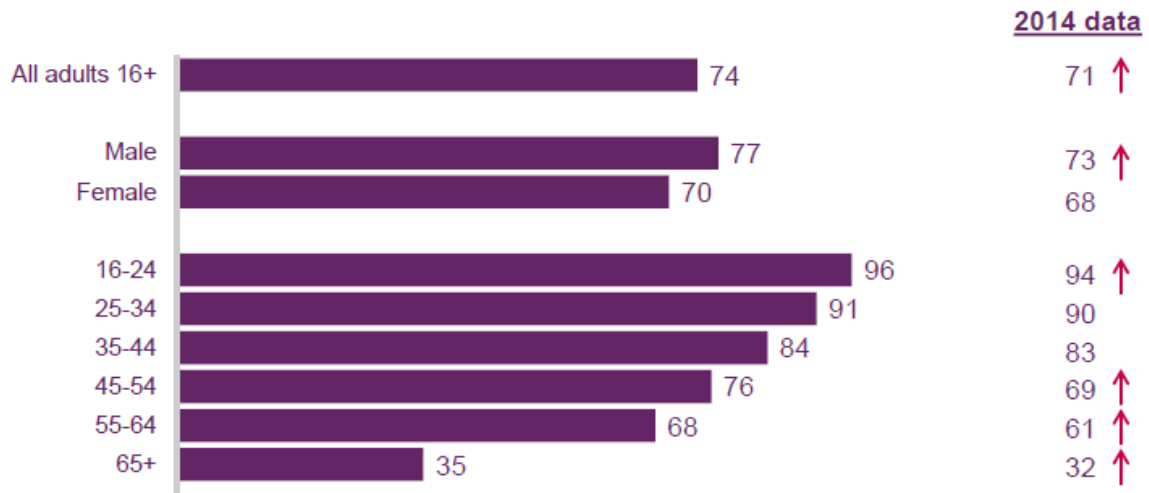


Fig. 1.1 Proportion of Adults who Using On-Demand and Online Service in the UK
(Ofcom, 2016)

This wide spread and growing consumption of online video content is coupled to an acceleration in the move to watching television content using non-traditional viewing devices such as laptops, tablets and smart phones. Currently 57% of young people in the UK who are between 16-24 years old now watch on-demand and catch up programmes on their computers and smartphones rather than on TV connected to a set-top-box whilst 45% watch on their smartphones. Moreover, 42% of them watch short videos from services including YouTube (Ofcom, 2015). This trend has, in turn, led content providers to invest in advertising their services via social networks. It is now common for programmes and channels to encourage their viewers to follow and interact through social networks such as Twitter and Facebook. That level of interaction then turns television watching into a social experience where viewers can share their opinions and thoughts in real time whilst watching. An interesting study by Yahoo advertising in 2014 showed that 86% of users who have Internet access through their smart phones used their phones whilst watching

television and 25% of them said that they had browsed for relevant content related to what they were watching (Holanda, Guilherme, da Silva, & Goussevskaia, 2015).

According to the previous studies, social TV and consumption of online video content have become an important part of people’s digital lives. Although there is much research about online viewing, little of it has focused on providing search and recommendation services for online viewers. A study by Lee claimed that online video services were displacing other media such as cable TV and showed that the time spent accessing online video services reduced the time spent with conventional, linear TV viewing. Hence, the implication being that watching online video will replace alternative modes of delivery in the course of time (Lee & Lee, 2015). The Ofcom Communication Market Report 2016 (Ofcom, 2016) also supported this claim which showed the decreasing reliance on live TV as a function of age, with the younger age group accessing more on-demand services as shown in figure 1.2.

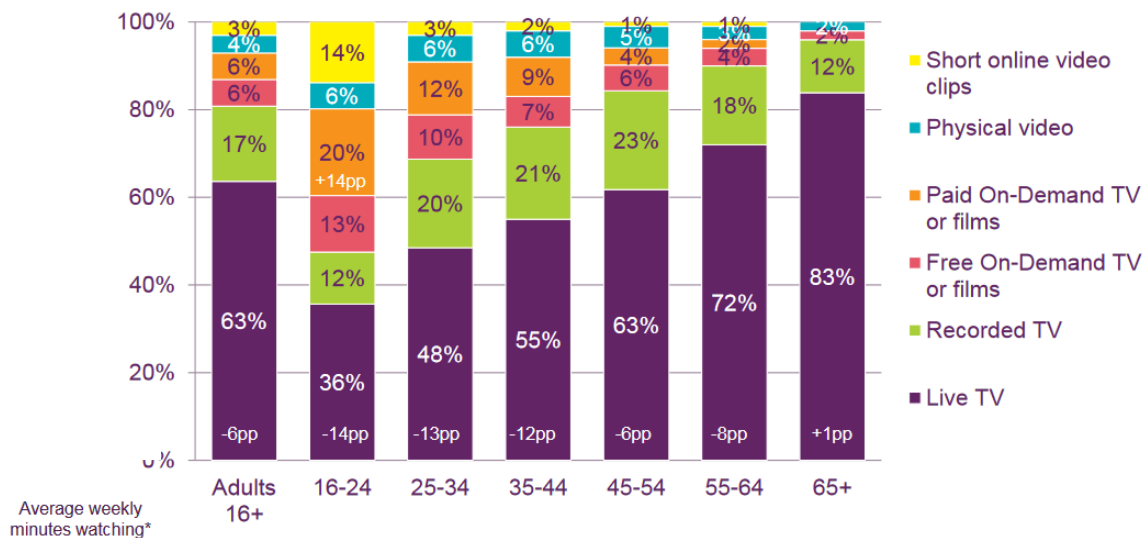


Fig. 1.2 Proportion of Time Spent Watching, attributed to activities, by Age Group (Ofcom, 2016)

According to a report published in YouTube about the statistics at the end of 2016, YouTube was receiving more than 900 million unique visits every month (YouTube, 2016, STATS, 2017). In 2015, users were spending an average of 1 hr 55 mins in watching

online video each day and only 1 hr 44 mins on social networks. Table 1.1 shows how the average time spent per day on different online activities by US adults has changed between 2011 and 2015 (Walgrove, 2015). This wide spread consumption of online video content and the large number of video content providers makes finding content that is of interest to the viewer less easy and consequently increases the risk of missing a vital item.

| Activity | 2011 | 2012 | 2013 | 2014 | 2015 |
|-----------------|-------------|-------------|-------------|-------------|-------------|
| Online Video | 00:39 | 01:03 | 01:22 | 01:39 | 01:55 |
| Social Networks | 01:11 | 01:28 | 01:36 | 01:39 | 01:44 |
| Digital Radio | 00:53 | 00:55 | 00:56 | 01:00 | 01:05 |
| Facebook | 00:30 | 00:33 | 00:35 | 00:38 | 00:40 |

Table 1.1. Average Time Spent per day on Different Online Activities (hrs:mins)
(Walgrove, 2015)

Moreover, many research investigated the ability of watching online television content using different types of devices such as smartphones and tablets. Cisco predicts in their research that consumer online video traffic will exceed 80% of overall Internet traffic by 2019 and especially in the mobile data traffic. Additionally, the calculated Compound Annual Growth Rate (CAGR) of the expected Internet video consuming for mobile networks reaches to 67% by 2019. Table 1.2 shows the expected Internet video consumed until 2019 where the values are represented in Petabytes per month (Cisco VNI, 2015).

| Network Type | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | CAGR 2014 and 2019 |
|---------------------|-------------|-------------|-------------|-------------|-------------|-------------|---------------------------|
| Fixed | 20485 | 25452 | 32981 | 43226 | 56771 | 74319 | 29% |
| Mobile | 1139 | 2014 | 3475 | 5842 | 9407 | 14999 | 67% |

Table 1.2. Internet Video Consumed (Petabyte per month) until 2019 (Cisco VNI, 2015)

Note that the CAGR has been calculated according to the equation shown below.

$$CAGR = \left(\frac{\text{Ending Value}}{\text{Beginning Value}} \right)^{\left(\frac{1}{\text{Number of Years}} \right)} - 1 \quad \dots\dots\dots \text{Equation 1.1}$$

Where in the previous table, beginning value equal to (1139), ending value equal to (14999) and number of years is 5.

This trend towards an increased consumption of online content suggests that there is a growing need for a system has the ability to help viewers search and find all the available content that matches their viewing pattern and desires. Indeed, with so many content sources now available and a trend towards on-demand viewing, the individual programme will start to become the dominant artefact rather than the more traditional channel through which programmes are delivered. However, even though some sites provide a search tool for their services, there is no centralized electronic programme guide available for the Internet users which can combine all of the available video content sources within the same guide.

1.2 Research Problem

Traditionally, broadcast channels provide an Electronic Programme Guide (EPG) to allow viewers to see what is being transmitted and when. These EPGs are used in the live TV broadcast channels (including terrestrial, satellite and IPTV) to present a time schedule of programmes for a specified time period which might be for one day or span several days. Figure 1.3 shows the typical structure of a conventional EPG that lists channel and programme information against time. Similarly, many catch-up and IPTV providers offer their equivalent version of this EPG. However, all of these current EPG systems are limited in the range of programmes and channels they include and hence, make available to viewers. Due to the proliferation of Smart TVs which integrate the Internet and television, new interfaces have been developed that provide more facilities and interaction with the viewer. However, these are still limited in scope both in terms of the channels they cover and the search options provided (M.-E. Kim, Cho, Yoo, Hong, & Kim, 2013).



Figure 1.3 General Diagram of EPG Provided by Sky

The goal of this PhD is therefore to develop a generic EPG that is able to take as its input any, and all, available online video sources and present to the viewer a recommendation list that is both tailored to their preferences and means of being able to view content. Achieving this will require advances to be made in terms of providing a generic interface to content sources, the flexibility of the search scheme and recommendation process, and the overall responsiveness of such a system. In addition, it is important that any recommendations are tailored to the viewing device being used. For example, a smart TV, PC, smartphone or Tablet are all capable of playing online video content but each has quite different capabilities that do affect what content can be played. This variety creates another challenge to ensure that chosen content is capable of being played on the viewing device and its current level of network connectivity. Therefore, the overarching challenge is to support the viewer with an EPG that provides their preferred programmes regardless of their origin and which are chosen to match their current viewing environment.

1.3 Research Aim

The aim of this research is to determine if it is possible to create a generic EPG system that is able to access all available online video content sources and provide a set of recommendations, regardless of the source, which is better tailored to the viewers' viewing pattern taking into consideration their viewing context and exploiting personal connections resulting from the viewer's interactions with social networks.

1.4 Research Questions

Carrying out this PhD has raised a number of research questions which are:

1. How can a generic EPG system be designed that can interface with all the available online video content sources and include any new source without changing the system architecture?
2. How can a searching system be designed to locate all sources of online video content given the diversity of such sources and the large variation in metadata which is used to describe programme content?
3. Can a unified form of programme metadata be devised that works across all online video content sources?
4. How can social networks be exploited within the recommendation process and what impact could this have?
5. What aspects of a user's context are important in recommending video content and how can these be integrated into the system design?
6. How can the time required to execute the functions of the EPG be optimised to deliver a quality of user experience?
7. How can this generic EPG be designed to serve different displaying devices such as TV's, laptops, tablets and smartphones?

1.5 Research Methodology

This research adopts the scientific methodology based on a hypothesis of whether it is possible to develop a generic EPG system. To achieve this, the following steps have been adopted.

1. Carrying out a Literature Review: which examined published work which is related to online video content providers, EPGs and underlying technologies to understand the research area and construct a clear appreciation of the whole system.
2. Identifying the research problem areas: which emerged from the literature review.
3. Defining the requirements and performance parameters that a generic EPG has to satisfy.
4. Proposing a system architecture for a generic EPG system that has the ability to interface with any online video content source.
5. Developing the searching process, metadata processing and recommendation operation for the generic EPG.
6. Extending the EPG architecture to integrate social networks with a view to enhancing the recommendation process.
7. Evaluating and validating the proposed design: by implementing an experimental prototype for the system to assess its effectiveness at improving the EPG and user's interactivity with the EPG system.
8. Ongoing modification of the system based on evaluation to improve the system performance and to optimise its design.
9. Finalising the conclusions, publishing results and submitting a PhD thesis for examination.

These steps are summarised in figure 1.4.

1.6 Contribution to Knowledge

This section will illustrate the main contributions to knowledge which are being claimed by this research. The first contribution is the generic online EPG system architecture. The EPG system architecture depends on a modular structure and the modules that are connected with the external sources are extensible to include any number of online sources without changing the main structure. That is implemented through adopting a new unified metadata representation for the different metadata formats which retrieved from the online sources where this unified format will be used within the entire system.

The second contribution is exploiting features of social networks in a new way to support the recommendation system. The recommendation system has also been uniquely designed

to employ four recommendation schemes to enhance the recommendation process and obtain a personalized list of programmes for each user.

Finally, the EPG system has been designed as a cloud-based service, rather than a locally hosted service, which can serve any device that, has the ability to connect to the Internet. Therefore, the EPG provides the ability to appropriately match content to the viewing device and its context.

1.7 Thesis Layout

The remainder of this thesis is organized as follows. Chapter 2 gives an overview of the related work which describes the contribution of other researchers in the development of the components that are used in the EPG systems. Based on the experience that has been achieved from the previous research, the system requirements have been proposed in chapter 3. The architecture of proposed system is then described in chapter 4 which contains a full explanation of the different functional blocks which represent the system components. Chapter 5 includes the implementation of the generic EPG system and the initial results that were obtained during system testing. The final results obtained from the EPG system trial have been presented and evaluated in chapter 6. Finally, chapter 7 concludes the work that has been done by this research and identifies relevant future work aspects.

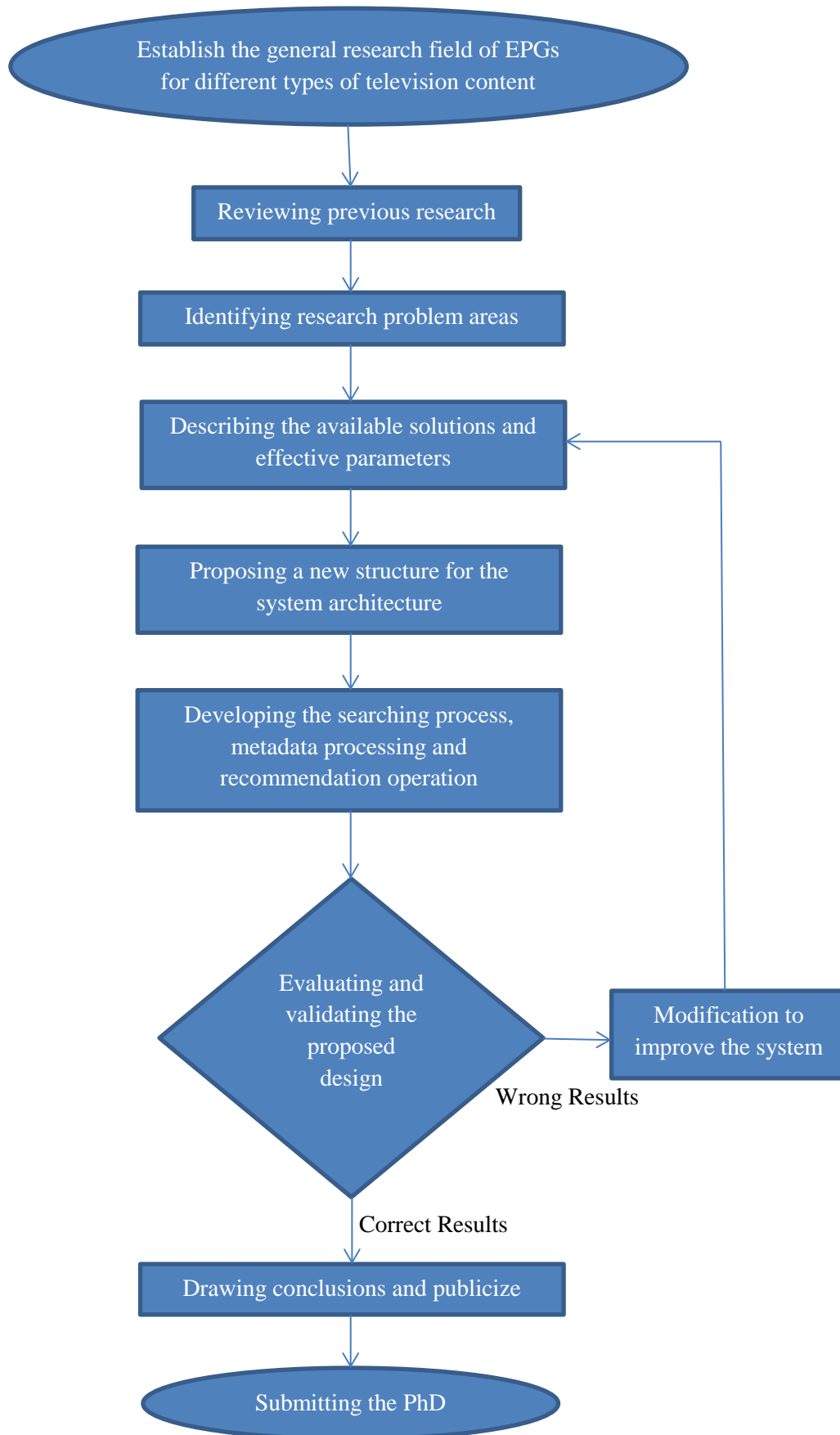


Fig. 1.4 Stages of Research Methodology

CHAPTER TWO

LITERATURE REVIEW

Chapter Two

2.1 Introduction

This chapter will survey relevant research relating to the development, structure and evolution of Electronic Programme Guide systems. Particular focus is given to research which has investigated the different operations and algorithms that are used in improving the performance of the EPG systems. This includes EPG system design for TV and IPTV, their searching, recommendation and context-aware systems together with social network integration.

2.2 Electronic Programme Guide Systems

An Electronic Programme Guide (EPG) is simply a tool for browsing the available information about programmes normally, via a list of channels and usually based on time. This includes the title, duration, a description or picture relevant to the programme. They have been created to help TV viewers find programmes that match their interests without having to navigate all of the available channels and programme providers. However, because EPG information is collected from more than one programme provider, it is necessary to present this information in a unified and standard format. Therefore, attempts have been made to create a standard format to represent EPG information from all content providers. The standards include information about the metadata structure, the encoding and decoding and content metadata of broadcasted programmes which means this information could potentially be processed by any device. The content metadata further includes different information about the programmes and mainly classified into two parts; content description metadata such as title, genre and textual description and instance description metadata such as content location, video format and usage rules (Rosengren, 1996).

Programme metadata could be defined as an organized collection of data items which provide information about programmes, channels, software, interactive elements and other relevant data. Content metadata consists of several elements to represent the basic format, structure and content description information. In order to use this metadata in automated processing systems, a number of conventions about metadata items and their organization that should be defined. The most important one is naming. In order to allow a scalable

interchange and use of the metadata, each item should have a unique name which ensures that merging the metadata collections will occur without errors or duplications.

XML is a standard infrastructure which is used to describe metadata and takes advantages of the similarity between the core idea of XML and metadata description concepts. An XML file consists a sequence of tagged data sections. The tag is simply a name and refers to the external description of the meaning for that section. As an example, < channel > refers to the name of the channel that is broadcasting that particular programme. Each section may contain multiple embedded sections to define a particular metadata item as shown in figure 2.1. Metadata could be described using other types of formats such as JSON. However, programme metadata is usually converted to a different format for reasons of usage space and easily use in the processing (Barton, 2000; ETSI, 2007; Huh & Kang, 2010).

```
<BasicDescription>
  <Title type="main">
    The one where Fox jumps in the Potomac
  </Title>
  <Synopsis>
    Fox goes to Washington and jumps in the Potomac
  </Synopsis>
  <Keyword>Fox</Keyword>
  <Keyword>Washington</Keyword>
  <Keyword>Potomac</Keyword>
  <Genre href="urn:tva:metadata:cs:FormatCS:2007:3.5.7.3" type="main"/>
</BasicDescription>
```

Fig. 2.1 Part of Content Metadata showing the Basic Description

In the field of metadata, it is worth mentioning that Metabroadcast has made a significant contribution in aggregating and organizing the metadata from different content sources. Metabroadcast has offered Atlas as a new service to take over UK listing data from several content sources. Atlas is a metadata aggregator which has built for multiple data sets. It ingests the metadata from several content sources, indexes, matches items and makes a set of data available through a RESTful API in a standard format. It provides the data using XMLTV. Atlas is working with different types of sources which have been classified into three types; international long-form, UK long-form and international short-form. Table 2.1 represents the sources of each type of the content sources that Atlas is dealing with.

However, this service is still limited for a number of content sources and their API is available for only the commercial applications (Metabroadcast, 2017, Atlas, 2009, Chris, 2016 & Jonathan, 2012).

| International Long-Form | UK Long-Form | International Short-Form |
|--------------------------------|---------------------|---------------------------------|
| Hulu | BBC | Archive.org |
| iTunes | Channel 4 | Dailymotion |
| TV Blob | Five | Blip.tv |
| | ITV | Podcasts |
| | MSN Video | TED.com |
| | Press Association | Vimeo |
| | | YouTube |

Table 2.1 List of sources that connected to Atlas (Metabroadcast)

With an increased number of channel and programme providers, EPG data that can be retrieved has increased. This has led to an increase in the management and analytical operations needed to extract relevant programme information from the EPG and display it to a user. Additionally, the required connection bandwidth must also increase to cover the EPG data for seven days or across multiple providers. Stettner produced a technique for client terminals such as set-top-boxes (STB) where requests to an EPG server could be made to download only the needed part of the EPG, instead of having to download the entire EPG. Then if the user wishes to view any additional information, the STB can issue a request to download the other parts of the EPG from the server. However, providing the EPG in this manner does limit its ability to interact with the users and increase the needed time if the user does a request for a particular programme or category (Stettner, 2012).

Recently, new technologies have been employed to add services to the existing EPG to improve a user's satisfaction and offer more facilities that could make watching TV more interesting. Ellis presented an EPG system with a scan feature. The system utilizes the television receiver with a tuner to tune the receiver to a selected channel and a data

processor to parse and store the television programme's schedule information. Then the system offers a service to the user which can scan the programme schedule list for multiple programmes with a single command (Ellis, Davis, & Knudson, 2014).

Generally, generating and implementing the EPG for digital TV or IPTV is based on receiving information that originates from TV content providers. Ideally, this information should be delivered to the receiver side in a standard format used by all providers. An EPG is a software application running on the receiver device such as a set-top-box (STB). The design and implementation of these EPG systems and graphical user interface (GUI) depends on the designer in choosing the appropriate software. Cui has designed an EPG system and GUI which is based on a Linux embedded operation system that is used within a digital TV set-top-box. The major functions of the EPG system are programme management, programme searching and system setting. However, this EPG depends entirely on the information from the channel providers and without any consideration of the users' needs and preferences (Cui, Sun, Wang, & Li, 2012; Dajda, Ciślak, Heldak, & Pacyna, 1996).

The development in EPG services has continued by proposing more facilities and offering more personalisation and interaction with the user side. Another EPG design has been proposed to present the recent and future TV programmes with an additional list of the past programmes. The designed EPG offers to users the ability to view the past programmes which have a link with the recent and future programmes in EPG such as series or parts of movies. Therefore, and in order to give the users another chance to watch the missed programmes, a list of past programmes is provided for those programmes that will be repeated in the coming few days whether on the same channel or another channel. Additionally, the viewers can select any past programme and the system automatically notifies the viewer when a repeat of that selected programme is about to be broadcast. This system is supported with a user interface which has the ability to receive and forward information to the system and includes management and recording options (Drazin & Kram, 2014).

Recently, broadcasters and providers have developed a hybrid TV platform which integrated the free-to-air digital terrestrial channels with TV on demand and catch up services. Youview is an example of these hybrid TV platforms in United Kingdom. It combines two main things; over 70 digital terrestrial channels (including BBC, ITV,

Channel 4 and Channel 5) and catch up and on demand services via a broadband connection. Moreover, it offers an EPG for users with streaming applications to particular providers such as Netflix, Now TV and Amazon Instant Video. Youview is available either as a standalone set-top-box or as part of a TV package from the partnership broadband providers. The provided EPG looks like a normal EPG but the user can watch previously broadcast programmes rather than just what is being broadcast. This EPG has deep links with the catch-up services from the major Freeview television channels and on demand providers such as Netflix and YouTube.

Freeview Play is another hybrid TV platform example in United Kingdom. It adds connected services to the free-to-air digital TV within the EPG. The users are able to scroll backwards through the last seven days of programming within the EPG to watch the programmes that have been missed. It has connected to BBC-iPlayer, ITV Player, All 4, Demand 5 and UKTV Play. Moreover, the users can browse forward to see what's coming up to watch in the seven days ahead where it connected to over 70 free-to-air digital channels. However, for the catch-up EPG, there is no search functionality. Therefore, if the user doesn't know the broadcaster or date of what they want to watch, they can't just search for it within the EPG. In fact, this TV platform might give the impression that the EPG has integrated different sources into one EPG, but it has not because it just has links with the catch-up services.

However, there are several limitations for using this type of hybrid TV platform such as it needs a stable and fast download speed (minimum 3 Mb), it needs extra cost for the set-top-box and the limited number of channels that connected to the TV system (Youview, 2017, Kim, 2017, Andrew, 2012, Freeview, 2017, Rik & Max, 2017 & Andrew, 2017).

2.2.1 Electronic Programme Guide for IPTV System

Internet Protocol Television (IPTV) provides TV services to users through the infrastructure of an IP network. That means, you can watch television if you have IP orientated device which can connect to the provider's network. Recently, the set-top-box (STB) has been used to receive IPTV services from providers such as Virgin Media. This is different from Internet TV which does not provide any the Quality of Service guarantees for TV programmes, has no geographical limitation and is widely popular (Yim, Lee, Lee, & Jeon, 2014).

Recently, the propagation of IPTV has been led many researchers to more investigate and develop the performance and quality of service that offered to the users of the IPTV system. In the same way as for digital TV, it has become quite difficult for users to find the programmes that they are interested in due to the increase in available TV programmes and channel providers. Therefore, it has similarly, become necessary to support IPTV with the development of EPGs to help users find their preferred programmes.

An approach proposed by Concolato generates an EPG in a separate device different from the rendering device. It can serve IPTV and Terrestrial TV users, where the server received at first the programme metadata through the broadcasted channels and transformed the raw metadata into a presentation data. Then this presentation data is driven from the server to the EPG generating device. The generating device converts the received data into a presentation format which is streamed to the rendering device. This system provides the EPG to the user device as a client, therefore, the user device does not need to analyse the metadata itself. Moreover, it is possible to update the EPG presentation and change the look of the EPG when the user decides. However, this design has some disadvantages in that it requires more processing on the server side, the adaptation of the presentation with the terminal characteristics of different rendering devices is difficult, the user cannot control the displayed programmes to show a specified category or keyword and finally because the metadata is transformed into the presentation style, the transmission of the EPG requires high bandwidth (Concolato, 2009).

Another IPTV EPG system has been produced by Macedo to collect web data relevant to the programmes that are listed in the EPG. Firstly, the system classifies the programmes in the EPG into specified categories. Then the system integrates the gathered data with the information provided by the programme providers. The web data is gathered from several Internet sources and it contains textual information about the programmes that are listed in the EPG such as related news articles. The system extracts these news articles from a set of predetermined sources. This enriches the information that the EPG provides to users and helps them to make better decisions about whether they would like to watch a given programme or not (Macedo, Cardoso, & Pinto, 2014).

An EPG system proposed by Tan is based on an exchange of information between IPTV users to share their experiences and list of preferred programmes. The system consists of an EPG recommender engine, information exchange module, message module and filter

module. The EPG recommender engine is responsible for recommending programmes to the users based on their profiles and preferred programmes. Then, in order to improve the user experience and provide more choices, the information exchange module is designed to send requests to the other users through the message module. The message module is responsible for sending requests to the other users, receiving their responses and parsing them to specify the users that should be provided with the shared EPG. The information exchange module supports several ways to exchange EPG information. The user can choose not to exchange their EPG information which means that they will lose the system benefits. Alternatively, they can choose only to exchange their programme information with some of their friends by sending requests to specified users. Finally, they can share programme information with all users in a totally unrestricted manner. The shared programmes are then forwarded to the filter module to filter out the programmes that already exist in the user's own EPG (Tan, Zhu, Wang, & Chang, 2009).

The open IPTV system is another contribution that allows users to choose an EPG from multiple EPGs which belongs to a number of network providers. This system depends on giving the user the ability to access other IPTV service providers which together create what is called a 'neighbourhood garden'. The open IPTV strategy allows users to change to another EPG which belongs to a different IPTV provider. Users do however, need to register with at least one default service provider who then provides the primary contact, bill management, security, other service connection and login records management services. However, service providers seem not to prefer the open IPTV platform (T. Kim, Kim, & Hahn, 2010).

An Internet television's programme guide system has been proposed by Boyer which integrates electronic mail (e-mail) services. This system allows users to order and receive e-mail message reminders of selected television programmes and events. Moreover, the user can control when and how the reminders are generated and received, adding new programmes to the list and canceling previous reminders. Additionally, the user can select the type of reminders for free to view programmes or pay-per-view depending on the selected channels (Frankin E Boyer, Demers, & Blackwell, 2014). In the same field of research, McKissick developed an interactive television programme guide service which allows the service providers to deliver television programme listing information to the user's local set-top-box. Using this service, the user receives notifications about the

programmes that are outside the current display time frame. The user can then specify the type of notifications related either to the programmes or the time of broadcasting. The EPG offers the ability to manage the notification through listing, adding, canceling and changing type of notifications. Moreover, the EPG can provide any available information about the user's selected programmes even those programmes that are not in the TV channels' time schedule (as example a movie in cinematic release). Figure 2.2 shows the notifications that are provided to the user based on previously selected programmes (McKissick & Forrer, 2013).

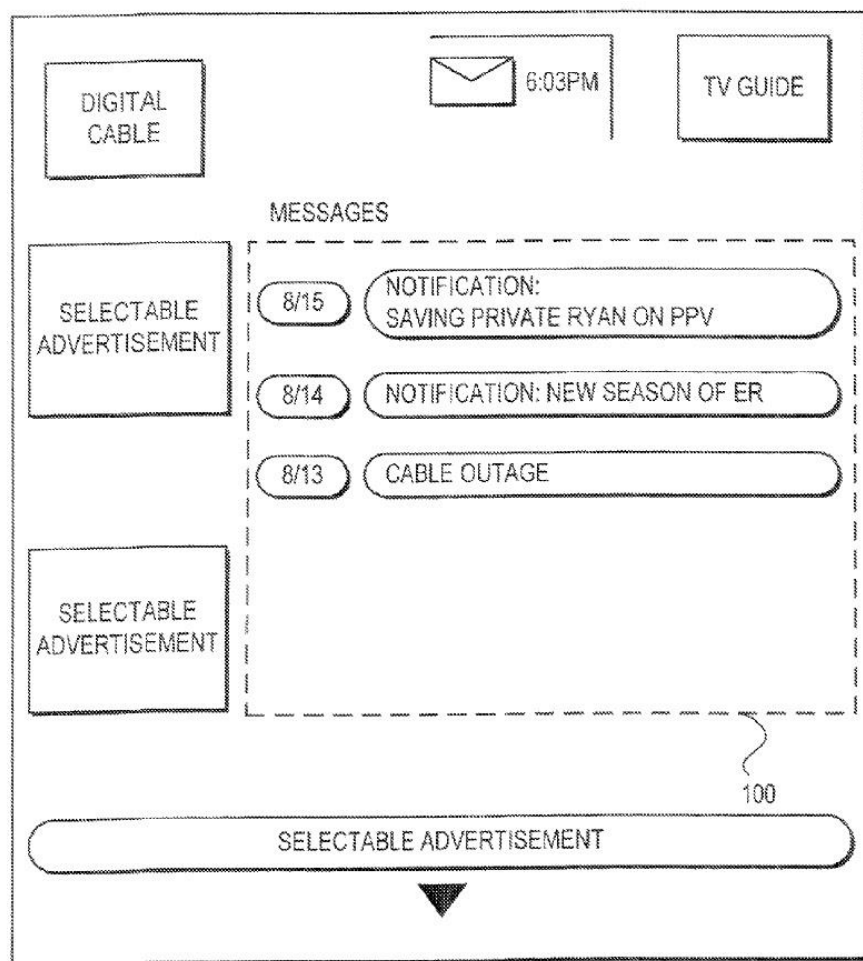


Fig. 2.2 User Notifications Service in the Developed EPG (McKissick & Forrer, 2013)

2.2.2 Personalized Electronic Programme Guide

With an increasing number of channels and programme providers, it has become difficult for users to find programmes of interest despite the existence of EPGs. This is because with traditional EPGs, the user needs to navigate through the whole EPG to find something interesting to watch and this wastes time. On the other hand, if a user becomes accustomed to watching certain kinds of programmes in certain channels, they might not experiment with different programmes in other channels which could also be of interest to them. Therefore, this problem needed to be addressed and one approach is to personalise the EPG based on a user's viewing history from which the EPG can recommend a list of programmes which is appropriate to their preferences. This new generation of EPG does not just list the information which is provided by content providers but can also process and recommend a list of programmes to users.

Another personalized EPG has been presented for digital TV and runs on the user's set-top-box. It creates a user model based on information supplied by the user. The system depends on three sources of information to manage the creation of a user model and these are the user's explicit preferences that should be declared by the user, information about the TV viewers classes which classified based on programmes' categories that are preferred by viewers and the user's viewing behaviour. Having completed the user profile, the system creates the user model which is then used in recommending programmes to the user. This is done by matching the programme information which is extracted from the TV stream and channel's EPG with the user model. Then the resulted programmes are ranked based on viewing behaviour (Ardissono et al., 2004). However, this proposed system doesn't have the ability to adapt to any update in the user viewing behaviour or preferences especially for events which only occur occasionally. Moreover, the user experience can suffer from over-specialization and hence, they will not experience any new programmes which may be of interest to them.

Smyth has proposed a web based personalized EPG. The user needs to create an account on a web server and register to create his profile comprising, for example, TV preferences, preferred viewing time, and genre preferences. The most important information is grading feedback provided by the user. Then based on this, the system recommends a list of programmes and their transmission times via an EPG page. The recommendation process depends on content-based and collaborative filtering techniques which help the system to

overcome the disadvantages of each technique individually. More details about those recommendation techniques will be described in the next sections. An overview of the system is shown in figure 2.3 (Smyth & Cotter, 2001).

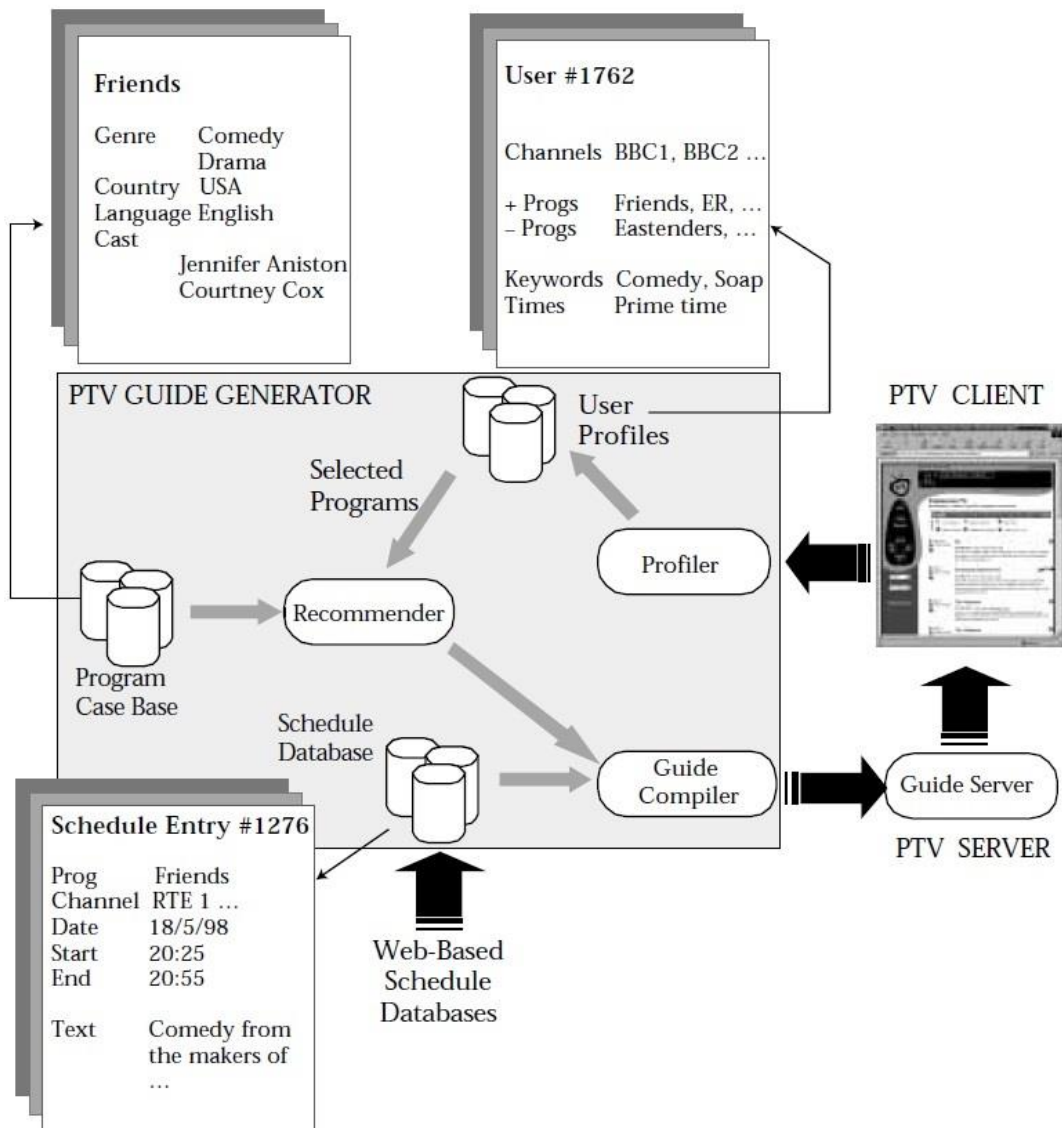


Fig. 2.3 Architecture of Personal TV Guide System (Smyth & Cotter, 2001)

Whilst personalized EPGs use different schemes to recommend programmes, it is still difficult to produce a precise list of programmes for each user because of the limited profiling information that is available. Sullivan proposed a personalized EPG that uses data

mining techniques to overcome the problem of sparsity in the collaborative filtering technique. That is because users are usually rating a small number of available programmes therefore, the expected rated programmes overlap between two random users is very low which cause the sparsity problem. Sullivan's system employs a user profile and information filtering techniques to learn about user preferences and personalize the recommendations according to their needs. Data mining methods have been applied to extract new programme metadata from user profiles in order to recommend similar programmes. However, this system doesn't address the other problems in the collaborative filtering technique namely, the first-rating problem and cold-start problem (O'Sullivan, Smyth, Wilson, Mcdonald, & Smeaton, 2004).

Another proposed system for personalizing the EPG was presented by Chen. This EPG is integrated into the viewing device and automatically monitors watched programmes and records the user's viewing habits. The users' feedback is then collected and forwarded to a community' server in order to cluster the users into several virtual communities. The EPG system comprises content search, content recommendation and content management modules. The content search module is responsible for providing programmes that are simply requested by the user via keywords or by selecting programme categories. The content recommendation module recommends programmes using a community-based recommendation process in which the system monitors the viewed programmes from each user and divides the users into different virtual communities based on their viewing habits. The content management module records the users' feedback about the recommended programmes and then processes and updates the recommendation process based on new information provided by the content management module. Figure 2.4 shows the block diagram of the community-based programme recommendation for the EPG system (Y.-C. Chen, Huang, & Huang, 2009).

Based on this existing body of research, in order to design a personalized EPG system which can interface with any video content source and provide personal recommendations to each individual user, major functional parts are needed. These comprise a searching system that can interface with any content provider, a recommendation system that offers a list of programmes that is matched to the user's viewing behaviour and preferences, and an ability to create, store and manage user profiles. Also, an EPG system should have the ability to extract and analyse the data collected from the social networks. Therefore, the

next sections of this chapter discuss research work relevant to these specific aspects of EPG design and operation.

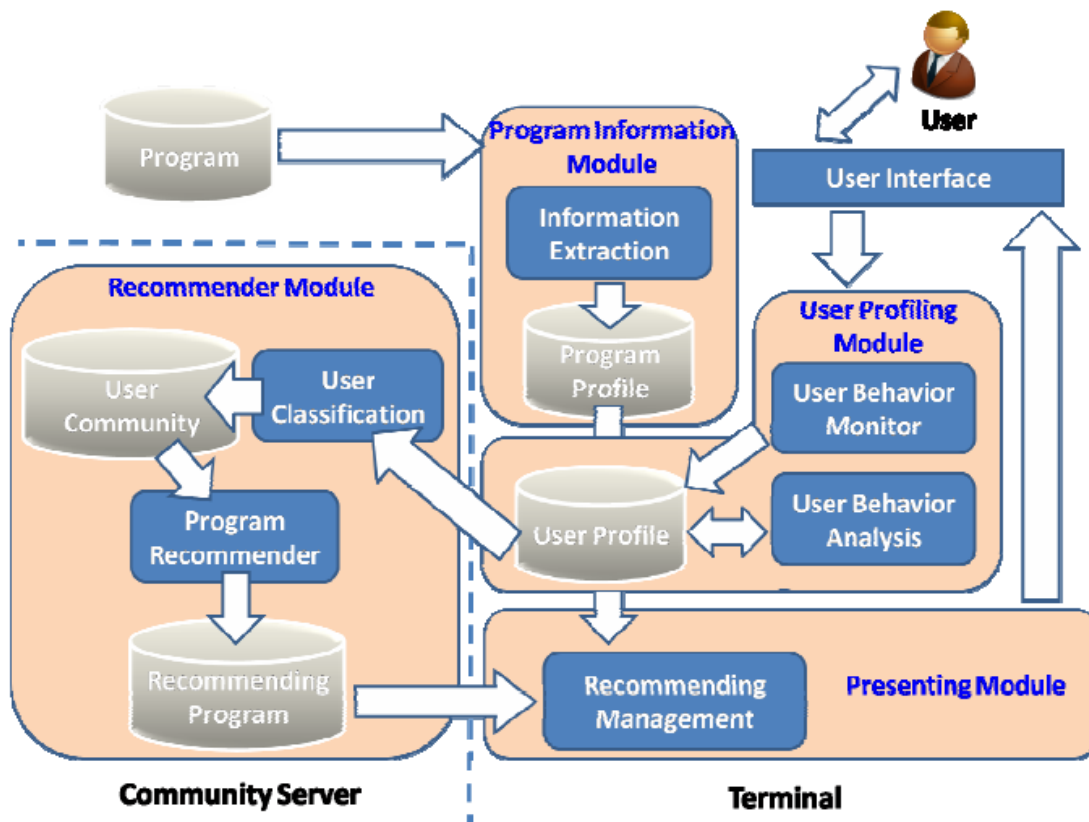


Fig. 2.4 Community-Based Programme Recommendation for EPG System (Y.-C. Chen et al., 2009)

2.3 IPTV System

Many researchers have discussed EPG design and implementation in the IPTV systems. This section presents a brief explanation of the IPTV system and its structure as an introduction to discuss the EPG within IPTV systems.

An IPTV system is a digital television service that is delivered over a dedicated provider network using an IP networking protocol. This system works on a TV with a supplementary device called a set-top-box which accesses channels and other multimedia content through a broadband connection which connects to the provider's network. It is different from Internet TV or Internet video streaming. Hence it is a secure, subscription service, local (limited operator coverage), that provides an end-to-end operator managed

service which guarantees the quality of service (QoS) over the provider's network. However, IPTV providers may also include Internet services such as Web access and Voice over IP (VoIP) in which case they become known as Triple Play.

The typical architecture of IPTV service comprises: a Super Head-End, Core Network, Access Network, Regional Head-End and Customer Premises. The Super Head-End encodes and processes the channels and transmits them to the Regional Head-End through the provider's core network allowing it to serve millions of subscribers across their serving area. The Regional Head-End handles local programming, delivering channels to a single city or geographic area. Programmes are then transmitted as data streams to the subscriber's premises through a range of access networks including Cable and Digital Subscriber Line (DSL) (Martinsson, 2006; Punchihewa, De Silva, & Diao, 2010). Figure 2.5 illustrates a simple structure of the main functional parts of an IPTV system.

IPTV services could be classified into four domains which are the consumer domain, network provider domain, service provider domain and content provider domain. The consumer domain presents the services to the end user while the network provider domain permits the connection between the consumer and service provider domain. The service provider domain is responsible for providing services for all subscribers whilst the content provider domain possesses licenses to sell or distribute the content. Currently, IPTV provides subscribers with linear TV programmes in the same way as traditional through the air broadcasters. However, many also supplement this with additional services including video on demand (VoD), personal video recorder (PVR), EPG and a traffic and weather service (Zeadally, Moustafa, & Siddiqui, 2011).

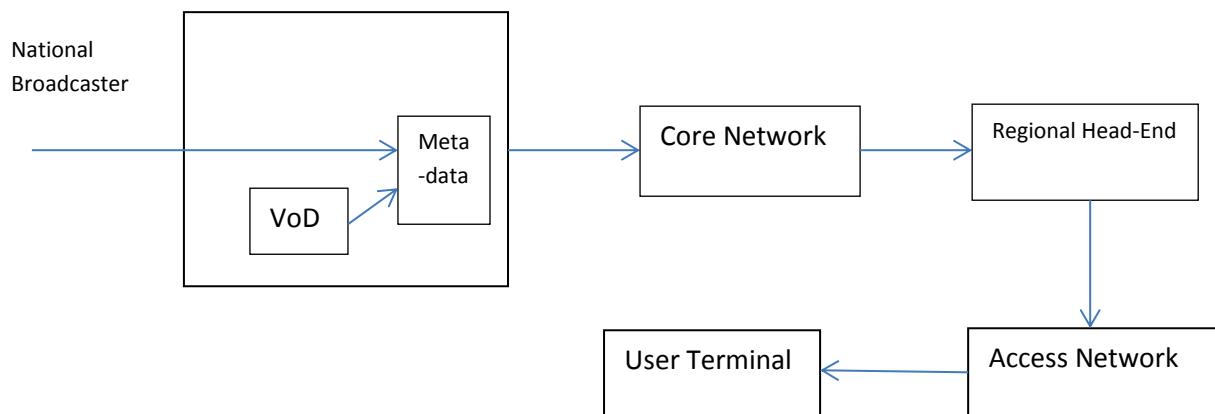


Fig. 2.5 Main Functional Parts of the IPTV System

Many researchers have discussed the performance of IPTV systems and how to improve the quality of service provided to users. Due to the requirements of an IPTV in terms of its performance and reliability, network bandwidth becomes critical (Mahimkar et al., 2009). A study has been carried out by Simsarian to discuss the traffic and required bandwidth that must be streamed to each user. It was found that the bandwidth depends on where Video-on-Demand (VoD) content, its cache and streaming server is located. Therefore, in order to decrease the traffic on the core network and the required bandwidth, VoD content is normally cached in the Regional Head-End (Simsarian & Duelk, 2007).

Due to the popularization of IPTV, many services have been added to offer more facilities for the users such as on demand and interactive multimedia services. With an increase in the number of channels and content providers, assisting users in making selection becomes crucial. However, creating the EPG in the user device or centralizing the generation of EPG in the main servers can cause more challenges especially when the user experience and available bandwidth will be discussed (Cha, Rodriguez, Crowcroft, Moon, & Amatriain, 2008).

IPTV services have been delivered to users through mobile devices to watch TV content based on a cloud network. Lai proposed a personalized mobile IPTV system which allows a user to configure a browsing interface with EPG information and select their favourite TV programmes through their mobile devices. Then, based on those selections, the manager server searches for multimedia videos in the cloud within the User Generating Content web site or VoD server. The management part of the system produces the list of content to match the EPG information and metadata parser engine. Based on the degree of similarity, the system chooses the most appropriate videos and presents them to the users through their mobile devices. This system is shown in figure 2.6 and comprises a management server, which includes the multimedia parsing engines and is connected to the VoD server, and a personalized user interface which is used by the user to select their favourite videos (Y. X. Lai, Lai, Hu, Chao, & Huang, 2011).

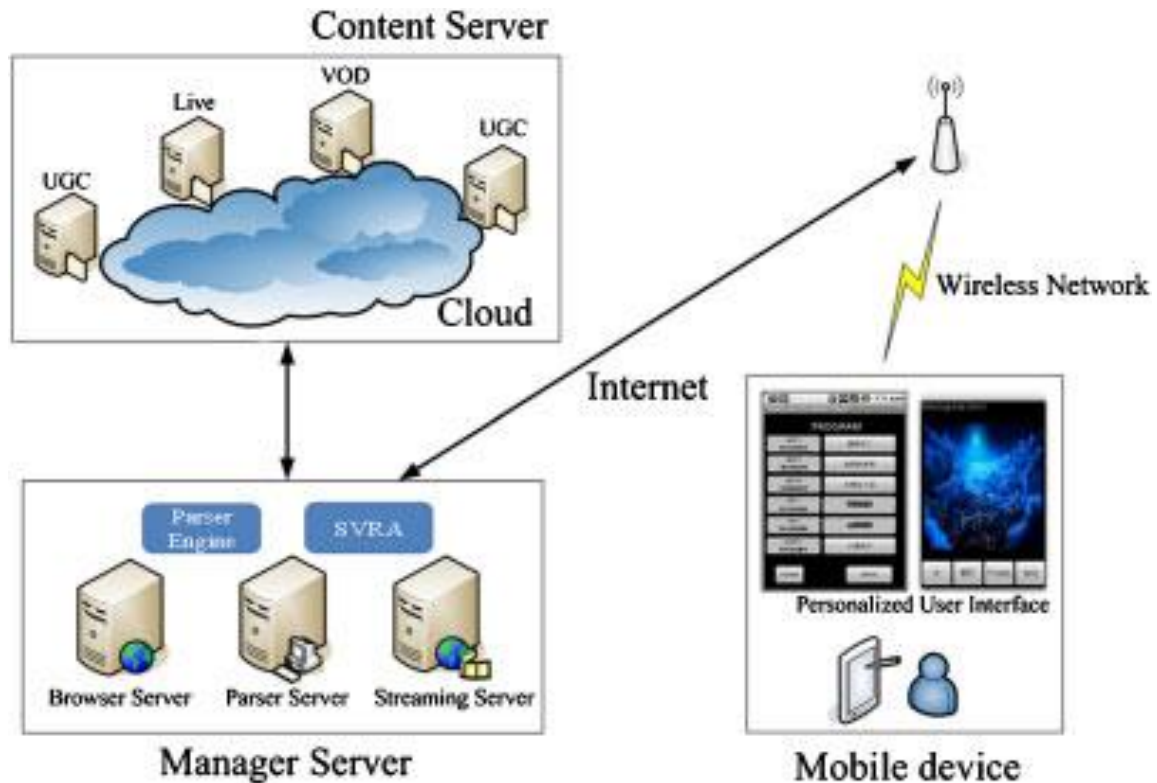


Fig. 2.6 Overall Structure of personalized mobile IPTV system (Y. X. Lai et al., 2011)

Many other researchers have discussed different aspects in the IPTV systems such as improving the EPG via more interactive services, user experience and improving the system transmission performance and quality of services to decrease the factors that affecting the service such as Jitter or packet loss. However, there are still many limitations relevant to improve the EPG especially the limited content channels and sources and difficulty of recommending an accurate personalized list of programmes that match the user viewing behaviour and wishes.

2.4 Searching Systems

Searching and retrieving the correct programme metadata is based on how the programme search module can interact with the range of content providers. Therefore, many researchers have discussed how to improve the searching operation. Moreover, due to the variety in the type and format of programme metadata delivered by different providers, search systems have to adapt with these types and formats. This section discusses some of

the research that has investigated how to improve the performance and accuracy of search operations.

2.4.1 World Wide Web Searching Systems

Searching the Web for information is managed through popular search sites such as Google. Using a non-real time process, these sites examine and categorise each web page they can find and build a database which is then searched in real-time when a user enters a search term. Such search engines are then evaluated based on factors such as speed, searching effectivity and ease of use. However, searches for images or non-textual content is more challenging. Smith has developed a Web searching agent to search images and videos by incorporating text-based searching and content-based technology. Content-based technique provides for automated estimation of salient visual features such as colours, shapes and spatial information contained in the visual scenes. Then, similarity calculation applied between images and videos based on these extracted visual features. The system retrieves images and videos from the Web and provides a tool to browse this collection. Then, the user needs to select and specify the most and least item from the resulted list that match his desires. Based on the user selection, the system will reformulate the searching and processing operations to better match with user's instruction. The system has several functionalities such as the search operation is based on content-based technique, updating the search operation using the user feedback and merging the text-based and image and video subject search and navigation (Smith & Chang, 1996).

A more complicated searching system has been developed in order to increase the accuracy of searching results. Liddy proposed a system that includes a language processor to determine the subject categories and important terms of query. Moreover, it includes an agent server which generates and trains a neural network according to the language processor query. This neural network is embedded in the search agent. The search agent server processes through the neural network the subject categories and query terms of each retrieved document. Then the addresses of the retrieved documents which are above a threshold retrieval value is displayed on a user interface. The agent server automatically selects which of those retrieved documents are more relevant and expands a search agent according to the selected documents to improve the ability to retrieve the documents that better match the query terms. The second group of resulted addresses is displayed on the same user interface (another displaying area in the interface) where the user is able to

select any one from the two resulted lists. Based on user selection, the system applies and evolves search agents to retrieve more documents relevant to the selected one and the same operation will be repeated (Liddy & Yu, 2001).

A searching system has been proposed by Tu to search the Web for information based on intelligent information retrieval agents which are designed to overcome the information overload problem. This problem arises when the user searches the Web for particular information to find too much relevant information and web sites. However, people often become frustrated due to the little useful information that is contained in the searched results and web sites. The system architecture is based on dividing the functions between several agents where one or more agents can cooperate to perform one task or function as a module while one or more subagents may cooperate to perform a feature of the agent. Each agent has a specified feature (task) such as intelligent searching where the agent can interact with user preferences for each category based on initial knowledge about user categories and filter the resulted pages to pinpoint exactly the user needs.

The system depends on creating a profile for the user based on information relevant to their preferences obtained through the personal agent. In the same time the system categorizes users into groups using its group agent. Based on the collected information from each user profile, initial knowledge will be built to be used in filtering the search results to improve what is presented to the user. Additional services offered by the system include auto-notification sending a notification to the user when any update occurs in the retrieved pages, navigation guide to help users to find the way for relevant pages and information management when updating the user profile based on the selected pages. An overview for intellegent information retrieval agent is illustrated in figure 2.7. (Tu & Hsiang, 2000).

Another searching system developed by Funk is designed to search programmes within Internet based VOD services. This system is based on saving the previous searching keywords in the system servers, then performing a search using those keywords and sending the results to the user device to be displayed without user intervention. This system applied for providing the user with any updating information about the programmes that the user watched previously automatically without the need to request a new search operation. Then, the user can select one of the programmes in the automated search results or execute new search operation based on new keyword entered by the user.

The system used a client-server model and works equally well on TVs and tablets (Funk & Noffsinger, 2014).

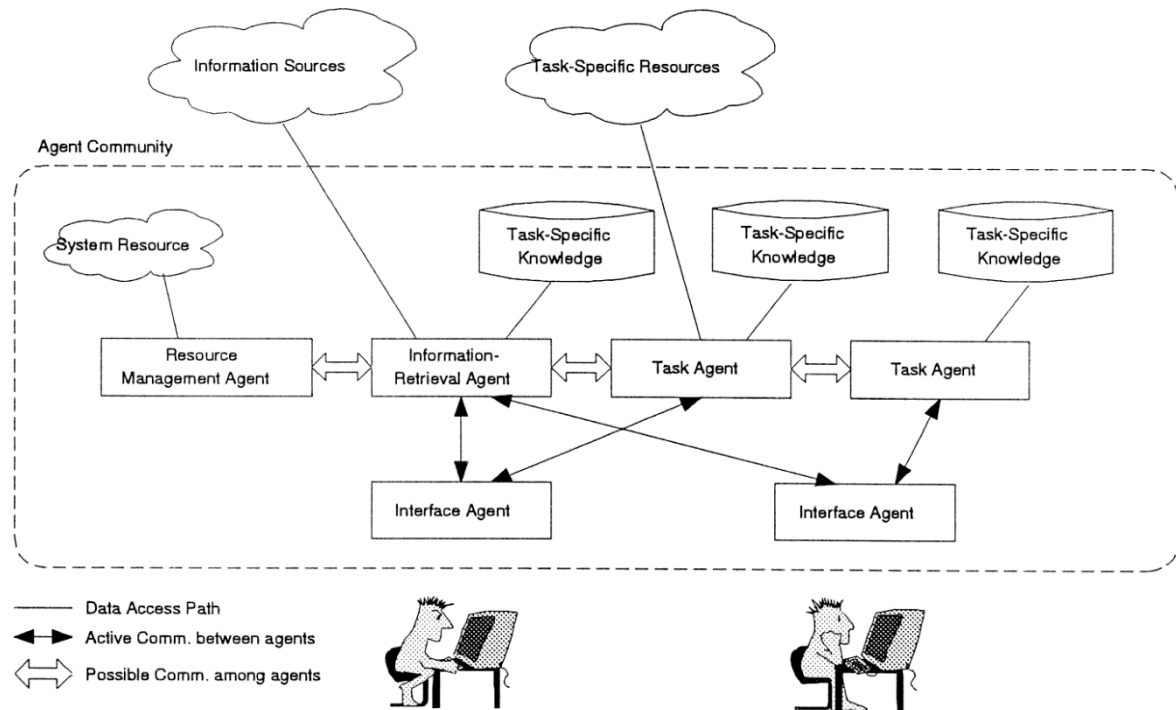


Fig. 2.7 Intelligent Information Retrieval Agent (Tu & Hsiang, 2000)

2.4.2 Multimedia Searching Systems

Many systems have been designed in order to search for multimedia content whether that is searching and retrieving programmes or information relevant to programmes. These systems use different methods to analyse the query terms and obtain accurate results for the searching operation. A system has been designed by Coden to search different multimedia data types that are relevant to a particular query term in a multimedia database. The system provides the user with a collection of multimedia documents including a mixture of media types such as images, text and video clips in a single list. This system uses separate search engines to search the different media type where each search engine supports an interface to each type. The search engine searches the database for a particular media type and then translates the results to a query type that can be understood by the application programming interface that is designed for that search engine. The query

objects are represented by four elements which are the part number, document number, rank and media type. All these objects are combined into a single result list according to the user requirements (So, Coden, & Mak, 1999).

Mankovitz has designed a system and method to search and extract information from a programme database and to use this information in a video recording device. The system could be built into the television or as a separate device. This system consists of a decoder, processor, controller located in the remote control device, memory and infrared transmitter. The decoder decodes the encoded programmes' data after receiving it and converts it into an accessible format and then saves this data in the device memory. Then the processor analyses the data to extract the needed information such as the channel, title, date, time and length. The user can access the programme's metadata to select the programmes that should be recorded whether they are currently being broadcast or scheduled in the future. Thereafter, the system saves the relevant information with all the selected programmes in the device memory. When the time and date of the selected programme coincides with real time, the processor sends a command to the controller by using the infrared transmitter to change the channel if it is not the same channel and start recording the programme. This system is entirely depending on the received programmes' metadata which is sent by the broadcasters (Mankovitz, 2014).

Digital systems now offer hundreds of channels and thousands of VoD movies. Although an EPG offers time schedule information for the available programmes in a given set of channels, it is still hard to scan all the channels to find programmes of interest. A searching system has been presented as separate device by Jerding to search the digital broadcasting channels and VoD for a specified user input. This system consists of a processor, communication components and memory. The system is connected to the receiver to download and process the programme metadata and then offers a user interface to enter the search criteria. Then, based on the user input which could be a channel, title, or time the system will search the stored media information and provide the user with a search result that is related to their input. (Jerding, Rodriguez, & Banker, 2013).

Another approach has been proposed to search directly the EPG data for a particular query term. The user needs to enter text into the search window to search the EPG data and decides whether the search operation should be for channels or both channels and content. This approach has been applied to an IPTV EPG to help users find the channels and

content that are relevant to a specified search request (keyword) among the available channels in the EPG. Figure 2.8 shows the block diagram of the designed system (Rao, Sloo, Danker, & Nyako, 2015).

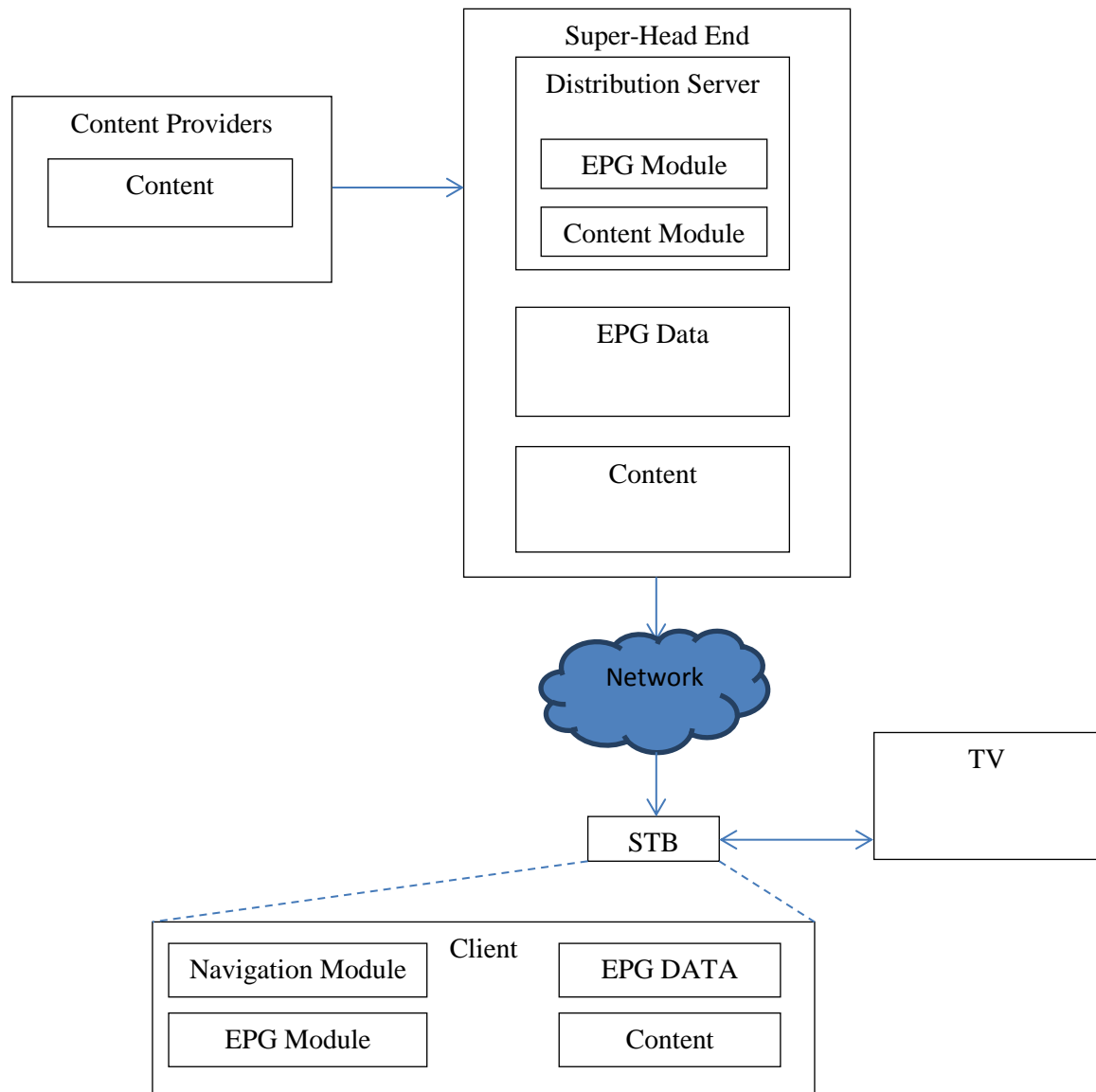


Fig 2.8 EPG Searching System (Rao et al., 2015)

In order to increase the interactivity between the television set and users, a searching system has been designed to search the programme metadata for relevant programmes using the current and last viewed one. This system monitors the user's viewing activities and saves the attributes of the programmes that are viewed. Then, the system searches the

database of the broadcast programmes to find suitable matches which most closely match the programme attributes. At the same time, another search operation is implemented to find programmes that match the attributes of currently viewed programmes. The system may suggest a list of programmes based on the current viewed programme, the last viewed programme or in general based on the programme that has recently been viewed. A neural algorithm is used to make suggestions and rank the programmes based on those that have the most matching attributes (Franklin E Boyer & Demers, 2014).

2.4.3 Integrating the Internet with Multimedia Searching Systems

Miyamori proposed a search engine that can search for Web content and recorded TV programmes (stored in a PC) in the same search operation. This system can handle Web content as an information source and TV programmes as another source which is considered as a first step towards developing the next-generation of search engines. The integrated search is achieved by generating integrated indices depending on keywords extracted from the TV programmes and Web content. Thus, when the user is watching a TV programme, the system searches the Web for content with content relevant to that programme. In contrast, when the user is browsing the Web, the system searches for TV programmes that have correspondingly relevant content. By selecting another item, the system updates the search results based on the selected item (Miyamori et al., 2006).

Kulas also presented a search system that integrated searching an EPG and information resources on the Internet. A user interface has been created to select filtering elements for the search operation and to display the results. The system has the ability to search information contained in the EPG based on the selected elements and extract appropriate information such as title or subject to then use as search criteria on the Internet. Then the system will search the Internet and provide a group of Web sites represented by URL's which are relevant to the search topics. The user interface displays both the EPG results and Internet information for the same search topics represented by web sites in separated windows. The user is then able to select programmes to watch directly through the Interface or to open Web sites in the supplementary window (Legall, Masli, & Kulas, 1999).

Internet search systems have been developed for television to support the user with additional information relevant to the programme that is currently being viewed. This system identifies search keywords extracted from the programme that is currently viewed

by the user. These keywords may be extracted from the closed captions (such as news titles or programmes short descriptions) or the EPG data which represents the programme metadata. Then, the keywords are presented to the user as recommended search queries via the user interface. When the user selects one of these queries, the system sends the keyword to the Internet and displays the resulted web page links. Moreover, the system allows users to refine the recommended keywords by offering the category as another choice. Therefore, the user can choose a category which better represents the targeted information. Additionally, the system offers a supplementary user interface for wireless devices such as tablets to enable the users to access the Internet recommendations as shown in figure 2.9 (Messer et al., 2008).

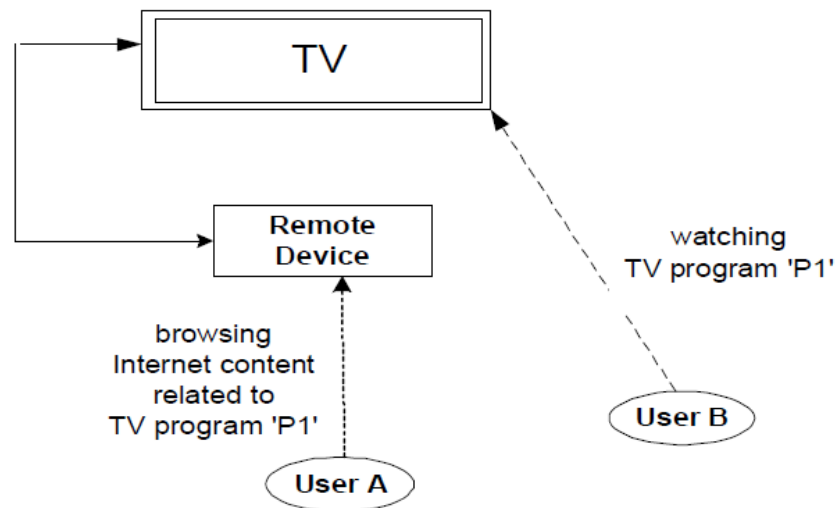


Fig. 2.9 Supplementary User Interface of Internet Searching system on TV (Messer et al., 2008)

2.4.4 Smart TV Searching Systems

The Smart TV combines a TV system with the Internet. Many researchers have discussed the methods to improve the searching system in order to provide additional information relevant to the viewed programme or to obtain more precise results that are relevant to the search query term. A new searching system has been proposed by Rothschild which provides content relevant on the actor, programme title and so on. These contents could be

movie, television programme, web content (information) or e-commerce website. The searching results are displayed to the user in a different window and stored in the system memory, so the user is able to view this content later. The system is designed to work as a separate physical device connected to the Smart TV and includes memory, processor, infrared transmitter and user interface. However, the system designed to serve the Smart TV viewers through providing different search methods. These methods include providing the viewer with multimedia relevant to the programme that currently viewed or searching based on a query term entered by the viewer through an input device (such as keyboard or remote control) represents actor name or commercial product (Rothschild, 2011).

To improve the searching operation in the Smart TV, another approach has been proposed to minimize the human efforts in searching multimedia content. This approach is based on using the mouse as a user interface to the Smart TV. The system analyses the movement of the mouse cursor while the user is navigating the initial results to define the intent of the user. By observing the mouse movement pattern, the system predicts the user's search intent and generates a new list of multimedia content. The new list contains more relevant content to the user search intent (Jeong & Lee, 2014).

Another system has been proposed for Smart TVs that is based on a multi-level searching process. The system sends keywords that have been entered by the user through their Smart TV browser to the searching system to search the local repository for matching content as shown in figure 2.10. Thereafter, the Smart TV agent server will send those keywords to the EPG repository to search the scheduled information of programmes for any matched content. At the same time, these keywords will be sent to the global repository to search the User Created Content website (UCC) for any matches. The results that are received from the other repositories are combined with the local repository searching results. Then, all of the results are ranked, indexed and displayed to the user through their Smart TV browser. However, the system is still based on the user providing keywords (M.-E. Kim et al., 2013).

Hong developed a personalized refinement technique for more accurate searching of multimedia content in a Smart TV. His system is based on the semantic analysis of the query term, user information and multimedia content. The system measures the relatedness in terms of similarity between the user query, multimedia content and user information. This system consists of a user analyser, content analyser and personalized refinement

module. The user analyser saves the user preferences as viewing history and receives the user query term in order to search for the targeted multimedia content. The content analyser extracts the multimedia content information such as the title, category and description. The personalized refinement module is responsible for calculating the similarity between the two information groups, then filtering the irrelevant content and removing them. Several functions included in the personalized refinement module such as grouping and ranking the results are based on the semantic relatedness (Hong & Lee, 2015).

Based on the previous research, in order to create an accurate search system, it is clear that the search system should adopt many analysing operations and information collecting processes. Moreover, the search system should be followed by a recommendation system in order to select or rank the relevant items (programmes) based on specific criteria and filter out the irrelevant programmes. Therefore, the next sections will discuss research that has investigated the development of recommendation systems and the collection of user information which forms a key part of context-awareness.

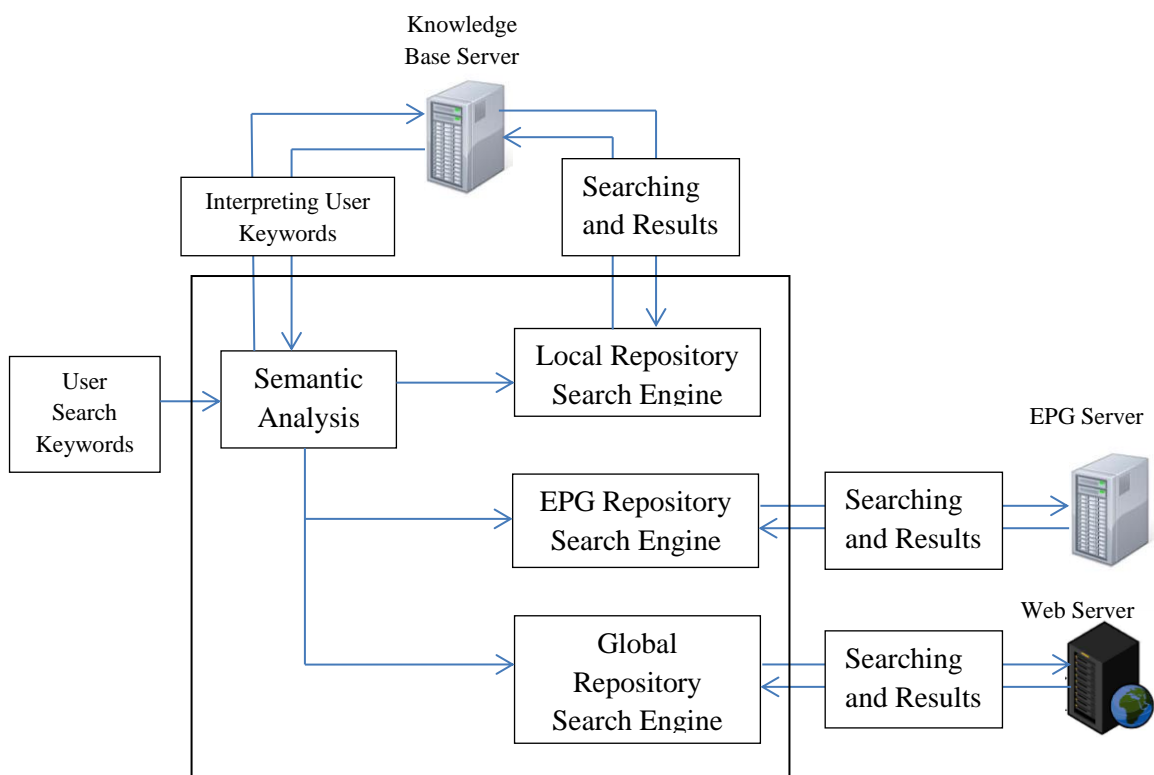


Fig. 2.10 Search Processing in the Smart TV system (M.-E. Kim et al., 2013)

2.5 Recommendation Systems

The recommendation system is one of the most important parts in personalized EPG systems. It has become a major research area due to the spread of using the recommender systems in many other applications such as Internet shopping and tourism websites. Therefore, this section discusses the relevant research that has investigated the development of recommendation operations that could be applied to online TV systems.

2.5.1 Traditional Recommendation Approaches

Basically, there are three main traditional recommendation techniques used to recommend items for users: content-based, collaborative filtering and hybrid recommendation approaches. A content-based approach depends on understanding the commonalities between the user's highly rated items or programmes. This could be done by creating a user profile that contains the user preferences, choices and their rating on each item or programme. Then, based on the commonalities of the user choices such as a specific actor, genre or subject, only those items that have a high degree of similarity with these preferences will be recommended. However, there are several limitations and drawbacks that can affect using this recommendation technique. The content-based approach is limited by the features that describe the recommended item or programme. This limits the content analysis if there are different items that are described by the same set of features. Additionally, the system can only recommend items that are similar to the items that have been rated by the user. Therefore, the user will not be able to experiment with anything different from what they have seen before – a problem known as overspecialization. Another problem is when the user only has a few rated items or new users join. The content-based recommender needs a sufficient number of items before it can understand the user preferences and present accurate recommendations for the user.

In contrast, the collaborative filtering recommendation approach tries to predict items or programmes for a particular user based on the rating of other users who have similar tastes of previously selected items. That means according to this approach, the recommender system will seek for users who have rated items similar to the particular user items and then recommend the other items of those users that have not being watched or selected to the particular user. Collaborative filtering can be classified into two methods called the memory-based and model-based method. The memory-based method makes rating

predictions based on the entire collection of items previously rated by users where all ratings, items and users should be stored in memory. On the other hand, the model-based method uses the collection of rated items to create a model from which the similar users will be grouped into sets or clusters. Then the system uses these sets to predict rated items for any new user after decide which set that this new user should belong to. More detail about the similarity calculations and clustering methods will be discussed in the next sections.

However, there are some limitations in using a collaborative filtering approach in recommending items. The cold-start problem is caused by new users who have not submitted any rating for any item. That is because in order to predict accurate recommendations, the system needs to learn the user preferences from ratings given to a sufficient number of items. Similarly, new items cause the same problem which is called the first-rater problem. That is because new items cannot be recommended to the users until some other users have assigned ratings to them. Moreover, rating sparsity could also affect the accuracy of the recommendations because the system cannot find a sufficient number of rated items or users. In this case the system has to depend on the user profile information in order to calculate a users' similarity.

In order to avoid the drawbacks of the two approaches previously described, a hybrid approach has been used which combines the content-based and collaborative filtering methods. There are different ways to combine the two approaches. The first one is by implementing each approach separately and then combining their predictions into one list or constructing a general unifying model that incorporates both approaches. Moreover, it could be implemented by incorporating some content-based characteristics into a collaborative approach or vice versa (Adomavicius & Tuzhilin, 2005; Bobadilla, Serradilla, & Bernal, 2010; Martínez, Arias, Vilas, Duque, & Nores, 2009; Wu, Liu, & Luo, 2008).

2.5.1.1 Content-based Recommendation Systems

Generally, content-based recommendation approaches depend on calculating the similarity between the previous user choices and some available items to identify and recommend items for that particular user. Many methods have been used in order to calculate the similarity between any two items such as cosine similarity and Pearson correlation.

In order to provide a list of recommended programmes, information should be collected that is relevant to the user and programme content. This information will be subjected to several analysing operations in order to extract the needed parts of this information and to use it within the recommendation processes. The recommendation needs to adapt to the differences in the formats of the collected information. As an example, the TV content metadata may be described by one provider using XML whereas others may use a different format or different tags for the same features. On the other hand, the collected information about the user may be described and presented in a totally different format. Therefore, the recommendation systems should contain different parts that can handle these formats to execute the analysing operations (Ricci, Rokach, & Shapira, 2011; Zhang & Zheng, 2005). Moreover, the content metadata should include the main features that describe the content such as the category, genre, actor, language, production year and originating channel (Scholl, Thelen, Kneissler, & Kellner, 2009).

As an example of content-based recommendation systems, a recommendation system for digital TV has been proposed by Shin which is based on extracting the content metadata from different sources. This system depends on performing similarity measurements between the collected content descriptive metadata and information from user profile. The user profile contains two types of information. Explicit information includes information entered by the user about their preferences and rated items and implicit information extracted from the user's viewing history. The system mainly consists of two subsystems, a user profiling subsystem and content rating subsystem. The user profile is a database containing keywords, genre, channels or languages preferred by the user. Additionally, the user can select a stereotypical profile which is pre-defined by the system and includes some preferred choices. On the other hand, the content rating subsystem is responsible for extracting and storing content description data which is used later for recommendation purposes. This system uses the content-based recommendation approach but with two different methods: explicitly by the user themselves and implicitly by analysing the history of viewing behaviour (H. Shin, Lee, & Kim, 2009).

2.5.1.2 Collaborative Filtering Recommendation Systems

The collaborative filtering (CF) approach is considered as one of the most popular recommendation approaches that is used commercially especially in the Internet shopping and movies web sites. Generally, it depends on clustering users into several groups based

on their previous choices. Different algorithms are then used in order to cluster the users into their appropriate groups such as k-means and nearest neighbour algorithms. In the television content field, this approach has been used to provide more un-experienced programmes based on the other users' selections in order to overcome the overspecialization problem.

An automatic recommendation system has been proposed by Kim for IPTV and traditional TV which recommends a personalized list of TV programmes based on a collaborative filtering approach. The system consists of three main parts, user profile reasoning, user clustering and recommendation part. This system doesn't rely on the explicit user feedback but instead implicitly learns the users' interests based on the TV programmes and genres. Then, the system clusters the users into groups based on the TV programmes genres that are extracted from the users' viewing history. Two methods are then used to cluster the users which are demographic clustering and k-means clustering. The programmes that belong to each group are recommended to the users after a ranking operation which is based on the preferred programmes by the user, popular programmes and TV programmes which have been watched recently. However, each user' recommended TV programmes represent the programmes that belong to the user group but only those programmes that haven't been watched by that particular user (E. Kim, Pyo, Park, & Kim, 2011).

In order to get a more accurate clustering method in collaborative filtering approach, a hybrid fuzzy-genetic approach has been used by Al-Shamri to retain the accuracy of memory-based CF and the scalability of model-based CF. This hybrid system is proposed to exploit both the user ratings for highly rated items and some content features taken from their previous choices. The system hybridization is introduced at four levels known as the features level, model level, CF algorithm and approach level. The features level collects and classifies the user demographic information and other information relevant to their preferred programmes such as genre or actor name. At the model level, the system builds the user model from the collected information. The CF algorithm level is responsible for the hybridization between model-based and memory-based approaches. Then the final level includes the proposed fuzzy-genetic approach where a fuzzy distance function is used for finding users' matches and creating recommendations for each user. The recommendation operation consists of four steps namely, data collection, user model formation, neighbourhood set selection and making recommendations (Al-Shamri &

Bharadwaj, 2008). Another system proposed by Mai used the same principles but used a different approach for clustering the users. Here a neural network based clustering CF algorithm was used to classify the user models based on a Back-Propagation (BP) neural network. This system improved the CF approach and overcame the disadvantages of this method such as the sparsity and cold-start problems (Mai, Fan, & Shen, 2009).

Chang undertook a study for recommending TV programmes to users that developed a system that can recommend programmes relevant to the user preferences based on a cloud computing architecture. The system filtered the viewed programmes according to the viewing duration and programme content. Moreover, it collected personal information about the users such as age and location. The system offered an EPG to build the user behaviour simulation system for evaluation purpose in order to generate a large amount of user behaviour data and then saved the collected data as users' profile data. When the user selects any programme from the EPG, the system saves the programme information such as channel name, programme name, location, viewing date and age in the user profile. Map-Reduce programming framework has been used which is efficient for processing the massive quantity of data. This system used a clustering algorithm based on K-means and K-nearest neighbour to cluster the TV programmes by viewer age and rank them based on their popularity (J.-H. Chang, Lai, Wang, & Wu, 2013; C.-F. Lai, Chang, Hu, Huang, & Chao, 2011). Another cloud-based recommendation system has proposed by Yijun for mobile users. This system exploits the user behaviour information in order to recommend multimedia items for users. The users are classified into several groups based on their context. Additionally, a video-sharing website has been used to collect the users' relationships and user profiles to generate the recommendation rules. The system is implemented on a Hadoop platform where the user clusters and multimedia content are collected. The system framework included four components; the user behaviour collecting, user content clustering, dynamic recommendation rule generating and optimized real-time recommendation components. This system collects the user context (such as time and location) and user interests (such as browsed contents, preferred keywords and category), and then defines the content description and content access pattern to cluster the contents into different user interest clusters. Moreover, the recommendation rules are extracted from user context clusters and user content clusters to decrease the necessary analysing processes and optimize the execution time (Mo, Chen, Xie, Luo, & Yang, 2014).

2.5.1.3 Hybrid Recommendation Approach

A hybrid approach has been used for a recommender for TV programmes proposed by Barragáns-Martínez. This recommending system provides the users, through a website, with a list of TV programmes for specified channels based on building a profile for each user through the same website. The system collects TV programme information of all available channels with any tags or keywords describing their content. Then, this information is filtered according to the user preferences which are contained within the user profile. The profile of each user consists of different types of information whether this information has been collected implicitly or explicitly. This profile may contain the types of user's interesting programmes, demographic information or some rated programmes that previously recommended for the user. Moreover, other information is extracted by observing the user's viewing behaviour to learn their preferred TV programmes and schedule. Through these rated items and preferences, the system analyses and uses the hybrid recommendation approach to recommend a list of programmes to each user through their account. Figure 2.11 shows the hybrid recommendation approach used by this system. The system implemented the operations that need a large amount of computation when the user was offline. This included downloading the list of TV programmes and storing them in a database. Providing the predicted list of programmes and ranking was then only implemented when the user was online (Barragáns-Martínez et al., 2010).

Another system proposed by Debnath hybridized the content-based and collaborative filtering recommendation approaches. Each item was represented by a feature vector where those features representing a specified aspect of item. These features have different weight values based on the users' judgement. This system has been applied on Internet Movie Database (IMDB) where the features represent the movie metadata such as genre, country, language, director and so on. The weights values estimated from social network graph where the graph represents human judgement of similarity between items aggregated over a large population of users. This system provided recommendations without using explicit user feedback on their preferences (Debnath, Ganguly, & Mitra, 2008).

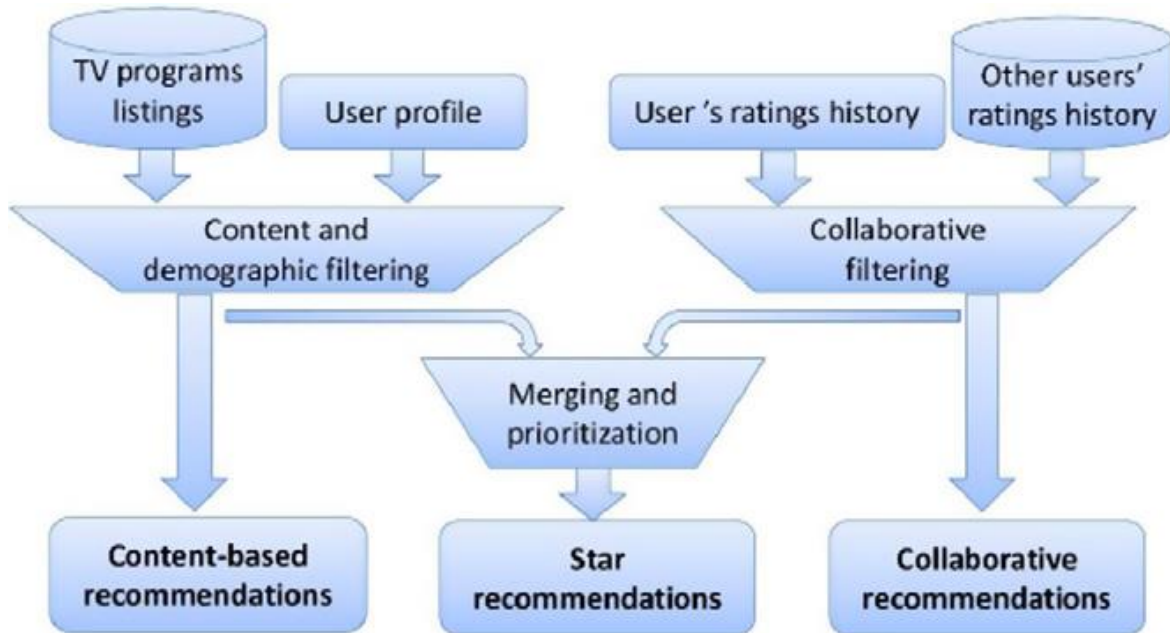


Fig. 2.11 Hybrid Recommendation Approach (Barragáns-Martínez et al., 2010)

De Campos presented a hybrid recommendation approach that used Bayesian networks to deal with the problem of the uncertainty in recommendations. It used artificial intelligence to predict how a particular user should rate a given item. Bayesian network formalism is used to represent the relationship between users, items and item features which describe the item. This recommendation system is represented by two parts. The first one represents the user profile and information about how the user rates the items. The second part represents the relationships between the users based on given items (previous selections) which represent how the content-based information will be used in improving the knowledge in the collaborative approach (De Campos, Fernández-Luna, Huete, & Rueda-Morales, 2010). The same principle has been used by Song to recommend TV programmes for smart TV users. A Bayesian probability model was used to mine the enormous set of information about programmes and extract the user feature information in order to identify those programmes that meet their demands. The system consists of four parts; data acquisition, data filtering, model building and results prediction. The first part is responsible for retrieving the TV programme information through the set-top-box and user preferences information. Then using a Bayesian data mining model, the system analyses

both programmes and user features information, builds the user model and predicts and recommends the results (Z. Song, Wei, & Jia, 2012).

Another recommendation system has been studied by Xu for web-based catch-up TV in collaboration with the Australian national TV. The system is designed to provide recommendations for the users based on their previous choices and those of other users. The system consists of two recommendation approaches. The first one is the content-based approach which recommends programmes based on what user has already watched. This approach is used to recommend things such as new episodes of a series that the user is known to be watching. The second approach is collaborative filtering which recommends programmes based on the preferences of the entire community of users. Then both recommendations that are generated by these two approaches are combined together to be displayed to the users in a single list (M. Xu et al., 2013).

Another recommendation system has been proposed that is based on content-based and collaborative filtering techniques but also uses another principle in clustering the users. This system uses partial similarity in order to cluster users into several communities where each community is represented by an attribute represent a certain aspect of user's interests (tastes). Based on the different attributes that describe each item and user, the user might belong to more than one community and so they can change the recommendations by changing the selected community. This method helped the user to experience new and unexpected programmes. The system has been evaluated through an experiment where the participants responded the questionnaires that have been prepared to asses the recommended list of programmes (Kamahara, Asakawa, Shimojo, & Miyahara, 2005).

2.5.2 Personalized Recommendation

Recently, recommendation systems have been designed to personalize the recommended list of programmes based on four strategies. These strategies are explicit profile which contains identity information related to the user and provided by the user themselves, implicit profile which describes the viewing history of the user and is used to predict the user's preferences based on a statistical analysis of the saved information, user feedback which represents user ratings for watched programmes which helps the system to recommend programmes which are similar to the user's preferred programmes and user context which is used to determine the viewing pattern and appropriate features for the

user's environment. Based on these strategies, the personalized recommendation systems generally comprise of four modules as shown in figure 2.12 which are; user profile module, programme representation module, CF module and recommendation module. The programme representation defines the content features that describe each programme such as category, genre, description and broadcast information. The user profile should consist of a fixed part and variable part. The fixed information includes gender, age and other statistical information about the user. The variable information includes the preferred programmes, keywords and their wishes and demands. Different algorithms are used to bring similar users into groups. Similarly, different mechanisms are used to find the appropriate programmes for each user based on similarity calculations (Knee, Reynolds, Ellis, & Hassell, 2009; Schaffer, Lee, & Gutta, 2011; J. Xu, Zhang, Lu, & Li, 2002).

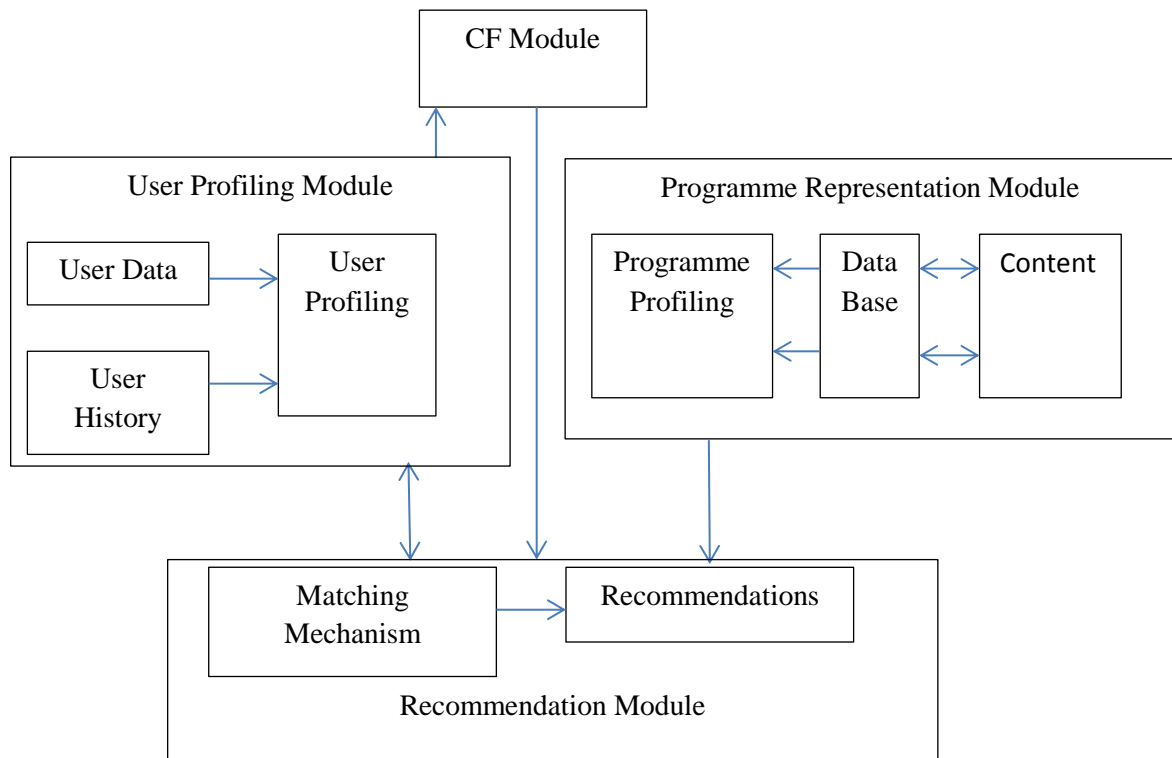


Fig. 2.12 Typical Personalized Recommendation System (J. Xu, Zhang, Lu, & Li, 2002)

In order to personalize the recommender system and provide a list of TV programmes that match the viewer preferences and address all the issues that could affect the performance of the recommendation operation, Chang has published a literature review covering TV

programme recommender systems and within it proposes a smart TV programme recommender system. The suggested system depends on implicit and explicit information collected from the user feedback and the watched programmes in addition to social recommendations. The system consists of a TV programme content analysis module, user profile analysis module and user preference learning module. The TV programme content analysis module is responsible of extracting programme information such as the title, genre and keywords. Moreover, the user can rate the programme using online social media and then the system can exploit these ratings in recommending the programme to the other users. The user profile analysis module collects the user demographic information, viewing history and preferences. Also, it provides social relationships by selecting other users as friends or family. On the other hand, it is suggested that the user preference learning module could comprise of three learning approaches; learning from the viewing history, learning from implicit network which means those users who have a common viewing pattern or taste and learning from explicit networks such as friends and family. This idea has opened the door to more sophisticated recommendation systems that can interact with the users and combine both the Internet and the TV programmes within the same system (N. Chang, Irvan, & Terano, 2013).

A dynamic recommendation system has been proposed by Dai to personalize the recommendations and adapt the changes in the user viewing pattern. The proposed system combines different recommendation schemes by using multiple recommendation agents to create a recommendation manager. Using content-based and collaborative filtering techniques, the system creates an explicit and implicit user profile based on the user feedback and viewing history. Moreover, the system provides an explicit recommendation system where the user can specify the programmes that they want to watch alongwith the time, date, channel, etc. Then, the system will recommend those programmes that match these criteria. In addition, the user needs to provide a feedback for the watched programmes and based on this feedback the system creates a list of recommended programmes. The recommendation system is located on a server which provides a centralized control for the offered services. The recommended list can then be displayed on a TV or through a web page. However, this system has not yet been implemented and so is currently only a proposed (Dai & Cohen, 2003).

In addition, other analytical methods have been proposed for the recommendation process within digital TV such as taking into consideration the current time, the available programmes at that time and deciding on the recommendation time. Oh proposed a TV recommender system called Show Time that determines the timing as well as the items for recommendation. This recommender system is continuously analysing the TV programme that the user is watching at that moment and the other programmes on the other channels by analysing the EPG information to create a user profile. Then, the system decides the time to recommend the selected programme or programmes based on the show time or the time remaining. The recommendation results are displayed on the TV screen and the user can then select or ignore the recommended programme or programmes. Thereafter, the system will analyse the user's decision and feedback and update the user profile to recommend a new programme or programmes. Figure 2.13 shows an overview of this system (J. Oh, Kim, Kim, & Yu, 2014).

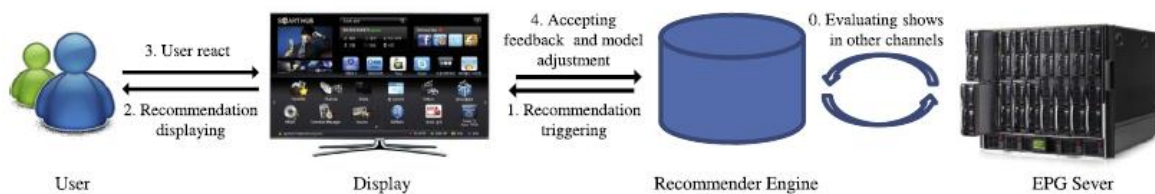


Fig. 2.13 Overview of Show Time System (J. Oh et al., 2014)

Another recommendation system has been proposed by Shang which recommends a personalized list of programmes based on activities, interests, moods, experience and demographic information. The system is called the AIMED model. It consists of four modules; user profile and user stereotype, viewing communities, programme metadata and viewing context module. This system relies on collecting different information relevant to the user through a questionnaire in the registration stage which includes their preferred activities, interests and demographic information. Based on this information the system clusters users into groups using a K-means algorithm to develop viewer group models. Moreover, the system used the content-based and collaborative filtering techniques as additional recommendation methods. The system generates the viewing behaviour by

analysing the viewed programmes where the programme attributes such as channel number, programme name, category, language of programme and cast are stored in the user profile. On the other hand, the system determines some key characteristics to find a match between each user and the viewing communities according to the similarities in those characteristics. Additionally, when the user is viewing any programme, they can provide rate and mood information to the system to be saved in their profile where it will be used later in the other operations. Although this system provided a personalized list of programmes based on different features and characteristics, it is still needed the user to enter different information and ratings for the viewed programmes in order to build the user model and provide an accurate recommendations to each user (Hsu, Wen, Lin, Lee, & Lee, 2007).

2.5.3 Group Recommendation

Because the TV is often viewed by a group of people such as family in a living room or a group of students in lounge, a new recommendation system has been proposed by Yu to recommend TV programmes for a group of viewers. In order to achieve a programme recommender for multiple television viewers, there are three strategies that could be used. The first strategy is a group agent where the users register a common account for themselves. Thereafter, they select their preferences to create a common profile. When the users want to watch TV together, they need to log in using their common account and the group agent will then learn the group's preferences and recommend programmes for them. The second strategy is merging recommendations. The system can use any recommendation approach to recommend programmes to the users based on their profiles. Then, it merges all the recommendation lists together to find the common programmes and generate the recommendation for the group. The third strategy is merging user profiles to generate a common profile. The final list of recommended programmes is then based on the features of the common profile. Figure 2.14 shows the schematic of this recommendation strategy.

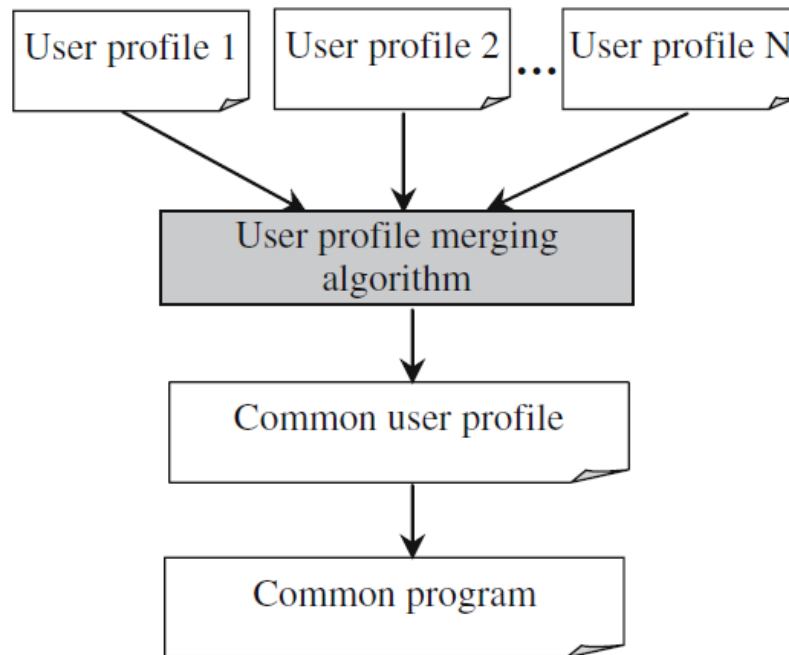


Fig. 2.14 Merging User Profiles Strategy

Several criteria should be determined by the system in order to merge the users' profiles such as to select the part of the user profile that should be merged or weight the features that will be used to recommend the programmes. Additionally, the system offers a user interface for multiple viewers to specify the users who log into the system as shown in figure 2.15 (Yu, Zhou, Hao, & Gu, 2006).

Another method has been presented by Kim for group recommendations. This is based on generating a profile for each group of people who share the same activity or same interests. The group profile represents the aggregated profiles of the members of that group. Then the traditional collaborative filtering technique is implemented using the nearest neighbour algorithm to find the neighbour group and generate a set of recommendations for each group. The second step is to remove the irrelevant items by calculating the similarity between the set of recommendations and the user's interests which are contained in the user profile. However, this method has been developed for recommending online books in reading clubs and not for TV programmes (J. K. Kim, Kim, Oh, & Ryu, 2010). The same idea has been presented by Chen for group recommendations where the difference between individual recommendation and group recommendation has been considered. Group decisions may be affected through social influence or opinion leadership influences that

can directly or indirectly influence the thoughts, feelings and actions of others. Additionally, a genetic algorithm was used to learn the group rating for items where the interactions between the users were learned from the group ratings (Y.-L. Chen, Cheng, & Chuang, 2008).



Fig. 2.15 Multiple Viewers Graphical User Interface (Yu et al., 2006)

2.5.4 Context-awareness Recommendation

Another technique that has been used to enhance the recommendation process is based on context-aware recommendations. Context definition could be different based on the type of information that is needed. There are many types of context that should be gathered such as user context (identity, age, and gender), device context (screen size or resolution), network context (available bandwidth), content context (content description) and dynamic

context (time and location). By including this information into the recommendation systems, the recommendation operation can provide more personalization in the selected list of programmes through adapting to the user's environment. Shin proposed a context-aware recommendation system which depends on the time and date as dynamic user context. They used the user context as a recommendation criterion in the similarity calculations in addition to the programme metadata. This recommendation operation is based on three layers; the user, item and user context. As a result, it was verified that the context-aware recommendation achieved an improvement in the recommendation performance compared with a traditional hybrid recommendation approach (D. Shin, Lee, Yeon, & Lee, 2009).

Context-aware recommendation systems model and predict user behaviour and preferences by incorporating the available contextual information into the recommendation process. Context-aware recommendation systems are described as three-dimensional recommendation systems because the rating function is based on a user-item-context relationship. In contrast a traditional recommendation system is described as a two-dimensional system because it only depends on the user-item relationship.

The contextual information could be obtained in different ways; explicitly by asking the people directly and gathering the information from them or implicitly by gathering the information from a database or environment. Implicit gathering is often implemented using an application or device that interacts with the user environment. Additionally, the contextual information could be gathered by inferring the context using statistical or data mining methods (Adomavicius & Tuzhilin, 2011).

Context-awareness recommendation can be classified into three approaches; contextual pre-filtering, contextual post-filtering and contextual modeling. Contextual pre-filtering filters the rating data for a particular context before calculating the predictions by the recommending system. That means, the items that will be recommended are predicted from a set of items that are relevant with only the particular context. The contextual post-filtering approach means generating the recommendations based on any recommendation algorithm and thereafter filtering these recommendations according to the user context. The contextual modelling incorporates the context information into the model that is used for generating recommendations. Figure 2.16 shows the different ways for incorporating the contextual information in the recommendation systems.

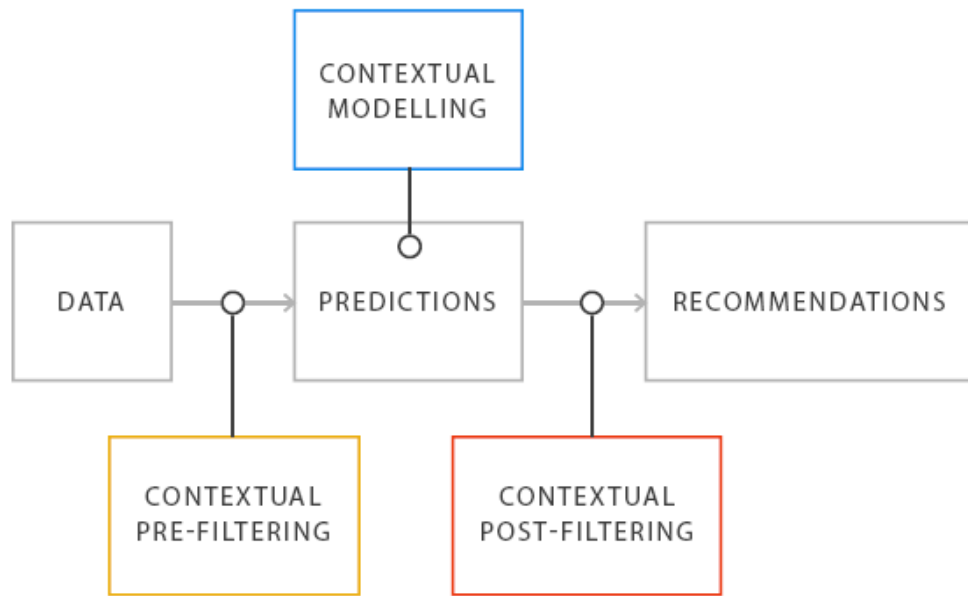


Fig. 2.16 The Possible Ways for Incorporating the Contextual Information into the Recommendation Systems (Steeg, 2015)

Campos presented a study that drew a comparison between these three approaches to find the best approach that provides the best recommendation performance. In this study, the context that has been chosen is time and social companion. The users were asked to rate a group of movies and state the context which includes the time of day such as morning, afternoon or night, period of the week such as working day or weekend and with whom they preferred to watch the rated movie such as alone, as a couple, with my family or with my friends. However, by analysing the results that were obtained from the experiment, it showed that no approach is superior to the others but that the combination of contextual recommendation with another recommendation algorithm can improve the recommendation performance and obtain more personalized list of programmes (Campos, Fernández-Tobías, Cantador, & Díez, 2013; Steeg, 2015).

A recommendation system based on combining the user profile and context has been proposed by Abbar. The user profile represents information that characterizes the user such as demographic data, domain of interests and quality and delivery requirements. The context is defined as a set of features which describe the environment that the user interacts with such as time, location and cognitive status. This system consists of two main

functions; the knowledge acquisition process and personalized recommendation. The knowledge acquisition collects the information from three sources which are the user profile, content description and context. All the selected programmes are then stored in a user log file which represents the user profile besides the user context and other information. The personalized recommendation contains the recommendation engine which is responsible for learning the user's viewing behaviour and finding the appropriate programmes. However, the users need to provide the system explicitly with information about users' contexts and ratings for the previous watched programmes in order to create the personalized recommended list (Abbar, Bouzeghoub, & Lopez, 2009).

Another context-aware recommendation system has been proposed by Song which takes into consideration the different types of context. The system gathered four types of context in order to personalize the recommended list of programmes. The first type is the user context information which includes static information such as name, age, gender, language and preferred actors and dynamic information that frequently changes such as user location and emotions. The second type is device context information which includes information about the device such as screen size, capacity and connectivity. The system also gathered network context information which included network access type, bandwidth information and quality of service information. The last type is the service context information that includes information about the services such as content description and requirements for devices and networks. The main functional blocks of the context-aware system are shown in figure 2.17.

However, this system depended on real-time context collecting and personalizing, therefore, there is a time delay in service initiation due to the time taken to collect context information and perform the personalized content selection leading to the generation of a list of recommendations (S. Song et al., 2012).

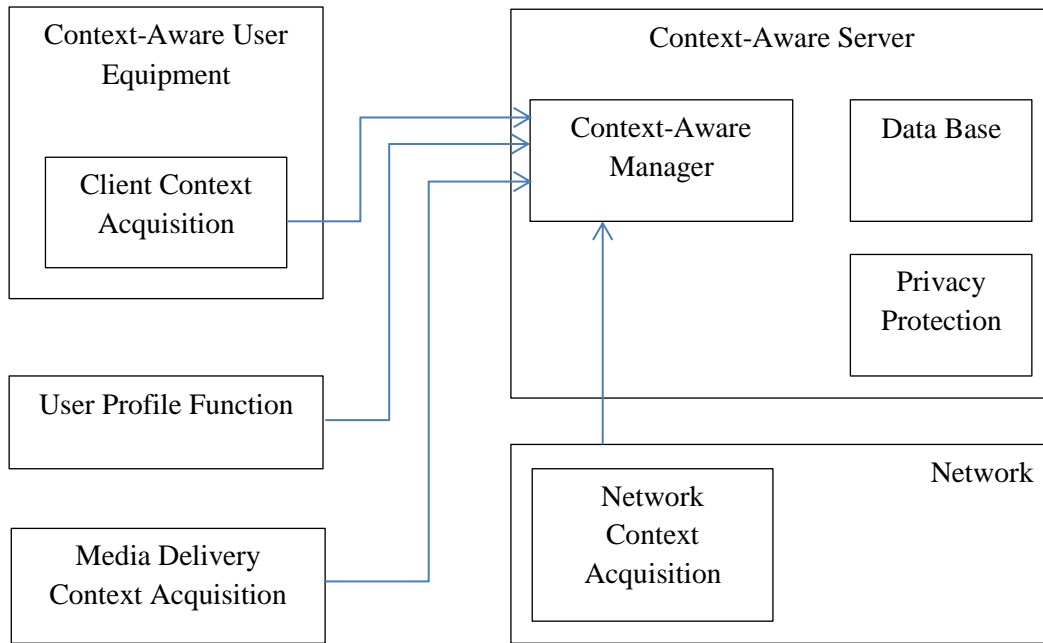


Fig. 2.17 Context-aware Recommendation System (S. Song, Moustafa, & Afifi, 2012)

2.6 Social Networks

Nowadays, using social networks (SN) has become very popular and one of the most important ways of keeping in contact with friends and family members. Online social networks could be defined as web-based applications that allow their users to create a profile that other users can see and use to create a friend list of other users to make contact. Besides connecting people with each other, social networks are used to exchange their ideas, interests and important moments. Consequently, many researchers have investigated the effect of these relationships on users' choices and interests. Several aspects have been studied such as the psychological outcomes from engaging in these social networks, the effects of the number of friends and the amount of SN use on the user behaviour. Additionally, social support is considered as one of the most widely used concepts across various domains in social networks and has the attention as a context that need further studies to explain the positive outcomes of social networks in recommendation operations. Social support could be defined as exchanging the resources or aids between the individuals through the social networks and is considered as a key benefit that users can obtain from these networks. Another psychological outcome from using social networks is

the sense of community. It is defined as a feeling of belonging to a group or community within which the members can perceive themselves as similar in characteristics and dependent on each other. Many studies have confirmed that social support and sense of community can be exploited as a good predictor for sharing items between users who belong to the same community or group.

Other studies have investigated the relationship between various social features of users (such as geographic distance and friend similarity) and the similarity in their interests. Moreover, what are the aspects of users, who have relationship, that should be similar to estimate the user interests when the another user's interests are known. A study has been implemented on three interest domain; music, movies and TV shows to reveal the effect of the social features on users interests. The study showed that the people tend to be interested in the same movies, TV shows and music if they share similarities in their demographic information such as age, gender and location. Moreover, friends have a higher similarity than strangers in their choices and interests and this similarity increases when those friends share more than one common friend. Furthermore, people with a high diversity of interests are more likely to share their interests with others. This study discussed how to infer the interests' similarity of two users when one of the users' interests is unknown. The psychological effects of social network relationships illustrate that a high proportion of social network friends who share the same interests can interact with each other more than strangers. This can influence their selections and can be exploited in recommendation systems especially using collaborative techniques (Ellison, Steinfield, & Lampe, 2007, 2011; Han, Wang, Crespi, Park, & Cuevas, 2015; H. J. Oh, Ozkaya, & LaRose, 2014).

Due to the growing importance of social networks, many studies have been focused on the nature of relationships between users and how these relationships can affect the selections and help users make decisions. For this reason, social recommendation has been investigated on the basis of psychology and sociology studies. The study discussed the effects of the social relationships on users' selections and their interaction with the recommendation systems. Two factors have been specified as social context which are individual preferences and interpersonal influences. The individual preferences describe the user behaviour toward the selected items which depends on whether the user previously liked or not similar items. The interpersonal influences describe whether the

user has a relationship with the person who recommends the item or not. This information has been extracted from historical data relevant to the target user. The social recommender system has been built to incorporate social contexts and has been evaluated through collecting real data from websites similar to Facebook and Twitter. The results showed that using the social recommender system improved the performance of the recommendation operation compared to the traditional recommendation techniques (Jiang et al., 2012).

2.6.1 Social Networks-based Recommendation Systems

Many researchers have discussed the effect of social networks as recommendation approach and compared the recommender performance with other types of recommendation systems. A study presented by Bonhard discussed the effect of relationships between advice-seeker and recommender and incorporated social networking features within the recommendation process. The study examined the effect of knowing the recommender on an advice-seeker selections and the effect of knowing the advice-seeker on the recommendations of the recommender and how this in turn helped them to make decisions, and hence, improve the recommendation process. The study applied considered group of participants that knew each other and were asked to provide basic profile information and rate 20-30 films and then to rate some profiles in terms of familiarity, similarity and trust. The results showed that an advice-seeker selected programmes that were recommended (rated) by known persons faster than without knowing them. In addition to knowing the recommender, two conditions have been recognized that affected decision making. The first one was when the advice-seeker knows that the recommender has similar tastes and secondly when both have mutual knowledge about each other's tastes. Even when they are different in tastes, the recommender will know what the advice-seeker will like (Bonhard & Sasse, 2006). In the same field, Mao discussed how the social influence between friends could affect the decision-making of item selection. The author developed a new parameter learning algorithm based on expectation maximization. This is based on integrating user behaviour, social influence and item content in order to enhance the recommendation performance. The system has been evaluated through a set of experiments using datasets collected from different web sites. The results showed that social influence has a valuable enhancement on the recommendation process and users are more likely to prefer friends' recommendations

than those from other recommendation schemes but combining both user preferences and social influence is more effective than depending on social influence only (Ye, Liu, & Lee, 2012).

Another recommendation system has been proposed based on social networks related to items used in electronic commerce. This system relied on making recommendations for items which have been purchased or rated by members of different social networks. As an example, when someone purchases any item online, the system provides recommendations from trusted friends and members in the same social network and members in other networks. The ranking of those recommendations is based on the relationship proximity of the members with the target user. That means the system relies on friends followed by other members who have mutual friends with the target user within the same social network then the strangers (members within other social networks) will be at the end. This system consists of a social network module, recommendation engine and data repository as shown in figure 2.18.

The social network module is configured to connect the recommendation application to the social network through an API provided by that social network. The data repository contains information about items that should be recommended within the social network. The recommendation engine provides the recommendations based on the social networks relationship proximity that is extracted from the social network (Berman, Iyer, Richardson, Rahrkar, & Seetharaman, 2012). However, this system was applied to purchasing items and exploited only one aspect of social network information (relationship proximity) in ranking the recommended items.

A recommendation system based on social networks only has been proposed by Fischer. This system relied on sharing the users' activities with friends who belong to one or more social network. The recommended items are weighted based on the closeness of the relationship with the target user which is determined based on the history of content sharing activity between users. This system comprises a group of services that are managed by a main application which communicates with the different web servers. Through these services, the application retrieves the needed data that is relevant to each user and their relationships with other users in each network (Fischer, 2012).

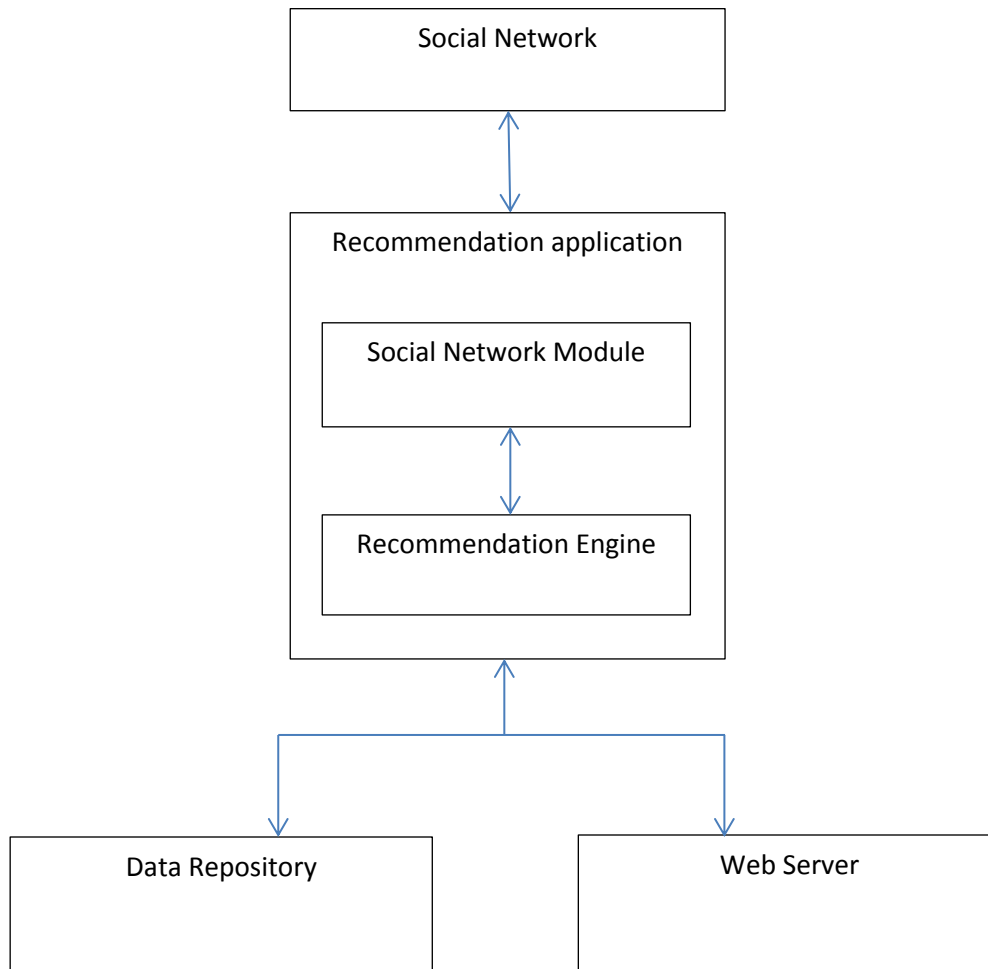


Fig. 2.18 Block Diagram Describes the Idea of Social Network Recommendation System
(Berman et al., 2012)

Another study proposed a social network recommender for an IPTV system where the recommendations for content had been determined based on watched items, likes or those that have been tagged by other social network members. This system could receive data about content items from other sources or rating engines such as IMDB or social networks such as Facebook. Figure 2.19 shows a block diagram for the designed social network recommender. The recommendation system includes a tracker module to track user interactions with content and to collect further information that may be relevant to provide recommendations to other users who are members of the same social network. This system can perform a keyword search to find content matching that keyword in the content database or user's social network. This keyword could be a genre, actor name or words in

the title or description belonging to the viewed programmes by the user themselves or members of their social network (Mathur, 2013). However, this system was applied to a limited content source which was available in the IPTV system and relied on the users' activities on the social networks such as comments or tags.

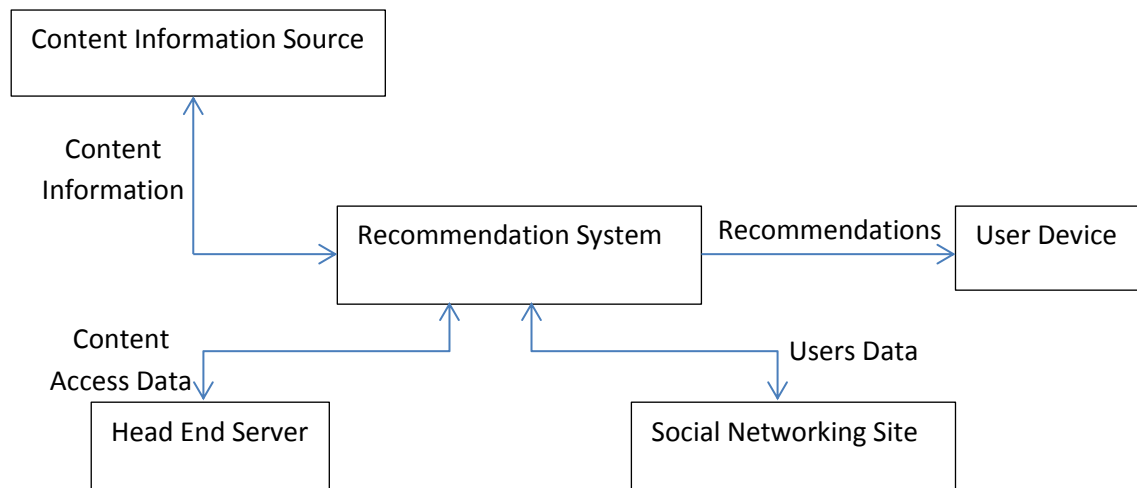


Fig. 2.19 IPTV Social Network Recommendation System (Mathur, 2013)

Another proposal including created a web site for recommending films to users based on a trust relationships. The web site has its own social network which offers the users a list of friends who are in the network and to be able to rate those friends, based on the strength of the relationship that exists such as acquaintance, good friend or best friend. Moreover, the web site asked the users to rate as much as possible different films and allowed them to write reviews (opinion) about each film. Based on these explicit ratings for both movies and friends, the system recommends and ranks a list of movies to each user within their own account. Additionally, the system presents the film reviews of the most trustworthy people where the trust ratings of those reviewers are displayed to the user. The results obtained from user trials showed that the accuracy of predicting films based on the trust associated with social networks were better than those based on the simple traditional recommendation techniques (Golbeck, 2006).

2.6.2 Hybridization the Recommendation Operation with Social Network

Research has also considered how social influence, as an additional source of information about the user and their friends, can help in understanding the user behaviour through observing the user preferences and interests. The social networks offer information about friends, therefore, it is not necessary to find similar users to the target user through analysing their rating and use the similarity algorithms because the fact that any two persons are friends indicates that they have things common between them. This idea could offer a solution to the data sparsity problem. Moreover, social networks can offer a solution to the cold-start problem where even though there is no history about the user's preferences, the recommending system can make recommendations based on the preferences of their friends. Social networks can provide not only the user's preferences and interests, but also list their friends. This potential has been investigated by He through a recommendation system that takes into consideration these social relationships by combining a user's preferences with general information about the recommended item and the opinions of friends within the social network. It has been applied to a real data set from the Yelp.com website to help users find restaurants, bars, shops and it also provides its own social network to invite friends or to find friends on the web site. The study showed that the recommendation performance was improved by using the social influence within the recommendation process (He & Chu, 2010).

Another approach has been presented by Liu to incorporate the social networks within the collaborative filtering method. This approach has been developed to increase the effectiveness of the CF method by solving one of its drawbacks in that it is unable to distinguish between friends and strangers when the system identifies two or more people share similar tastes. The approach is based on collecting data about the users' preferences, ratings and their social network relationships from the Cyworld website. Cyworld is a social network web site similar to Facebook and popular in South Korea and China. User's data was collected by distributing a survey to specific users of Cyworld. Thereafter, through using an algorithm to handle the collected information, the system found the nearest neighbours and associated social network members. If the social network members were also in the nearest neighbour group, then the preferences of those members were highly ranked. Otherwise, the other members' preferences are ranked based on number of viewers. Figure 2.20 shows a clarification of the presented approach. The results indicated

that combining the collaborative filtering technique and social network improved the recommendation accuracy and decreased the number of times needed to repeat the computation to get the optimum nearest neighbours (F. Liu & Lee, 2010).

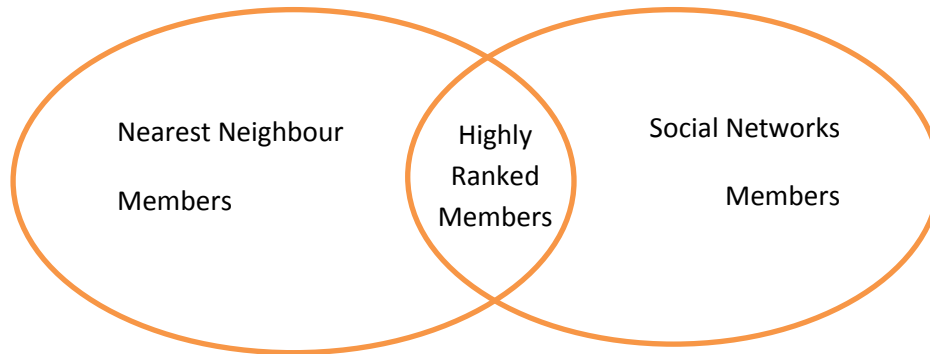


Fig. 2.20 Incorporating Social contribution and collaborative filtering (F. Liu & Lee, 2010)

A movie recommender system presented by Carrer-Neto offered a hybrid recommendation approach by combining the knowledge-based recommending technique with social networks instead of the collaborative filtering technique in order to improve the recommendation process. In the knowledge-based recommending technique, the user profile is used to identify the correlation between their preferences and existing content using inference algorithms. The system then allows the users to create friendship links with other users and uses information collected from the friends' profiles, which is relevant to their preferences, to recommend movies to the target user. The proposed mechanism identified the dominant content that has received the highest ranking from most of their friends. This system relied on IMDB (Internet Movies Database) as a domain ontology to provide the system with the necessary information about movies. The recommender system presented promising results in enhancing the recommendation process by taking into account the social influence on a user of selections made by friends (Carrer-Neto, Hernández-Alcaraz, Valencia-García, & García-Sánchez, 2012).

Another recommendation approach was proposed by Liu that presented a novel context-aware recommendation system incorporating processed social network information. This

system handled the user's context information which was collected either explicitly or implicitly to predict the user's preferences. This system used two categories of context; static context (such as age, gender and membership) and dynamic context such as mood and location. Additionally, the system selected the user's friends who share a similar taste with the target user through analysing their historical ratings. This helped the system to predict the missing preferences based on the social information. This system was applied to one of the largest Chinese social platforms for sharing reviews and recommendations for books, movies and music called Douban. The users could then follow other users and provide ratings to books, movies and music to determine their preferred items. The results showed that combining contextual information and social network information improved the quality of recommendations (X. Liu & Aberer, 2013).

A trust-based recommendation model has also been investigated by Walter to identify the impact of social network relationships on the recommendation performance. This model has been suggested to solve some problems in the traditional recommendation schemes such as sparsity. The recommender system uses the collaborative filtering technique to recommend items for each user and then it exploits the social relationships in filtering the final items (Walter, Battiston, & Schweitzer, 2008). Another study used the social networking and social tagging in collaborative recommendation. This study was implemented based on analysing the dataset that was collected from the Last.fm social network which includes users' preferences, tracks, tags and bonds of friendship. This site allows the users to create a profile and add their preferred music tracks either from the web site or from their private music collections. A Random Walk with Restarts (RWR) model was applied to the collected data to evaluate the incorporation of friendship and social tagging in terms of enhancing the recommendation performance. According to the system evaluation, it was shown that the extra knowledge provided by the users' social activity could improve the recommendation performance (Konstas, Stathopoulos, & Jose, 2009).

2.6.3 Integrating Facebook into Recommendation Operation

One of the most popular online social networks is Facebook where the number of participants in this web site reaches millions of people. It contains a huge amount of personal information that can be analysed in order to exploit it in the recommendation process. However, it is not easy to obtain all the needed information from the Facebook website because of the limitations of Facebook API. One of the methods that has been

used is to create an application in the online social network and ask users to join it if they want to participate. One of the studies that exploited Facebook in the recommendation process has been implemented by Quijano-Sanchez. A group recommender was generated by creating an application in Facebook and asking the users to join it in order to have the benefits of its services. The application is dependent on questionnaires that the user firstly should complete in order to create a user profile. The application employs three important features which are personality, social trust and historical recommendations in order to recommend new events or movies to the users' group. The application includes a create an event feature (such as going to the cinema), an option to invite friends, facilities to accept an invitation and to create a preferences profile of selected movies as shown in figure 2.21. The application gathers information from all group members and analyses it to create the group recommendations based on the selected items and events. All the recommendations are stored in the user profile in order to avoid repeating the same recommendations for the same user. Moreover, the application updates the gathered information in order to erase events or movies that have passed their deadline. This method is based on the collected information from multiple tests and questionnaires and can provide the recommendation within the Facebook website (Quijano-Sanchez, Recio-Garcia, & Diaz-Agudo, 2011).

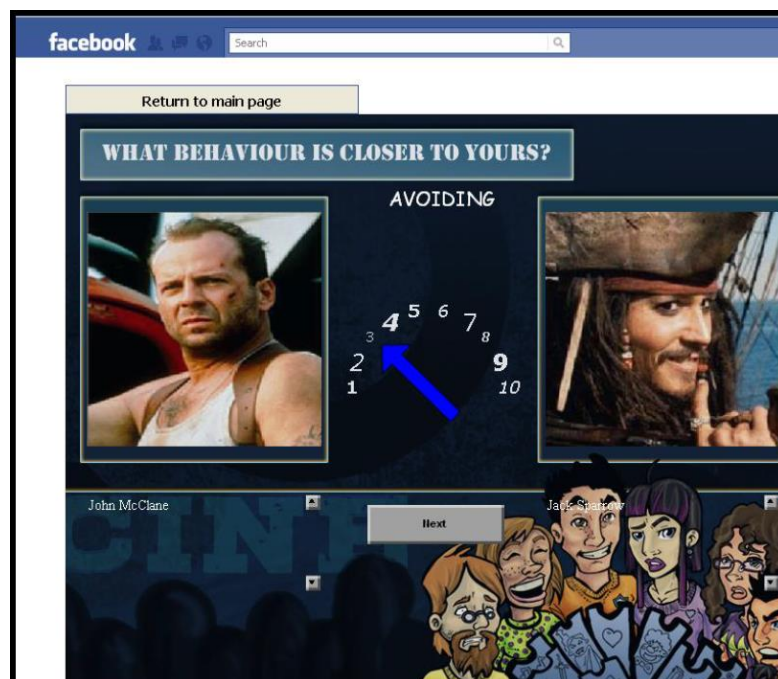


Fig. 2.21 Selecting and Rating Preferred Movies within the Facebook Application
(Quijano-Sanchez et al., 2011)

User context has also been used to improve the recommendation process where the results show that location, time and other contextual information has a big influence on user selections. The user needs to specify their status through some questions about their favourite categories, their mood, their location and so on to recommend the appropriate programmes.

A study has been conducted using the IPTV system by Mitchell where social networks have been employed to provide social awareness to the system users. Firstly, this study was based on a survey (questionnaire) organised to obtain some initial information from each user relevant to TV viewing habits, equipment and online services consumption. This information was used to categorize the user contexts such as user identity, time of watching, historical activities and access device. The system then provided an EPG to display the available channels with recommendation based on viewing history. Moreover, this system integrated Facebook into the EPG. The system provided an EPG that recommends channels for the IPTV system users based on channel popularity which is calculated by the number of clicks on each channel in the last 24 hours or by a number of simultaneous viewers at that time. In order to employ Facebook as social awareness and take the benefit of the information included on that website, the system asked the users to log into their Facebook accounts through the user interface of the system. Then the system server found the user's Facebook ID and list of their friends' IDs. Based on the information that was retrieved from each user, the system calculated statistical data about the number of viewers for each channel and identified where those view intersected with the friends' list of each user. The system ranked the displayed channels in the EPG based on the statistical data such that the channel with higher viewers appeared in the top of the list. Viewers were shown how many people are watching each channel and how many friends are watching each channel together with a listing of their names as shown in figure 2.22. Results show that the social effects play an important role in selecting the next programme where friends always prefer to watch the same programmes. This system supports users with statistical data for each channel where the recommended list is based purely on the channel popularity (Mitchell, Jones, Ishmael, & Race, 2011).

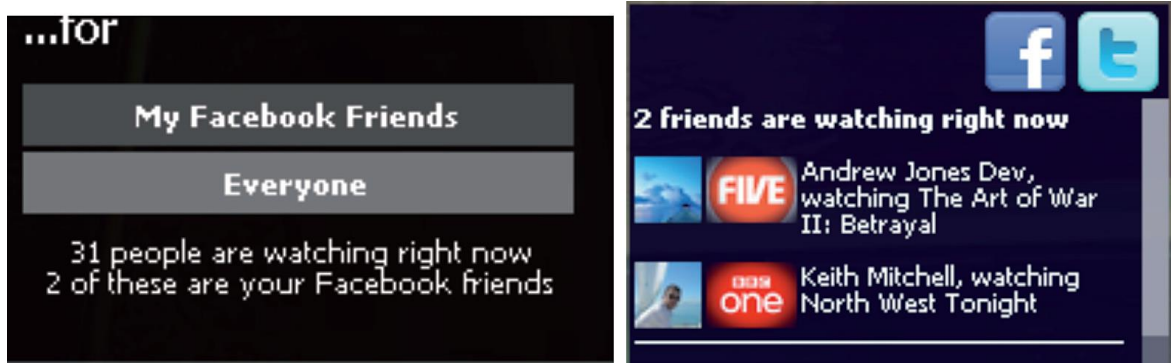


Fig. 2.22 Interface Showing Viewing Statistics (Mitchell et al., 2011)

Moreover, Facebook contains different information that could be obtained which is relevant to the operation of an EPG which can be very beneficial for recommending systems and substitute the explicit rating to achieve not only competitive but superior results. A study presented by Shapira compared the quality of recommendations between using traditional rating information and extracted data from Facebook. Facebook contains information published by the user which relevant to favourite items and preferences. Additionally, participants were asked to rate a list of movies via a specially developed Web form. Their ratings have utilized instead of the users' explicit feedback after watching any video contents. The extracted information was integrated into the traditional CF recommendation technique. Thereafter, a comparison has been implemented between the obtained results with those obtained from the traditional CF. The results revealed that the recommendations based on Facebook data were at least as good as those based on explicit ratings and even more accurate in some situations. Moreover, the integration of Facebook data into the recommendation process for users solves some of the known problems found in traditional CF especially those associated with cold start and high sparsity (Shapira, Rokach, & Freilikhman, 2013).

2.7 Summary

This section summarizes the contributions of the relevant research that has been reviewed in this chapter and discusses the new areas that should be developed in order to evolve the EPG to serve online viewing. The relevant research in this context is that which discussed improvements to the entire EPG systems rather than that which simply focused on improving one aspect of the EPG system. This work is summarized in Table 2.2 which lists the main contribution made alongside a summary of its limitations.

| Authors | Contributions | Limitation for Online Viewing |
|--|---|--|
| Stettner (2012), Ellis et al. (2014), Yuen et al. (2000), Cui et al. (1996), Drazin & Kram (2014). | Adding a new service to the existing EPG such as display part of the EPG, scan feature, Picture in Picture (PIP) technology, new GUI using new operating system, display the past programmes' list. | <ul style="list-style-type: none"> • Doesn't have the ability to communicate with all available content sources. • Doesn't have the ability to recommend personalized list of programmes. |
| Macedo et al. (2014). | Retrieving textual information from the web that is relevant to the programmes in the EPG list. | <ul style="list-style-type: none"> • Doesn't have the ability to retrieve programmes relevant to the selected programmes. • Doesn't have the ability to recommend a personalized list of programmes. |
| Tan et al. (2009). | Recommending programmes for IPTV system's users through the EPG. | <ul style="list-style-type: none"> • Doesn't have the ability to communicate with all available content sources. • Depends on one recommendation approach only (CF). |
| Boyer et al. (2014), McKissick & Forrer (2013). | Adding services to the existing EPG such as issuing reminder messages for selected or preferred programmes. | <ul style="list-style-type: none"> • Limited to a set of channels that are included in the IPTV system service. • Needs an explicit request to |

| | | |
|---|---|--|
| | | <p>specify the preferred programmes.</p> <ul style="list-style-type: none"> • No real recommendation process included in the designed system. |
| Ehrmantraut et al. (1996), Ardissono et al. (2004), O'Sullivan et al. (2004). | Personalizing the recommended list which is created based on one of the recommendation approaches by using analysing methods such as fuzzy logic and data mining. | <ul style="list-style-type: none"> • Limited for a set of channels that are included in the TV EPG. • Using only one recommendation approach (CF). |
| Smyth & Cotter (2001), Uchyigit & Clark (2003). | Provide a personalized EPG for TV viewer as a web page. | <ul style="list-style-type: none"> • Limited to a set of channels that are included in the TV EPG. • EPG is separated from the viewing device. • Depending on user's explicit feedback. |
| Chen et al. (2009). | Personalizing the EPG for IPTV system users. | <ul style="list-style-type: none"> • Limited to a set of channels that are included in the IPTV system service. • Depending on user's explicit feedback. |
| Lai et al. (2011). | Personalizing a recommended list of programmes for mobile IPTV system based on cloud network. | <ul style="list-style-type: none"> • Recommending VoD content (stored in a specified server on the cloud) relevant with user's favourite programmes. • Depending on user's explicit feedback. |
| Funk & Noffsinger (2014). | Developed a searching system for multimedia programmes in remote Internet based server that | <ul style="list-style-type: none"> • Limited for channels connected with the specified server. |

| | | |
|-------------------------|--|---|
| | is based on saving previous keywords. | <ul style="list-style-type: none"> • Dependent on the last search term without taking into consideration the programme metadata and personalization. |
| Jerding et al. (2013). | Designing a search system for multimedia content to search the digital broadcasting channels and VoD for specified input. | <ul style="list-style-type: none"> • Limited to TV channels that provide information about their content. • User search term is required. |
| Franklin et al. (2014). | Designing a search system to search the metadata of broadcast programmes for relevant programmes to the viewed and last one. | <ul style="list-style-type: none"> • Limited to TV channels that provide information about their programmes. |
| Rothschild (2011). | Proposed a searching system for Smart TVs that can provide the viewer with multimedia content relevant to the viewed programme's features. | <ul style="list-style-type: none"> • The idea is based on keywords extracted from the viewed programme to use as search keywords on the Internet to find movie, web content or e-commercial web site relevant to those keywords. • The system worked as a separate device connected to the Smart TV. • Didn't connect to video content sources. In contrast, it searches the Internet directly using a search engine for any relevant information. |
| Kim et al. (2013). | Proposing a multi-level searching process for Smart TVs to search the local repository and Internet and combine the results | <ul style="list-style-type: none"> • System is limited to three permanent sources and can not include any new source. • Depends on an explicit user |

| | | |
|--|---|---|
| | into different windows. | <p>search request.</p> <ul style="list-style-type: none"> • Doesn't have a recommendation system. |
| Kim et al. (2011), Chang et al. (2013), Barragans et al. (2010), Song et al. (2012), Xu et al. (2013), Kamahara et al. (2005). | Improving the recommendation system through using hybrid methods in recommending programmes for users or using data analysing methods to obtain more precise recommendations. | <ul style="list-style-type: none"> • It is still limited to using the traditional recommendation scheme. • Limited to a set of TV channels. |
| Hsu et al. (2007). | Personalizing the recommended programmes based on user context such as interests, activities, moods and demographic information. | <ul style="list-style-type: none"> • Depends on explicit user inputs. • Limited to a set of TV channels. |
| Song et al. (2012). | Designing a recommendation system based on different types of context which includes user, device, network and service context. | <ul style="list-style-type: none"> • Limited to a set of TV channels. • Doesn't take into consideration the other recommendation approaches. • Depends on real-time context collecting and personalizing which causes a time delay. |
| Mitchell et al. (2011). | Employing Facebook in recommending programmes for IPTV system. | <ul style="list-style-type: none"> • Facebook has been used to recognize the users who are currently watching TV programmes in order to use this information in ranking the recommended programmes and has not exploited the information included in Facebook in |

| | | |
|----------------|---|--|
| | | <p>recommendation operation.</p> <ul style="list-style-type: none"> • Depends on recommending the programmes based on real-time viewing (recently viewing). • Limited to a set of IPTV channels. |
| Mathur (2013). | Proposing a Social network recommender for IPTV system. | <ul style="list-style-type: none"> • Depends on the activities on the social networks such as comments and tags. • Limited to a set of IPTV channels. |

Table 2.2 Summary of Literature Review

As a summary, the literature review has revealed a set of common limitations that remain to be resolved in order to create a generic EPG that can serve online viewing. Specifically these are:

1. All previous EPG systems are limited to a set of channels or specified content sources. A generic EPG system should not be limited in this way.
2. The EPG systems recommend available programmes in real time within the broadcast channels and that is not suitable for online viewing where users may access different types of content sources including channels, repository and user created content (UCC). Therefore, the online EPG should be programme not channel oriented regardless of the programme source.
3. Limitation in recommendation methods due to the difference in formats used to describe a programmes' metadata and collected data which is relevant to a users' context. An generic EPG should seek to adopt a common standard for all information used as part of its recommendation process, irrespective of its origins.
4. Hybridizing the recommendation methods, especially that which depends on collecting data relevant to the user's contextual and environmental features, causes a time delay which increases when more hybrid recommendation methods are used

simultaneously. A generic EPG should seek to minimise the delay which the user perceives.

5. The difference in the displaying device should be taken into consideration. It is important that a generic EPG is able to deliver the same services irrespective of the viewing device which they are using.

The next chapter will discuss the above requirements in more detail and consider how a generic EPG system can be designed for online viewing.

CHAPTER THREE

REQUIREMENTS OF EPG SYSTEM DESIGN

Chapter Three

3.1 Introduction

The literature review that was presented in the previous chapter reviewed the research that has been undertaken to improve the EPG systems and services offered to viewers. This chapter will specify the requirements that are needed to design a generic EPG system that can be used to serve online viewing. These requirements are based on reviewing the contributions of previous research and determining the gaps that should be addressed in order to adapt to this mode of viewing. Generally, all EPG system designs comprise a means of retrieving data relating to the content, analysing this data, filtering the data and saving the resulting information. For a generic EPG designed to serve online viewing, the following requirements must be met.

1. The system can access online content from any and all online content sources.
2. The system can understand the different metadata formats that are used by the different sources.
3. The system can make recommendations based on programme content and not be restricted to specific channels.
4. The system should be scalable and responsive to both the number of users and content sources.
5. It should be employ contextual awareness to respond dynamically to the user's environment, network and social media interactions and incorporate these into its recommendations.

A generic EPG system should therefore contain a set of functions that comprise a content searching function, recommendation function and context-awareness function. Each one of these functions should then be characterized by features to satisfy the requirements and be implemented to provide the viewer with a list of recommended programmes that are collected from all the available online content sources and chosen to match their specific interests and viewing habits. Therefore, the next sections will discuss the requirements that should be met within each one of these main functions.

3.2 Content Searching System

Nowadays, searching for a specific piece of online content has become tedious because of the huge number of online content providers and the variety of the content they serve. Hence searching system must be able to interface with any and all content providers. Moreover, those providers may provide their metadata in different formats which in turn means that the searching system should be able to interpret and process programme metadata irrespective of how it is formatted. In summary, the searching system should satisfy the following requirements:

1. Scalability: the searching system should be able to connect with all the available online video content providers.
2. Extensibility: the system should have the ability to include any newly added online content provider. This therefore means that the architecture design of the EPG should be flexible to allow the addition of new content providers without affecting the remainder of the system.
3. Compatibility: the system should be able to communicate with all online content providers in spite of their differences in terms of the metadata formats they use. Therefore, the EPG system will need to be able to interpret and translate metadata to ensure compatibility across providers.
4. Standard Format: the system should have the ability to create a standard format for all the retrieved content metadata to describe that content regardless of their origin. This will help in preparing the retrieved metadata for the next stages of data processing within the EPG. Additionally, such a standard format should be able to keep the main features defined within the provider's existing metadata.
5. Analysing: the system should be able to analyse the retrieved metadata to combine it and remove duplicates.

3.3 Recommendation System

Based on the existing research that has investigated the design of recommendation systems and how the performance of the such processes can be improved, the generic EPG should be able to recommend programmes that are tailored to the user's interests and context. Moreover, the recommendation system should have the ability to predict what new programmes should be highlighted because they match the user's viewing pattern.

Evidence presented in the previous chapter suggested that to achieve this, multiple traditional recommendation schemes will need to be embedded within the recommendation system. Additionally, user context and dynamic environmental features should be taken into account to cater for viewing content on devices other than a Smart TV. Therefore, the designed recommendation system should satisfy the following requirements:

1. **Automated Recommendation:** the system should be able to recommend programmes automatically without the need for explicit feedback or ratings from the user.
2. **Extractability:** the system should have the ability to extract the needed information from data collected by the searching system that will be used in the recommendation process.
3. **Predictability:** the system should have the ability to analyse the extracted information in order to determine and predict programmes that should be recommended to the user.
4. **Content-based Recommendations:** the system should be able to recommend programmes based on the user's interests and preferred programmes. This in turn means that the system needs to maintain a record of the user's viewing history.
5. **Collaborative Recommendations:** the system should be able to recommend new programmes that differ from their historical viewing but that lie within their viewing pattern.
6. **Environmental Context:** in order to design a generic EPG system that can serve online viewing, the user's environmental features should be taken into account. This could help in recommending programmes that match the user's viewing device and network connectivity.
7. **User Context:** user context is one of the important features that used for improving the performance of recommendation systems. Therefore, the system should be able to use the user context (such as time, user device and network performance that is used by the user) into the recommendation process.
8. **Social Networks Recommendation:** social networks have become a significant source of additional information that can be used to tailor recommendations and therefore, the system should have the ability to interact with social networks.

3.4 Context-awareness Function

Based on the previous research, context-awareness has been used to enhance the performance of recommendation processes. In order to exploit context-awareness, a generic EPG system should be able to gather and process as much contextual information as possible including that which is static and dynamic. These requirements are:

1. **Collecting Contextual information:** the system should be able to collect different types of context such as the user context, device context and network context.
2. **Scalability:** the system should be able to store the context for all of its registered users which has implication for storage.
3. **Capability to Communicate:** due a requirement for providing a connection to social networks, the system should have the ability to communicate with social networks in order to retrieve relevant contextual information that can aid its recommendation process.
4. **Analysing:** the system should be able to analyse the collected contextual information and extract that which is relevant for each processing stage within the overall EPG.
5. **Automated System:** the system should have the ability to update contextual information in real-time.

3.5 Service Requirements

In order to allow the users to enjoy and benefit from the services offered by EPG, the system should have features which enable it to provide an easy way to display the final results. These system-requirements that should be satisfied are:

1. **Variety of Devices:** the system should be able to offer the same services for any device that has the ability to connect to the Internet and stream programme content.
2. **Timeliness:** the system should offer its services without minimal impact on the quality of the user's experience by ensuring that response times are as short as possible.
3. **Personalization:** the services should be personalized such that recommendations are tailored to each individual user.
4. **Privacy:** the system should provide user privacy for the offered services and recommendations.

5. User Interface: the EPG system should have a user interface to allow the user to access and interact with the services it offers in an easy communicate manner.

The next chapter will describe in detail the architecture of a generic EPG system that has been designed to meet this set of requirements.

CHAPTER FOUR

EPG SYSTEM

ARCHITECTURE

Chapter Four

4.1 Introduction

Based on the system requirements that have been discussed in the previous chapter, this chapter will describe in detail the proposed system architecture and design that can deliver a generic online EPG and satisfy the stated set of requirements. The system block diagram will be discussed by showing each part separately and describing the functionality of each block individually. Thereafter, the overall block diagram will be shown and the connection between the different parts will be illustrated.

4.2 General System Architecture

The system has been designed based on the experience and results of previous research that was presented in chapter two. Three core functions are implemented to serve a recommendation, searching and context-aware process. According to the previous requirements, a search system should be developed to access any and all the online content sources and handle the different metadata formats they use. Additionally, different types of information should be collected about the user which includes their interests, context, viewing pattern, viewing environmental features and the interests of their social network friends. This information will be used in personalizing the recommendations that are generated by the system. Figure 4.1 presents a high level block showing the core functions of the EPG system.

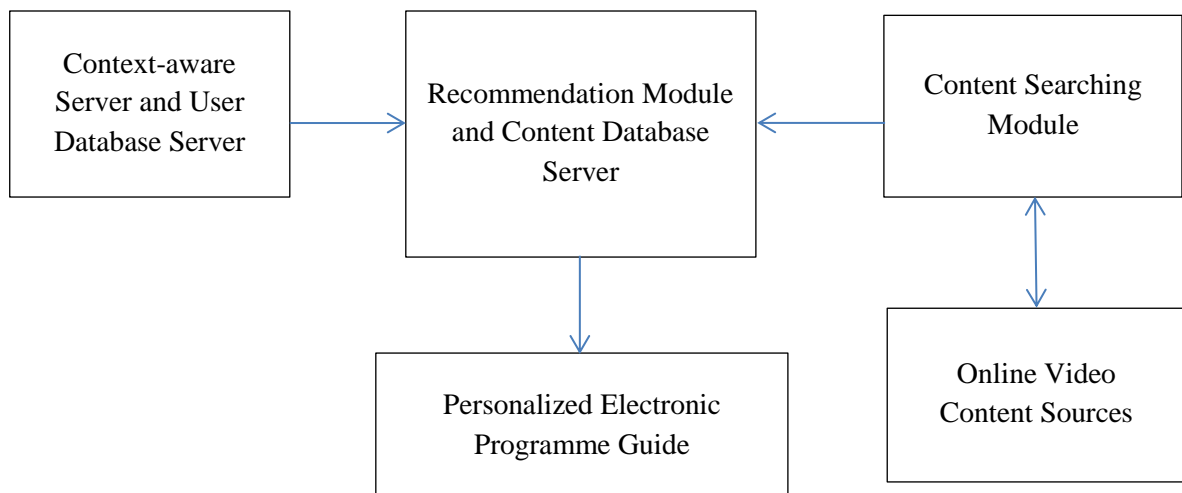


Fig. 4.1 Overview of the Proposed System

The proposed system consists of a content searching module, recommendation module, context-aware server and database server. The next sections will describe these main components separately with a detailed explanation for each component and its internal design.

4.3 Content Searching Module

The content searching module is the most important part in our proposed system. Due to the differences in metadata used by online content sources and their respective application programming interfaces (API), the main concept of the searching module is to provide multiple search agents where each agent connects to one specific online content source. The search agent is analogous to a TV tuner where it allows the system to receive and translate the retrieved information from each source. These search agents are then controlled by a search agent leader, which controls the overall searching operation and prepares the received metadata for the next stages. All of the different metadata formats that are used by the online providers are converted into a unified internal format which is used within the remainder of the EPG system.

An API approach is an architectural approach that revolves around providing programmable interfaces to a set of services to different applications serving different types of consumers. When used in the context of web development, an API is defined as a set of Hypertext Transfer Protocol (HTTP) request messages, along with a definition of the structure of a response message, which is usually in an Extensible Markup Language (XML) or JavaScript Object Notation (JSON) format (Basic Knowledge, 2015). Therefore, each search agent contained an interface block which ensure the interface with a specific online content provider. This block contains a program to give the system the ability to communicate (send requests and receive responses) with each online content source.

Each search agent consists of a search engine and interpreter as shown in figure 4.2. Each search engine receives a set of search keywords from the search agents' leader. Thereafter, the interpreter has a dual operation which is to convert the search keywords into a format that can be recognized by the online content source and then to convert the returned search results into a standardised format that can be used within the remainder of the EPG system. The conversion operation consists of two stages.

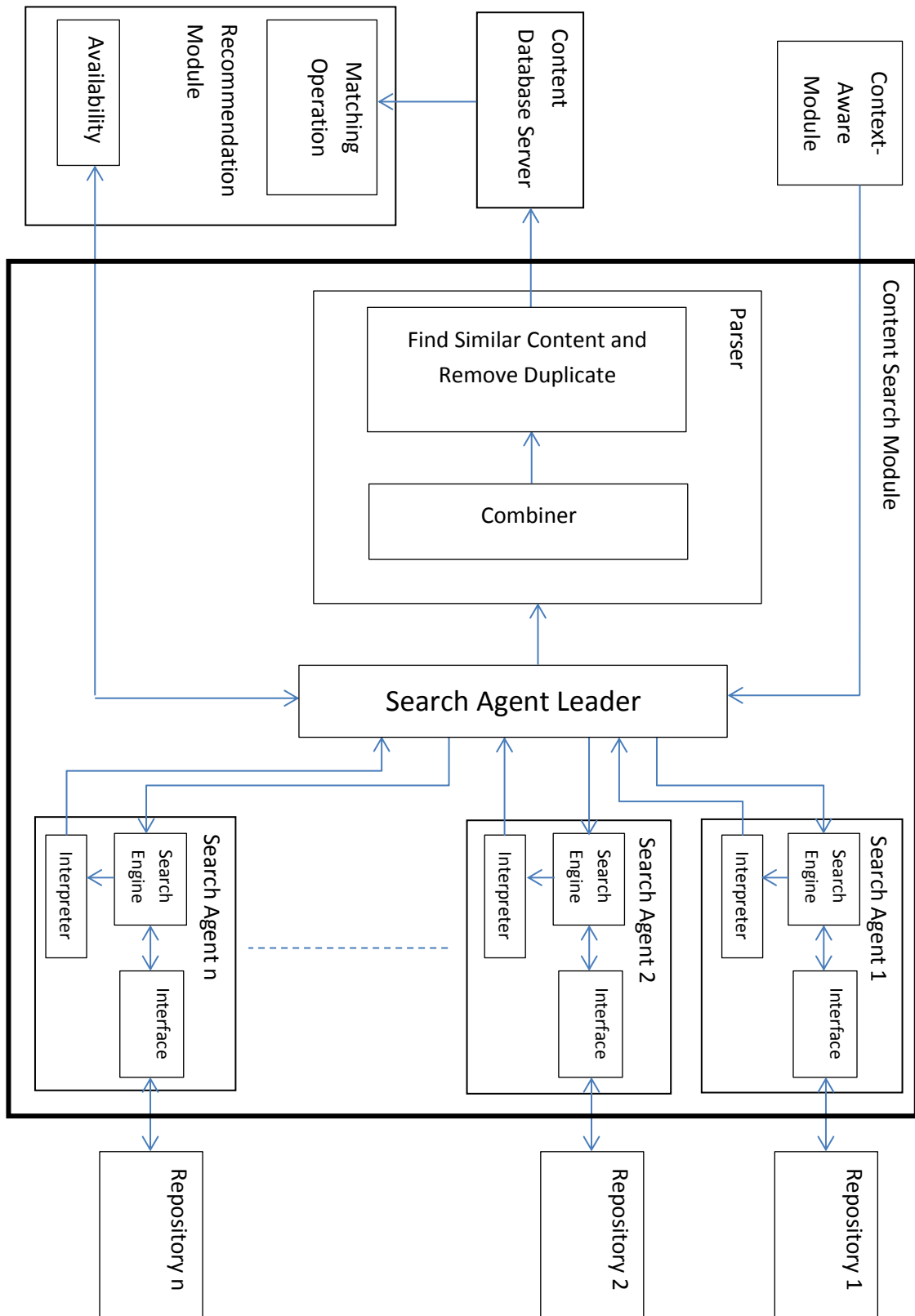


Fig. 4.2 Content Search Module

The first stage is to convert the retrieved metadata format from its raw form into XML format. This stage is implemented in the interpreter which is located within each search agent where it reads the raw metadata and extracts the key features of each programme. These key features are listed in Table 4.1. More details about these features will be displayed in the next subsection. The second stage is to convert the XML files which contain the programme's metadata into a matrix which contains the key features of each programme. The search agent leader implements this operation by reading the XML files and creates a matrix for each programme consisting of 21 rows where each row represents a specified programme's key feature.

| No. | Feature | Definition |
|-----|----------------------|--|
| 1 | Title | represents an identifying name given to a file or play or series. |
| 2 | Synopsis/Description | represents the words that use to explain the movie idea and clarify its context. |
| 3 | Genre | represents the categorization according to the setting of programme such as romantic, comedy, horror, cartoon, war, etc. |
| 4 | Type | represents the type of programme such as movie, news, show, documentary programme, series, etc. |
| 5 | Length of Content | represents the duration of the required content which might be short or long according to the user preferences. |
| 6 | Date | represents the date of publishing the programme or movie to the audience or on TV. |
| 7 | Popularity (Ratings) | represents the rate of the programme depending on the opinion of other users (previous ratings). |
| 8 | Format | represents the format of the video file (extension of the video file) which should be compatible with the user device. |
| 9 | Resolution | represents the definition of the picture or image such as Standard Definition, High Definition, etc. |
| 10 | Language | Represents the language of the required movie or play or series such as English, Chinese, French, etc. |

| | | |
|----|-----------------|--|
| 11 | Country | represents the country of movie or programme required by the user. |
| 12 | Producer's Name | represents those who prepare and then supervise the making of a film, and they may be an independent person or product company. |
| 13 | Actor | Represents the leading actor or actress (star) who plays the protagonist of a film or play or the singer of specified song. Sometimes it refers to the largest role in the film. |
| 14 | Director | represents the person who directs the making of a film. |
| 15 | Cast | represents the actors who were not playing main roles or parts in the film or programme. |
| 16 | Subtitle | represents a dialogue or commentary in a film or any TV programme and usually displayed in the bottom of the screen and it can be to a foreign language or the same language. |
| 17 | Author | represents the writer name. |
| 18 | Parental Level | represents the restrictions on what the kids can or cannot watch. |
| 19 | Age | represents for which target age group that this programme is focused. |
| 20 | Gender | represents the gender of the user. |
| 21 | Available | represents a tag defining either this programme or movie is available now or in another time. |

Table. 4.1 Programme Metadata Features

The search agent leader receives the search keywords from two locations, either from the context-aware module or from the availability functional block. The keywords that come from the context-aware module represent the titles of the watched programmes and the user preferences that have been provided by the user directly. The keywords that come from the availability functional block represent the titles of the programmes that have been recommended to the user from interactions with social networks or the clustering groups

(more details will be discussed in the next sections). Thereafter, the search agent leader forwards those keywords to all of the interpreters that are located in the search agents.

The search agent leader is programmed to receive search results from all of the search agents irrespective of how many there are. Therefore, when any new online content source is added to the system, all that is required is to develop a new search agent module with its functional parts (which includes a search engine, interpreter and interface). In this way, the system architecture is extensible and there is no need to change any other parts of the system when an extension of this type is required.

Having collected the responses from each search agent, the search agent leader sends the search results to the parser which contains two operations. The first operation is to combine the programmes' metadata matrices into a single matrix. This matrix named the content matrix and belongs to a specified user. The content matrix contains all of the programmes that are relevant to the target user and which also includes the programmes that have resulted from the different search operations carried out in the previous days. Each column of this content matrix represents a programme's metadata that itself consists of 21 key features. Thereafter, the content matrix is forwarded to the second operation which is to find the similar content and remove duplicates. This operation is needed to avoid repeating the same programme information in the database storage area. After this, the content matrix is stored in the content database server from where it is used by the remainder of the EPG system.

4.3.1 Extracting Metadata

Metadata is generally defined as data about data. It is textual information that is used to ensure that the contents of a file can be identified correctly. However, programme metadata can be divided into two categories; structural metadata and descriptive metadata. Structural metadata describes the technical format of the file itself which is usually added automatically. On the other hand, descriptive metadata includes information which will be read by the users of the file and is usually added by the producer of the file (ITV, 2015). However, it is important to note that content providers adopt different schemes to describe their content. Therefore, to achieve that, seamless interoperability standards should be developed that could be used for representing programme metadata. Interoperability means that any metadata provider using the standard representation would ensure that their

content is appropriately interpreted and processed on different platform implementations. As an example, the TV-anytime forum has adopted an XML-based and MPEG-7 Description Definition Language (DDL) as its representation format for metadata as shown in figure 4.3 (ETSI, 1983; Evain, 2000).

```

<complexType name="BasicContentDescriptionType">
  <sequence>
    <element name="Title" type="mpeg7:TitleType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="MediaTitle" type="mpeg7:TitleMediaType"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="ShortTitle" type="tva:ShortTitleType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="Synopsis" type="tva:SynopsisType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="PromotionalInformation" type="mpeg7:TextualType"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="Keyword" type="tva:KeywordType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="Genre" type="tva:GenreType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="ParentalGuidance"
      type="tva:TVAParentalGuidanceType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="Language" type="mpeg7:ExtendedLanguageType"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="CaptionLanguage" type="tva:CaptionLanguageType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="SignLanguage" type="tva:SignLanguageType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="CreditsList" type="tva:CreditsListType"
      minOccurs="0"/>
    <element name="AwardsList" type="tva:AwardsListType"
      minOccurs="0"/>
    <element name="RelatedMaterial" type="tva:RelatedMaterialType"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="ProductionDate" type="tva:TVATimeType"
      minOccurs="0"/>
    <element name="ProductionLocation" type="mpeg7:regionCode"
      minOccurs="0" maxOccurs="unbounded"/>
    <element name="CreationCoordinates" type="tva:CreationCoordinatesType"
      minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="DepictedCoordinates" type="tva:DepictedCoordinatesType"
      minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="ReleaseInformation" type="tva:ReleaseInformationType" minOccurs="0"
      maxOccurs="unbounded"/>
    <element name="Duration" type="duration" minOccurs="0"/>
    <element name="PurchaseList" type="tva:PurchaseListType" minOccurs="0"/>
  </sequence>
</complexType>

```

Fig. 4.3 XML-Based Metadata Example (ETSI, 1983)

However, these standardizations and structure are not used by everyone and certainly not by the online content providers. In fact, the set of online content providers adopt different

structures and formats in defining their programme metadata. Figure 4.4 shows the metadata format that is used by YouTube which is written in JSON format. Therefore, a key part of our research has been to investigate and analyse the metadata formats used by different online content providers such as BBC i-Player, ITV Player, Netflix, YouTube and IPTV systems (Netflix, 2016, Linux Centre, 2010, Open IPTV Forum, 2014, YouTube, 2014).

```
{
  "version": "1.0",
  "encoding": "UTF-8",
  "feed": {
    "xmlns": "http://www.w3.org/2005/Atom",
    "xmlns$openSearch": "http://a9.com/-/spec/opensearchrss/1.0/",
    "xmlns$gd": "http://schemas.google.com/g/2005",
    "xmlns$yt": "http://gdata.youtube.com/schemas/2007",
    "xmlns$media": "http://search.yahoo.com/mrss/",
    "id": {"$t": "http://gdata.youtube.com/feeds/api/videos/"},
    "updated": {"$t": "2015-01-29T16:36:38.158Z"},
    "category": [{"scheme": "http://schemas.google.com/g/2005#kind", "term": "http://gdata.youtube.com/schemas/2007#video"}],
    "title": {"$t": "Videos matching: messi", "type": "text"},
    "logo": {"$t": "http://www.gstatic.com/youtube/img/logo.png"},
    "link": [{"rel": "alternate", "type": "text/html", "href": "http://www.youtube.com"},
    {"rel": "http://schemas.google.com/g/2005#feed", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/"},
    {"rel": "http://schemas.google.com/g/2005#batch", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/batch/"},
    {"rel": "self", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos?alt=json&q=messi"},
    {"rel": "next", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos?alt=json&q=messi"}],
    "author": [{"name": {"$t": "YouTube"}, "uri": {"$t": "http://www.youtube.com/"}}],
    "generator": {"$t": "YouTube data API", "version": "2.1", "uri": {"$t": "http://gdata.youtube.com/"}},
    "openSearch$totalResults": {"$t": "1000000"},
    "openSearch$startIndex": {"$t": "1"},
    "openSearch$itemsPerPage": {"$t": "5"},
    "entry": [{"id": {"$t": "http://gdata.youtube.com/feeds/api/videos/gw2UwuPTekk"},
    "published": {"$t": "2014-10-06T14:59:25.000Z"}, "updated": {"$t": "2015-01-29T05:42:22.000Z"},
    "category": [{"scheme": "http://schemas.google.com/g/2005#kind", "term": "http://gdata.youtube.com/schemas/2007#video"}, {"scheme": "http://gdata.youtube.com/schemas/2007/categories.cat", "term": "Sports", "label": "Sport"}],
    "title": {"$t": "Lionel Messi ? Magic Skills 2014-2015 |HD|", "type": "text"},
    "content": {"$t": "Messi Best Skills and Goals so far in 2014-2015. Messi Magic skills, dribbling, passing, c", "type": "text"},
    "link": [{"rel": "alternate", "type": "text/html", "href": "http://www.youtube.com/watch?v=gw2UwuPTekk&feature=youtu"},
    {"rel": "http://gdata.youtube.com/schemas/2007#video.related", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/gw2UwuPTekk/related"},
    {"rel": "self", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/gw2UwuPTekk"},
    {"rel": "http://schemas.google.com/g/2005#mobile", "type": "text/html", "href": "http://m.youtube.com/watch?v=gw2UwuPTekk"}],
    "author": [{"name": {"$t": "I N M E S S I O N A N T E"}, "uri": {"$t": "http://gdata.youtube.com/feeds/api/users/INMESSEIONANTE"}}],
    "gd$comments": {"gd$feedLink": {"rel": "http://gdata.youtube.com/schemas/2007#comments", "href": "http://gdata.youtube.com/feeds/api/videos/gw2UwuPTekk/comments"}},
    "yt$hd": {"$t": "1080i"},
    "media$category": [{"$t": "Sports", "label": "Sport", "scheme": "http://gdata.youtube.com/schemas/2007/categories.cat", "term": "Sports", "label": "Sport"}],
    "media$content": [{"url": "http://www.youtube.com/v/gw2UwuPTekk?version=3&f=videos&app=youtu"}, {"url": "rtsp://r2--sn-cg07luez.c.youtube.com/CiILENY73wIaGQ1JtPjWpRtgRMYC"}, {"url": "rtsp://r2--sn-cg07luez.c.youtube.com/CiILENY73wIaGQ1JtPjWpRtgRMYC"}],
    "media$description": {"$t": "Messi Best Skills and Goals so far in 2014-2015. Messi Magic skil"},
    "media$keywords": {"$t": ""},
    "media$player": {"url": "http://www.youtube.com/watch?v=gw2UwuPTekk&feature=youtu"},
    "media$thumbnail": [{"url": "http://i.ytimg.com/vi/gw2UwuPTekk/0.jpg", "height": 360, "width": 480}, {"url": "http://i.ytimg.com/vi/gw2UwuPTekk/1.jpg", "height": 90, "width": 120}, {"url": "http://i.ytimg.com/vi/gw2UwuPTekk/2.jpg", "height": 90, "width": 120}, {"url": "http://i.ytimg.com/vi/gw2UwuPTekk/3.jpg", "height": 90, "width": 120}],
    "media$title": {"$t": "Lionel Messi ? Magic Skills 2014-2015 |HD|", "type": "plain"},
    "yt$duration": {"seconds": "7404"},
    "gd$rating": {"average": "4.77672", "max": "5", "min": "1", "numRaters": "13024", "rel": "http://schemas.google.com/g/2005#rating"},
    "yt$statistics": {"favoriteCount": "0", "viewCount": "3301330"}}, {"id": {"$t": "http://gdata.youtube.com/feeds/api/videos/e9B3eyJjSfc"},
    "published": {"$t": "2015-01-09T09:03:45.000Z"}, "updated": {"$t": "2015-01-28T01:13:18.000Z"},
    "category": [{"scheme": "http://schemas.google.com/g/2005#kind", "term": "http://gdata.youtube.com/schemas/2007#video"}, {"scheme": "http://gdata.youtube.com/schemas/2007/categories.cat", "term": "Sports", "label": "Sport"}],
    "title": {"$t": "Lionel Messi - The Record Breaker - Documentary (HD)", "type": "text"},
    "content": {"$t": "Lionel Messi - The Record Breaker - Documentary (HD) Lionel Messi - The Record Breaker - D", "type": "text"},
    "link": [{"rel": "alternate", "type": "text/html", "href": "http://www.youtube.com/watch?v=e9B3eyJjSfc&feature=youtu"}, {"rel": "http://gdata.youtube.com/schemas/2007#video.related", "type": "application/atom+xml", "href": "http://gdata.youtube.com/feeds/api/videos/e9B3eyJjSfc/related"}]}]}

```

Fig. 4.4 YouTube Metadata

A comparison has been made between provided metadata by several online and metadata providers by examining the features that are included in the metadata and the way of defining these features by those providers. Table 4.2 describes the programme's features that are included in the metadata provided by the examined online content providers and IPTV's EPG standardization.

| Parameters | BBC Iplayer | ITV Player | YouTube | Netflix | IPTV System | EPG Standardization |
|-------------------|--------------------|---------------------------|-------------------------------------|-----------------------------|--------------------------|-------------------------|
| Title | Name (name short) | Title (series, programme) | Title | Name | Title/Name | Title/short, Long Title |
| Synopsis | Desc (description) | Synopsis | Description | Synopsis | Description | Synopsis |
| Genre | Categories | Genre | Category | Genre | Genre | Genre |
| Type | Type | x | x | Type | Genre Type | Type |
| Subtitle | Subtitle | Closed Captions Present | x | x | Subtitle | Closed Captions Present |
| Date | Firstcast | Copyright Year | Posting Date | Release Year | Time of Production | Year/Day and Year |
| Available | Available | x | x | Run Time | Free or Subscribed | Purchase |
| Popularity | x | x | Number of Viewer /Number of Ratings | Rating and Average Rating | Users Profiles | Rating |
| Length of Content | Duration | Total Duration | Length | Duration | Duration | Duration |
| Actor | x | x | x | Actors | Credits | Credits List (Role) |
| Director | x | x | x | Director | Credits | Credits List (Role) |
| Producer's name | x | Originator | x | Producer and Production Co. | Credits | Credits List (Role) |
| Resolution | x | Picture Format | x | Resolution | Resolution | Frame Rate/Pic Form |
| Format | Modes | Video Codec | x | x | File Format | File Format |
| Language | x | Primary Audio Language | x | In Language | Primary Caption Language | Language |
| Country | x | x | x | x | Location of Production | Production Location |
| Parental Level | x | x | x | Is Family Friendly | Parental Guidance | Parental Guidance |
| Age | x | x | x | x | User Identifier | x |
| Gender | x | x | x | x | User Identifier | x |
| Cast | x | x | x | Actors (Cast) | Credits | Credits List (Role) |

Table 4.2 Programme's Features Provided by Several Online Content Providers

Based on the information that has been extracted from these providers, a unified format has been developed and adopted in our generic EPG design. This unified format represents the programme metadata as a matrix of 21 key features as shown in Table 4.1 which describes each programme. Inevitably, the features which are included in the unified format are not found in every online content provider's existing metadata. However, the EPG system has been designed to compensate for any missing elements. That means, the system can describe the programme with the available features even some of the features are missed. More details will be explained in the implementation sections.

Each parameter (of the 21 parameters) in the features' matrix can include multiple terms. As an example, the genre feature can include cartoon, comedy and adventures in the same feature where this programme could be recommended for any one of those genres. Similarly, other features have multiple terms such as resolution, format, type, actor, cast and so on. Note that features such as these require more processing. As an example, the programme title has two terms which are the main title (keyword) and the long title. Therefore the searching operation uses the main title, such as searching for programmes relevant to "Top Gear" while the matching operation that is applied in the recommendation module uses the long title such as (Top Gear: Series 24: Episode 12). Therefore, in the searching operation, the system will search for programmes relevant to the keyword and thereafter, the matching operation will compare the long title of each programme and remove any similar episodes.

Note that for features that contain multiple terms such as genre (which could be defined as a drama, romance and adventures), the system applies the matching operation that examines all terms with the programme's genre parameter. Therefore, the programme could be recommended for three different types of genre. In this way, the system will not repeat this programme three times. Similarly, this applies for all the other feature parameters which could contain multiple terms. The system can also recognize similar programmes that are retrieved from different online sources because the location of the programme is removed from the matching operation.

This unified format is used in different operations within the system such as storing content information, specifying the viewing history, matching operations and the recommendation process. This flexibility in the design of the searching module allows the

system to be generic with the ability to retrieve programme metadata from all online content providers.

4.4 Recommendation Module

The second main module in the proposed system is the recommendation module. Based on the experience and results of other researchers arising from the research covered in the literature review, multiple recommendation approaches have been employed in this module in order to enhance the recommendation process and provide the user with recommendations that are better tailored to their interests and preferences and which reflect the current viewing environmental features. Those approaches are content-based, collaborative filtering and context-awareness supplemented by an integration with social networks. Figure 4.5 shows a block diagram of the recommendation module. Three lists of recommendations are generated in this module. These lists are user preferences (based on user viewing history), user cluster (based on collaborative filtering approach) and social network recommendations (based on user's friends in social networks).

4.4.1 User Viewing History

The first list of recommended programmes is created based on the user's viewing history. This is done by matching the programmes' metadata which is saved in the content matrix (in the content database server) with the user's viewing history which is saved in the context-aware module. Here, both the programme's metadata and user's viewing history are represented by matrices.

The programmes' metadata columns represent the resulting programmes from searching operation based on viewing history plus the programmes that have been resulted from previous searching operations while the user's viewing history columns represent the programmes that have been watched by the user. The viewing history reflects the user's viewing pattern and what the user likes to watch at the different hours during the day. More details about the user's viewing history will be discussed in the context-aware module section. The matching operation is implemented by calculating the similarity between each column in the programmes' metadata and each column in the user's viewing history. The similarity is calculated by different methods and so the next section contains a brief illustration of the similarity measurements and some of the common methods which are used in similarity calculations.

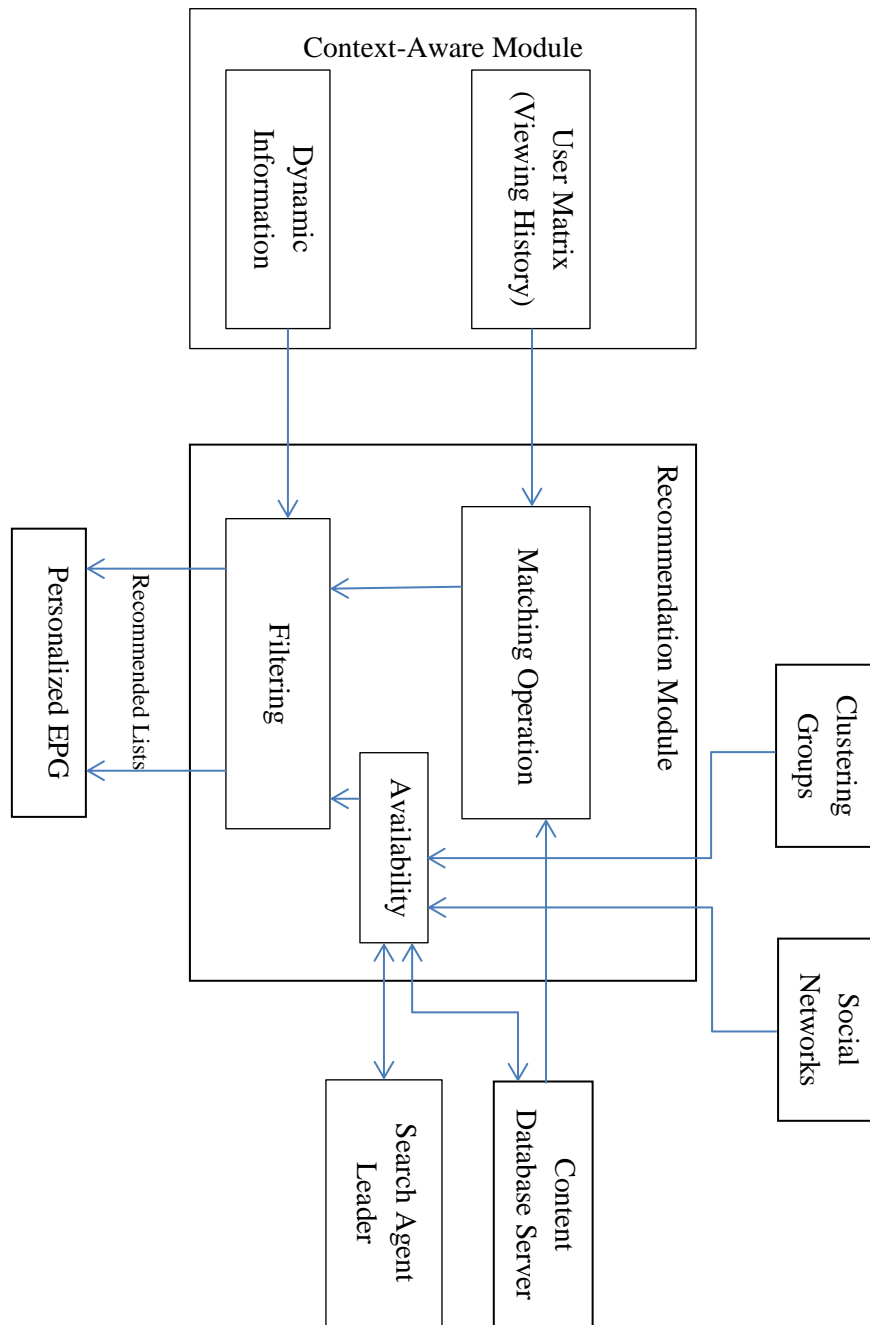


Fig. 4.5 Recommendation Module

4.4.1.1 Similarity Measures

Similarity measures are defined in statistics as a real-value function that quantifies the similarity between two objects. This measure reflects the degree of closeness of two objects in each feature when those objects are represented by group of features. Similarity measurements are needed for clustering, classification, feature selection and searching. Similarity measures are in effect the inverse of distance metrics which calculate the distance between two objects when those objects are represented by vectors. However, similarity measures have different categories depending on the data type that is used to represent the objects such as binary variables, quantitative variables or symbolic descriptions. As an example, similarity measures map the similarity between the symbolic descriptions of two objects into a single numeric value. However, there is no measure that is universally best for all kinds of objects. Different methods are used to measure the similarity or the distance between any two objects. The following are some of these methods that are used to measure the similarity between two objects when those objects are represented as vectors.

4.4.1.1. a Cosine Similarity

Cosine similarity is one of the most popular methods that is used to measure the similarity especially in information retrieval applications and clustering. The similarity corresponds to the correlation between the vectors that represent objects. Here, vector represents a collection of objects (vector space) which may be added together or multiplied by numbers. This is quantified as the cosine of the angle between two vectors, hence, cosine similarity. To calculate the cosine similarity, equation 4.1 is used where the two vectors of the target objects should be given.

$$similarity = \frac{A \cdot B}{|A| \times |B|} = \frac{\sum_{i=1}^n A_i B_i}{\sqrt{\sum_{i=1}^n A_i^2} \sqrt{\sum_{i=1}^n B_i^2}} \quad \dots\dots\dots \text{Equation 4.1}$$

Where A and B are the vectors while A_i and B_i are components (objects) of those vectors. As a result, the cosine similarity is non-negative and bounded between [0, 1].

4.4.1.1. b Jaccard Coefficient

The Jaccard similarity coefficient, also known as the Jaccard index, is a statistic used to compare the similarity of objects or sets. Jaccard coefficient measures the similarity as an intersection (number of objects which are shared between both sets) divided by the union of the sets (shared and un-shared objects) as defined by equation 4.2.

$$SIM_J(A, B) = \frac{|A \cap B|}{|A \cup B|} = \frac{|A \cap B|}{|A| + |B| - |A \cap B|} \dots\dots\dots \text{Equation 4.2}$$

Where *A* and *B* are the objects or sets. For a text document, Jaccard coefficient measures the similarity as comparing the sum weight of shared terms to the sum weight of terms that are exist either of the two documents but are not the shared terms. Its range is bounded between [0, 1].

4.4.1.1. c Pearson Correlation Coefficient

The Pearson Correlation coefficient is a measure of the linear correlation between two vectors. Unlike the other similarity measures, the Pearson correlation coefficient value is bounded between [+1, -1] where +1 means that the two vectors are similar. There are different forms of the Pearson correlation coefficient formula, but when applied to a set of data the common formula used is represented by equation 4.3:

$$SIM_P(A, B) = \frac{\sum_{i=1}^n (A_i - \bar{A})(B_i - \bar{B})}{\sqrt{\sum_{i=1}^n (A_i - \bar{A})^2} \sqrt{\sum_{i=1}^n (B_i - \bar{B})^2}} \dots\dots\dots \text{Equation 4.3}$$

Where *A* and *B* are vectors of *n* dimentions and *A_i* and *B_i* are the components of those vectors.

$$A = \{A_1, A_2, A_3, \dots, A_n\} \dots\dots\dots \text{Equation 4.4}$$

$$B = \{B_1, B_2, B_3, \dots, B_n\} \dots\dots\dots \text{Equation 4.5}$$

$$\bar{A} = \frac{1}{n} \sum_{i=1}^n A_i \dots\dots\dots \text{Equation 4.6}$$

$$\bar{B} = \frac{1}{n} \sum_{i=1}^n B_i \quad \dots\dots\dots \text{Equation 4.7}$$

Note that \bar{A} and \bar{B} are the sample mean of A and B respectively (Adomavicius & Tuzhilin, 2005; Huang, 2008).

4.4.2 User Clusters

The second list of recommended programmes is created using the collaborative filtering technique. This is done by clustering users into several groups which bring together all the viewers who have a similar viewing pattern. The recommendation module retrieves the required information from the context-aware module. This information includes the categories of the watched programmes which are stored in the viewing history of each user. The proposed system currently defines 36 programme categories as shown in table 4.3. These categories are specified based on information collected from different online content sources such as BBC-iPlayer, YouTube, The Movie Database and Dailymotion. Table 4.4 shows the categories that are used in these online content sources. Note that, some of those categories have been converted to obtain a unified set of categories which are used within the EPG system. However, the system has the ability to add any new category to the defined group of categories if a user selects a programme with a category not included in this group. Thereafter, the system applies a clustering algorithm where each group includes the most similar users in their selected categories. The next section describes examples of the clustering algorithms that are used for this purpose. From these groups, the system determines the programmes that have been watched by each group and suggests programmes from this group for the target user which they haven't yet watched. Figure 4.6 provides a block diagram of the user clustering process.

Then, the system checks the list of recommended programmes for each user in the content database server to decide if those programmes are already available or that they should be searched for in online video content sources. The programmes that already exist in the content database server (content matrix) are saved in the user profile in the context-aware module (more details will be described in the context-aware module section) while the programmes that do not exist in the content database server are forwarded to the search agent leader to search the online video content sources. The same steps are applied for searching those programmes as described previously in the content search module. The programmes that are resulted from the search operation are stored in the content database

server (in the content matrix that belongs to the target user). Thereafter, a list of those programmes is forwarded to the user profile to join the first group (those programmes that are found in the content database server before searching the online content sources) and be ready to be displayed to the user when they login to their EPG.

| No. | Category | No. | Category |
|-----|------------------|-----|---------------------|
| 1 | Autos/Vehicles | 19 | Weather |
| 2 | Comedy | 20 | News |
| 3 | Factual | 21 | Animals |
| 4 | Movies | 22 | Short Movies |
| 5 | HowTo/Style | 23 | Gaming |
| 6 | People | 24 | Videoblogging |
| 7 | Shows | 25 | Entertainment |
| 8 | Action/Adventure | 26 | Science/Technology |
| 9 | Romance | 27 | Classics |
| 10 | Documentary | 28 | Family |
| 11 | Drama | 29 | Foreign |
| 12 | Sport | 30 | Scientific Fiction |
| 13 | Travel/Events | 31 | Thriller |
| 14 | Shorts | 32 | Trailers |
| 15 | Music | 33 | History |
| 16 | Horror | 34 | Western |
| 17 | Animation | 35 | Religion and Ethics |
| 18 | Learning | 36 | Children |

Table 4.3 List of Used Categories

| BBC iPlayer | YouTube | TMDB | Dailymotion |
|---------------------|----------------------|-----------------|------------------------|
| Children | Film & Animation | Action | Movies |
| Comedy | Autos & Vehicles | Animation | Travel |
| Drama | Music | Crime | Webcam |
| Entertainment | Pets & Animals | Drama | Celeb |
| Factual | Sports | Fantasy | Lifestyle & Howto |
| Learning | Short Movies | History | News |
| Music | Travel & Events | Music | Music |
| News | Gaming | Romance | Creative |
| Religion and Ethics | Videoblogging | TV Movie | Sports |
| Sport | People & Blogs | War | Tech |
| Weather | Comedy | Adventure | TV |
| | Entertainment | Comedy | Comedy & Entertainment |
| | News & Politics | Documentary | Cars |
| | Howto & Style | Family | Gaming |
| | Education | Foreign | Animals |
| | Science & Technology | Horror | Education |
| | Movies | Mystery | Kids |
| | Anime/Animation | Science Fiction | |
| | Action/Adventure | Thriller | |
| | Classics | Western | |
| | Documentary | | |
| | Drama | | |
| | Family | | |
| | Foreign | | |
| | Horror | | |
| | Sci-Fi/Fantasy | | |
| | Thriller | | |
| | Shorts | | |
| | Shows | | |
| | Trailers | | |

Table 4.4 Online Content Sources' Categories

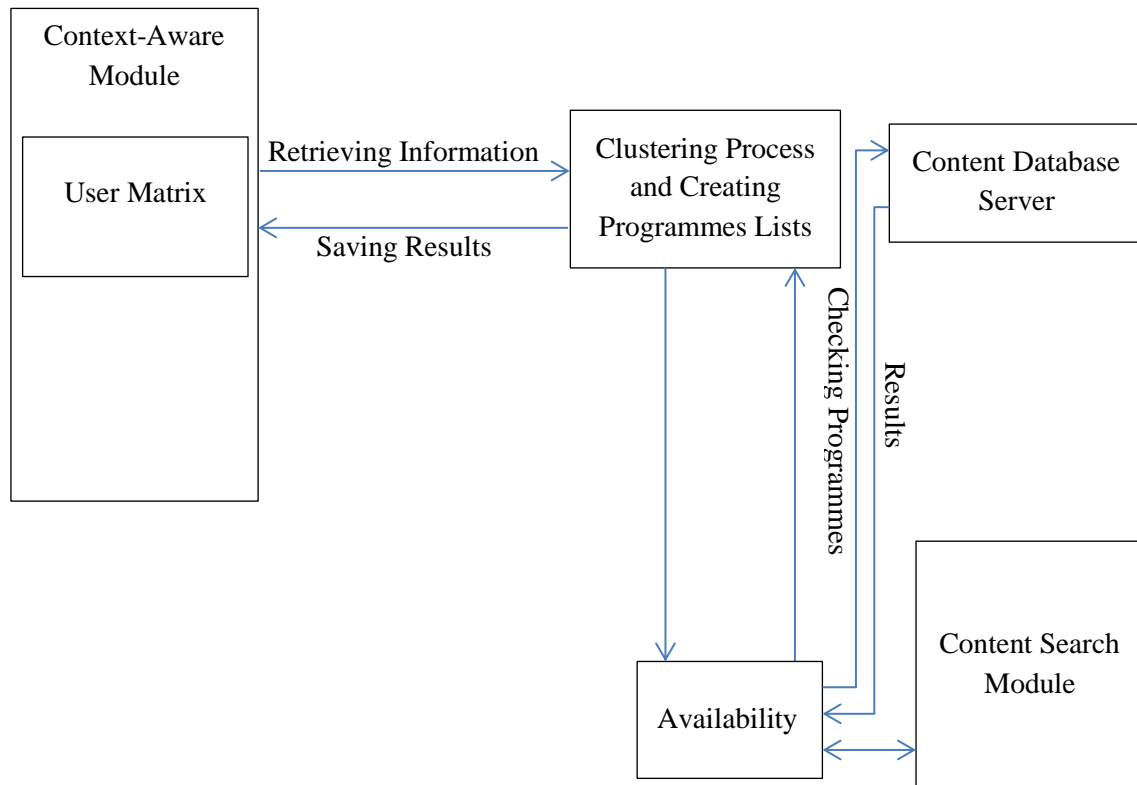


Fig. 4.6 Users' Clustering Process

4.4.2.1 Clustering Algorithm

Clustering is defined as grouping a set of objects in such a way that similar objects in some sense are grouped together in the same cluster. This can be achieved by various algorithms. One of the popular concepts in clustering is to measure the distance among the cluster members. Cluster analysis is not a one specific task, but it is an iterative process of knowledge discovery that involves trial and failure. This sub-section will give two examples of the simplest algorithms that are widely used in the clustering process.

4.4.2.1. a K-Means Clustering

K-means is one of the simplest learning algorithms for clustering and is popular in data mining. The main idea of K-means is to define k points (centroids) with one for each cluster. These centroids should be located as far as possible from each other in the area of data set where Euclidean distance is calculated between each centroid of clusters. Then,

each point belonging to the given data set will be taken and associated with the nearest centroid. The first step is completed when all the points which belong to the data set are included in the clusters. Thereafter, new k centroids will be calculated which should be different from the centroid used in the previous step. Again, a new binding will be done between each point in the data set and these new centroids. The same steps will be repeated (as a loop) until no more changes in the location of the specified centroids occur which means that centroids do not move any more. The quality of clustering is determined by the error function described in equation 4.8:

$$E = \sum_{j=1}^k \sum_{i_l \in C_j} |i_l - w_j|^2 \quad \dots\dots\dots \text{Equation 4.8}$$

Where i_l is the data point, $l \in \{1, \dots, n\}$, w_j is the cluster centre, $j \in \{1, \dots, k\}$ and C_j is the j th cluster (Alsabti, Ranka, & Singh, 1997; Ding & He, 2004; Likas, Vlassis, & Verbeek, 2003).

4.4.2.1. b K-Nearest Neighbour

K-nearest neighbour is an algorithm that stores all the available cases and classifies the new cases based on the similarity measurements with those previous cases. Its input consists of the K closest cases in the feature space while the output is the class member. The main idea of K-nearest neighbour is to identify the k nearest neighbours to any new case or object in order to estimate its class. This could be achieved by measuring the similarity (as an example distance measure) with exist cases by specifying the value of K which represent the number of required nearest neighbours in that algorithm. Then, based on the number of nearest neighbours, a decision is made for that case or object where it will be considered to belong to the class which has more nearest neighbours to that new case or object than the other classes (Theodoridis & Koutroumbas, 2003). Note that these algorithms are used in collaborative filtering approach for clustering users into groups based on clustering criteria such as programme's category.

4.4.3 Social Networks Recommendations

The third list of programmes results from integration of social networks into the recommendation process. Social networks add an important contribution to the recommendation process by supporting the recommendation system with another source of information which in this case is a list of programmes that have been suggested by a user's friends. This has been done through creating an application that runs within a specific social network and which asks the user to join. This application then retrieves information relevant to the programmes that the user has tagged as "like", "watched" or "want to watch" within their profile. Additionally, the application retrieves the user's list of friends and from those friends' profiles, the application retrieves a similar set of programme information. All the information will be listed under a 'nickname' created by the user during their registration operation while the system will use an ID number (defined within the social network) to recognize each user.

The EPG system therefore requires an interface with each social network to retrieve the required information. Figure 4.7 shows the block diagram for how the EPG connects to a social network. The system translates the retrieved information through the interpreter into the unified format (the same format used in the system which created in the content search module). This makes analysing and merging the information easier. The retrieved information includes the titles and categories of movies and TV shows that have been liked or watched by the user and their friends (but only the friends who joined the application). Additionally, it includes the movies that the user and their friends tagged as "want to watch". Hence, the same concept as used for the content search module applies here in that expanding the EPG to handle inputs from multiple social networks only requires the addition of an interface module for each social network that is to be added. Social networks, like content providers, use different formats for presenting their information through their own bespoke API.

The system then analyses the resulted information from the interpreter to find the titles of programmes that have been mentioned through the social network. This list of titles must then be checked through a 'checking availability block' to identify whether any of them is already contained within the content database server. This database server contains the programmes that have been searched and located previously. More details will be provided about the content database server in the next section. For those items that are not found in

the content database server, the availability block sends a list of required programmes to the search agent leader to search for relevant programmes in the online content providers.

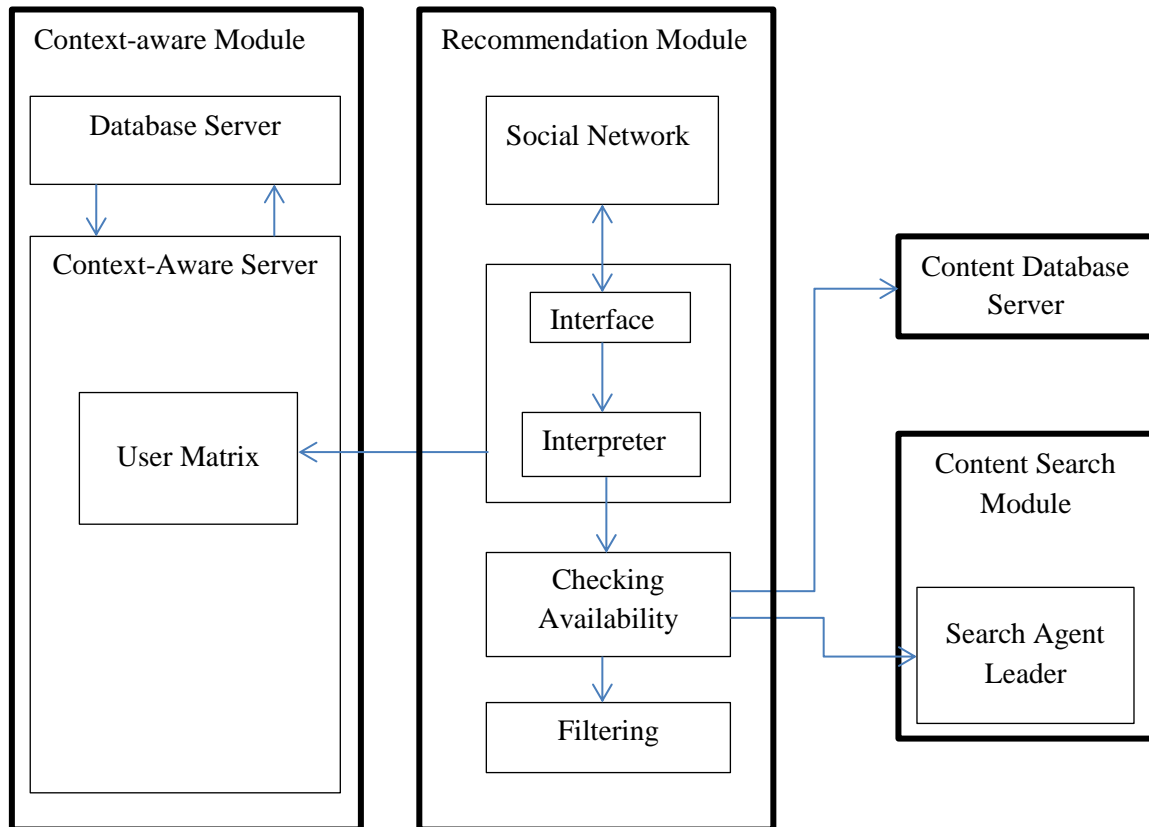


Fig. 4.7 Social Network Contribution

Additionally, the system analyses the retrieved information to determine the number of friends who have common liked or watched programmes. This information is used in ranking programmes in the final list that is displayed to users. The list of programmes, which is created based on the social network contribution, is stored in the user matrix contained within the context-aware module. Note that, the list of programmes will contain only the number of friends who watched or recommended this programme within the social network and not any information about the friends themselves. Despite the fact that several researchers have shown the benefits of knowing the identity of the person making a recommendation, such as the study by Bonhard (Bonhard & Sasse, 2006), a decision has

been taken to not reveal this within the EPG. The reason of this is to provide more privacy and to take into account that not all users are willing to share their viewing information.

On the other hand, if the user doesn't have a social network account or doesn't join the application within the social network for any reason, the system will continue to create the recommended list of programmes based on the other recommendation schemes. Similarly, for the cold-start problem, the system will depend on the collaborative filtering technique only to recommend programmes for the user when the system has no information about the user's viewing history.

This list together with those that have been created based on the other recommendation schemes is then filtered according to the user's environmental features. This is done by installing an application in the user device to extract information relevant to the viewing device and the connected delivery network. For example, the application provides the EPG system with information about the device type, screen resolution and available network bandwidth. More details will be mentioned about retrieving the user environmental features in the context-aware module.

In order to reduce the perceived time for recommending programmes for users, the EPG system has been designed to perform the search and recommendation operations as background tasks whilst the user is logged off. The only operation that is performed when the user logs onto the EPG is filtering of the final programme lists according to the environmental features.

4.5 Context-aware Module

The context-aware module is responsible for collecting and saving information required for the recommendation process. Figure 4.8 shows the main functional blocks of the context-aware module. It consists of a context-aware manager, information retrieval manager and user database server. The context-aware manager records each programme that has been watched by a user and saves it in their viewing history. The context-aware manager is responsible for creating the user matrix. This matrix includes the user's viewing history and other information relevant to the user such as the information retrieved from the social networks, the recommended list of programmes created by the different recommendation schemes and information resulted from the analysing processes such as the selected programmes' categories (which are used in users clustering) or user's

preferences (keywords). Hence, the context-aware manager has an internal connection with most of the other parts of the EPG architecture because all these other parts either produce information that should be saved in the user matrix or require some information that is generated by the context-aware manager.

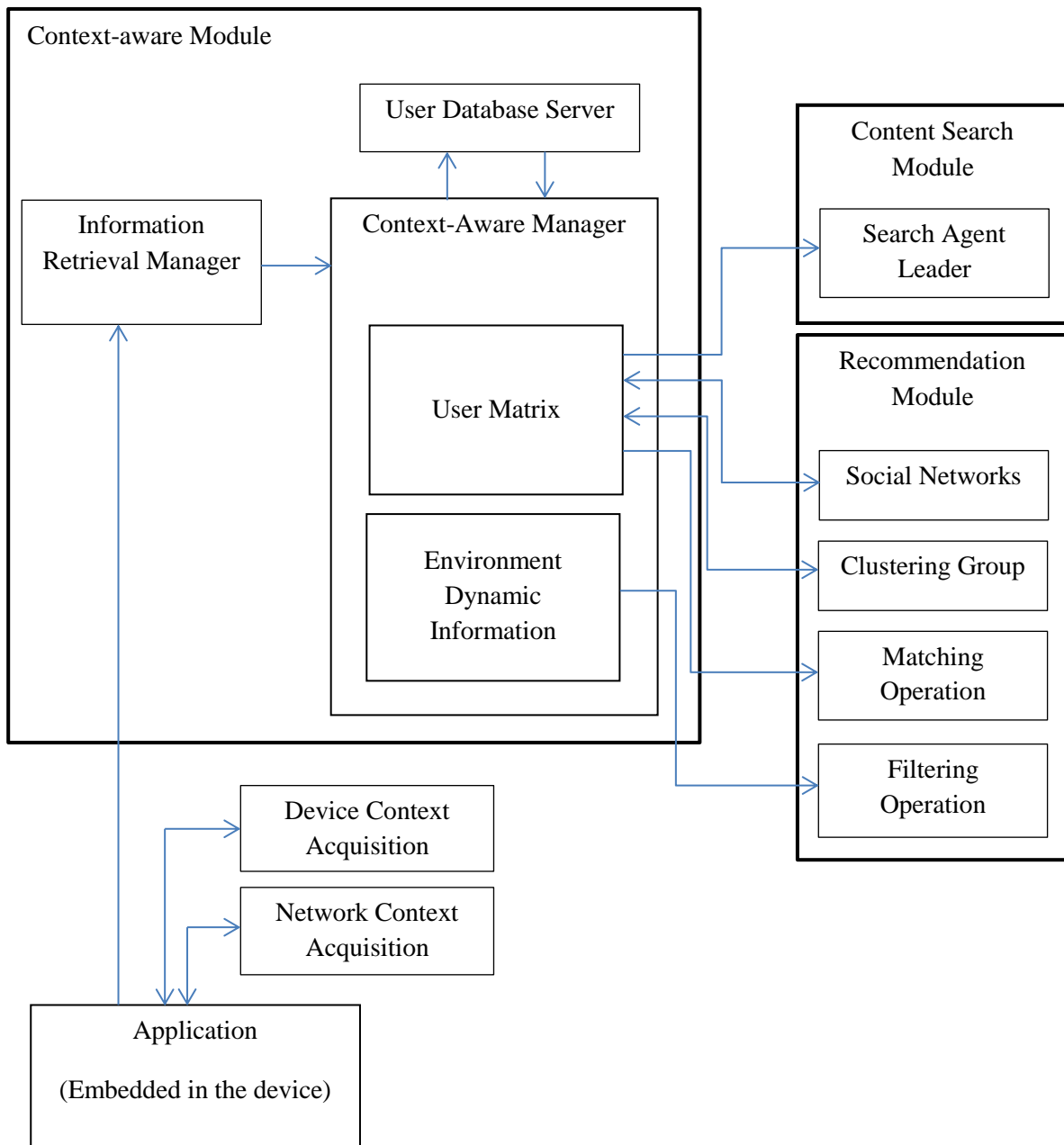


Fig. 4.8 Context-aware Module

The context-aware manager is connected to each social network in order to save retrieved information following the translation operation. That includes the user name, user ID, liked and watched items, list of friends and the friends' liked and watched items. The context-aware manager analyses the collected information to find items that are common and different between the user and their friends. Moreover, it determines the number of friends who liked or watched a specified item in order to give this item a higher rank in the recommended list. All of this information is stored in the user matrix which used in the recommendation operation.

Additionally, the context-aware manager is connected to the clustering process to define the users who are in the same group as the user and from those to collect programmes to recommend. Thereafter, the recommended list of programmes is saved in the user matrix which is displayed to the user when they log into the EPG system. Moreover, the context-aware manager is connected to the matching operation block where the recommendations based on the viewing history are created. The context-aware manager provides information about the user's viewing history which is saved in the user matrix and classified into 24 parts where each part covers one hour per day. In this way, a user's viewing pattern will be specified according to what they prefer to watch at each hour. Then, after completing the matching operation, the recommended list of programmes is classified as a 24 column matrix where there is a recommended list of programmes for each hour of the day. These recommendations will also be saved in the user matrix. Furthermore, the context-aware manager is connected to the filtering block to provide information about the user's viewing device and network connectivity.

The context-aware manager is also connected to the content search module to provide the search agent leader with the titles of programmes that should be used to search the online video content sources. Furthermore, the context-aware manager receives the search keywords that are provided directly by the user through the information retrieval manager and forwards them to the search agent leader.

The user matrix information is stored in the user database server which is located in the context-aware module. This matrix includes permanent information, such as the user name and user ID, and variable information such as the viewing history, information and recommendations from friends. The variable part is reviewed, and potentially changed, daily when the user uses their EPG resulting in the viewing history being updated and also

as a result of changes in the environment features. The information retrieval manager retrieves information relevant to the user device and network connectivity when the user logs onto the EPG. The information retrieval manager forwards information about the user's environmental features to the context-aware manager to use for filtering the final lists of recommended programmes to ensure such programmes match the user's device and network bandwidth and thus, filter out un-matched programmes. Moreover, any change or update to the user information done by the user themselves within the EPG is forwarded through the information retrieval manager.

The proposed EPG is designed to personalize the recommendations for each user, therefore, each user needs to have their own account (own EPG) within the EPG system. The designed EPG updates the user's viewing history automatically each time the user logs on to the EPG and watches a programme. This information is sent to the context-aware module through the information retrieval manager and is saved in the user's viewing history and updates all of the related information such as programme categories which will be used later when clustering the users.

4.6 Content Database Server

The content database server is responsible for storing the searched programmes that result from the different search operations. The content database server is connected to two parts in the overall EPG system namely, the content search module and the recommendation module. All the metadata of programmes which result from the search operations are stored in content database server. The programmes metadata is stored in the content database server using the unified format which was described previously in Table 4.1. Therefore, each user has a specified group of programmes in the content database server.

Furthermore, the content database server is connected to two blocks in the recommendation module namely, matching operation and availability block. The matching operation block is connected to the content database server to find those programmes that are most similar to the user's viewing history. The availability block is connected to the content database server to check whether the recommended lists of programmes created from the clustering groups and social networks are available in the content database server. The system is designed to remove the programmes from the content database server that have been stored and not accessed for one month. The same operation is performed on

items in user viewing history where items that have been watched one month ago are removed. These removal processes are important for managing potential data overload problems arising within the content database server as will be shown in the next chapter.

A complete system architecture for the whole EPG system is presented in figure 4.9.

4.7 Summary

In summary, the proposed architecture offers the following key:

1. The content search module is connected to multiple online content providers where the different metadata formats used by those providers are converted into a defined unified format that is used within the remainder of the system.
2. Social networks have been exploited in supporting the recommendation process where a list of programmes is compiled that is relevant to the user's profile and programmes watched or recommended by their friends.
3. The recommendation operation is based on multiple recommendation techniques which also take into consideration the user's viewing environment.
4. The system is centralized as a cloud service where all the operations have been implemented in a main server.
5. The flexibility of the system through a modular architecture, where the system has the ability to include any expansion in the number of online content providers or social networks without changing the main structure of the system.
6. Reducing the perceived time for creating the EPG due to implement the searching operations and recommendation processes as background tasks whilst the user logged off.
7. Providing a personalized EPG for online viewing by integrating the user's viewing history, their social network relationships and local environment features and context.
8. The EPG provides the user with option of searching for programmes that match a particular keyword.

The next chapter discusses the implementation of this EPG system and explains how the different functional blocks have been created for evaluation purposes.

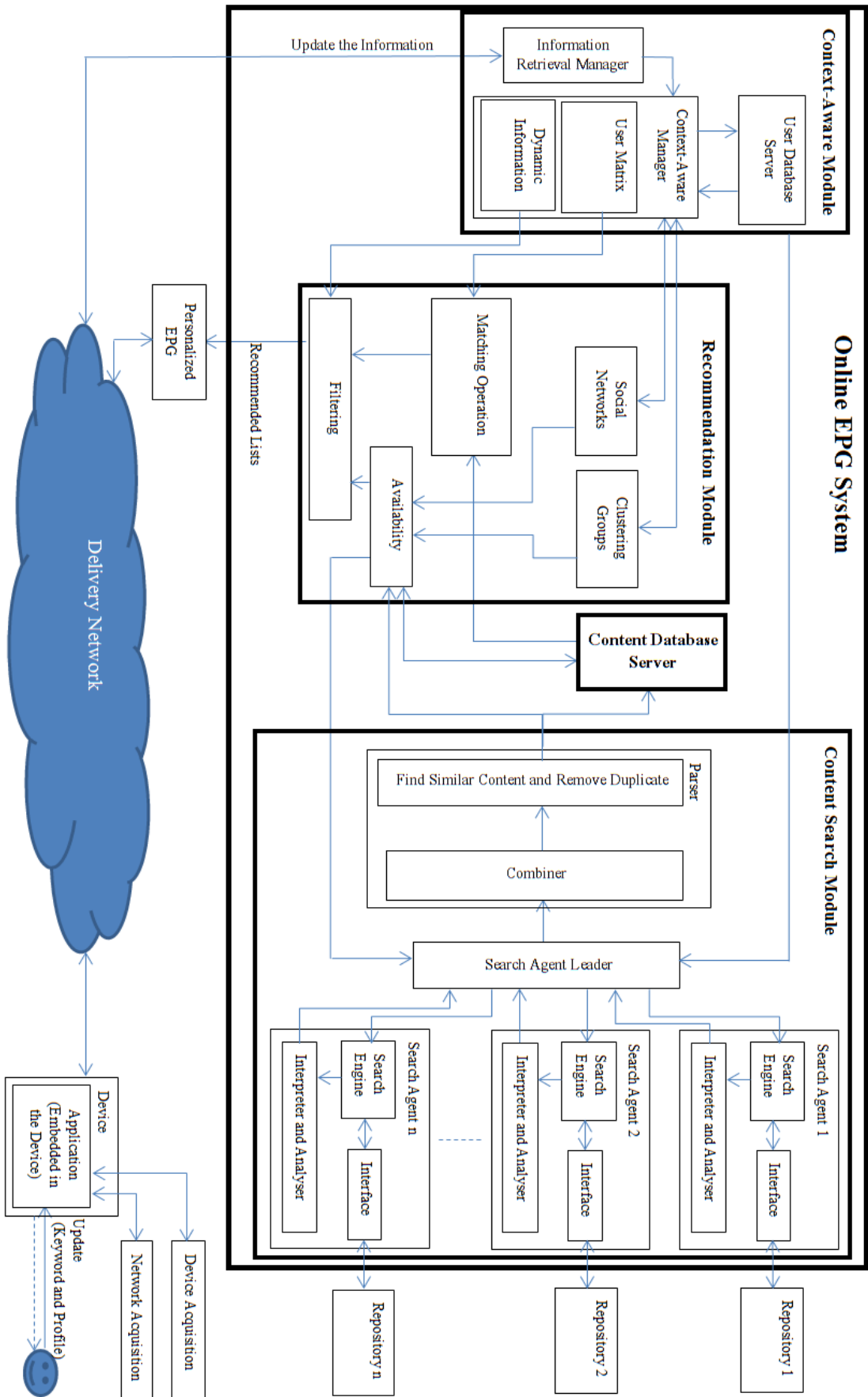


Fig. 4.9 Overall System Configuration

CHAPTER FIVE

IMPLEMENTATION AND FUNCTIONAL TESTING

Chapter Five

5.1 Introduction

This chapter presents the implementation of the EPG architecture described in the previous chapter and explains how the functionalities of the system blocks have been realised. The different parts of EPG system have all been implemented using a combination of Matlab and software written in PHP as described in the sections which follow.

5.2 Content Search Module

As previously shown in figure 4.2, the content search module handles multiple online video content sources. Implementation of this has focused on accessing BBC-iPlayer, YouTube, Dailymotion and The Movie Database as examples of online video content sources. Except for BBC-iPlayer which represents its metadata in XML format, the other sources provide their metadata using JSON format.

In order to communicate with these sources through their APIs, PHP code has been developed to retrieve the required programme metadata from each source and translate it into the standard used within the EPG and to save it in XML which is subsequently used by the other modules. This PHP code has been developed for each online content source and its function is to translate search keywords that are received from the search agent leader to an appropriate format which can be sent to the online content source through its associated API. Thereafter, the PHP code analyses and translates the retrieved programme metadata into XML and saves the results in a file which contains the required programme features that were discussed previously in Table 4.1. Figure 5.1 shows an example of raw metadata that was retrieved from The Movie Database through its API written with JSON while figure 5.2 shows the same metadata but after the analysing and translation operations that have been implemented using the PHP code (the search agent). As shown in figure 5.1, this piece of information contains many details including some pictures (those ended with extension .jpg) and numbers belong to the online content source. Therefore, this information should be treated in order to extract the required metadata. For example, the title (Happy Feet) is located in front of the bracket [original_title] (at the end of the seventh line) while there are two categories (Animation and Comedy) which are located in

front of the bracket [genres] (at the end of the forth line). It is clear that there are two arrays in that bracket, “0” which is “Animation” and “1” which is “Comedy”.

```
Array ( [adult] => [backdrop_path] => /xgLMY1VCwM8SJq6HB3ZMpYL6mic.jpg
[belongs_to_collection] => Array ( [id] => 92012 [name] => Happy Feet Collection
[poster_path] => /pYbF2FAi2MEVshl8jCDNTRNdTF0.jpg [backdrop_path] =>
/b9CLLxKsh63AHEkGE9nz9bupQqh.jpg ) [budget] => 100000000 [genres] => Array (
[0] => Array ( [id] => 16 [name] => Animation ) [1] => Array ( [id] => 35 [name] =>
Comedy ) ) [homepage] => http://www.warnerbros.com/happy-feet [id] => 9836
[imdb_id] => tt0366548 [original_language] => en [original_title] => Happy Feet
[overview] => Into the world of the Emperor Penguins, who find their soul mates through
song, a penguin is born who cannot sing. But he can tap dance something fierce!
[popularity] => 1.865194 [poster_path] => /8RLEzftZK87S2usLsQoYtyNYzoV.jpg
[production_companies] => Array ( [0] => Array ( [name] => Kennedy Miller
Productions [id] => 2537 ) [1] => Array ( [name] => Animal Logic [id] => 8089 ) )
[production_countries] => Array ( [0] => Array ( [iso_3166_1] => AU [name] =>
Australia ) ) [release_date] => 2006-11-16 [revenue] => 384335608 [runtime] => 108
[spoken_languages] => Array ( [0] => Array ( [iso_639_1] => en [name] => English ) )
[status] => Released [tagline] => WARNING: May Cause Toe-Tapping. [title] => Happy
Feet
```

Fig. 5.1 Retrieved Programme Raw Metadata From The Movie Database Source

```
<Metadata>
<Title>Happy Feet</Title>
<Description>Into the world of the Emperor Penguins, who find their soul mates
through song, a penguin is born who cannot sing. But he can tap dance something
fierce!</Description>
<Category>Animation, Comedy</Category>
<Language>en</Language>
<ProductionCountry>Australia</ProductionCountry>
<ProductionCompany>Kennedy Miller Productions, Animal Logic</ProductionCompany>
<Rating>1.865194</Rating>
<Date>2006-11-16</Date>
<Status>Released</Status>
<Address>http://www.warnerbros.com/happy-feet</Address>
```

Fig. 5.2 Retrieved Programme Metadata After Analysing and Translation Operation

Figure 5.3 shows a part of The Movie Database Interface PHP code for extracting the categories from the retrieved metadata. The first group of commands is used for sending the first request which includes the search keywords and decoding the response (which represented in JSON format) in order to read the Movie ID. The second group of commands is used to send the second request which includes the movie ID and decode the response to read other movie details. Note that for categories, a loop has been created for reading all the available categories within the movie metadata. Thereafter, the third group is for changing some of the category's names in order to match the unified format which is used within the EPG system.

```

- - -
/** Send the Request that include the Search Keyword and Decode the Response to get the Movie ID**
//-----
for ($i=1; $i <= $NofKeywords; $i++) {
    ${'NewKey' . $i} = preg_replace('/\s/', '+', rtrim($Keyword[$i-1]));
    $url = "https://api.themoviedb.org/3/search/movie?query=" . ${'NewKey' . $i} . "&api_key=e9a820";
    $data = json_decode(file_get_contents($url), true);
    $TotalResults = $data['total_results'];
    echo $TotalResults;
    $h=0;
foreach($data['results'] as $value) {
    $h++;
    ${'MovieID' . $h} = $value['id'];
}

/** Send the Second Request with Movie ID and Decode the Response to get the other details **
//-----
for ($j=1+(2*( $i-1)); $j <= (2*$i); $j++) {
    ${'TMdbDetails' . $j} = "https://api.themoviedb.org/3/movie/" . ${'MovieID' . ($j-(2*( $i-1)))};
    ${'txt' . $j} = file_get_contents("${'TMdbDetails' . $j}");
    ${'result' . $j} = json_decode(${'txt' . $j}, true);
    echo 'Title: ' . ${'result' . $j}['title'];
    $c=0;
    foreach(${'result' . $j}['genres'] as $value) {
        $c++;
        ${'TMovieCategory' . $c} = $value['name'];
    }

/** Change some Movie Categories to match the Categories of the Unified Format **
//-----
    if (preg_match("/Action/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Action/Adventure";
    }elseif(preg_match("/Crime/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Action/Adventure";
    }elseif(preg_match("/Fantasy/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Scientific Fiction";
    }elseif(preg_match("/TV Movie/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Movies";
    }elseif(preg_match("/War/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Documentary";
    }elseif(preg_match("/Adventure/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Action/Adventure";
    }elseif(preg_match("/Mystery/i", ${'TMovieCategory' . $c})) {
        ${'MovieCategory' . $c} = "Action/Adventure";
    }else {
        ${'MovieCategory' . $c} = ${'TMovieCategory' . $c};
    }
}
}

```

Fig. 5.3 Part of The Movie Database Interface PHP Code for Extracting the Categories

Each PHP code has to be different for each content provider according to the API commands and metadata used format. The code is divided into several parts where each part is responsible for a specific task such as get the authentication to communicate with the online source, send the request command, receive the response, analyse the retrieved information and translate the programmes metadata. Appendix I, II, III and IV contain the complete PHP code that has been developed for each selected online content source. Within the overall EPG system architecture, this PHP code represents the search agent which is the only part of the overall system which has to be variable because it needs to be created for each new online content source connected to the system. Note that in order to communicate with the online sources, an Apache web server has been installed and configured. Moreover, different methods have had to be used to establish the communication with these content sources through their API. For YouTube, an account in the YouTube developer section had to be created to obtain a developer key to use in requests for retrieving the programme metadata. Figure 5.4 shows an example of a part of PHP code which used to communicate with YouTube web site showing how this code includes where each part is responsible for specified task.

As shown in figure 5.4, the first group of commands is used to recall the required functions that needed to initiate the connection with YouTube server. These functions are located in the installed Apache web server location. The second group of commands represents recognizing the authentication key to ensure that this request is authorised by the developer section. Thereafter, the third group of commands initiates the request to YouTube by sending the keywords through a search command with specifying the number of required search results (number of videos that match the search keyword). Finally, the last group of commands is used to read the retrieved metadata and extract the required items.

On the other hand, the GET_IPLAYER software has had to be installed in order to communicate with the BBC-iPlayer web site and retrieve the required metadata from it. Figure 5.5 shows the start of the PHP code that used for that purpose. Again, the PHP code is divided into several parts to implement specified tasks. The first group of commands in this code represents creating the required requests that are used in GET-IPLAYER software to search for programmes relevant to a particular keyword. Then, after implementing the batch file that contains the requests, the second group of commands specifies the number of result's files where each programme is represented by separate


```

// **Implement the batch file to retrieve the XML files from BBC-iPlayer**
$MyFile = fopen("NofKeywords.txt", "r") or die("Unable to open file!");
$NofKeywords = fread($MyFile, filesize("NofKeywords.txt"));
fclose($MyFile);

$Keyword = file("Keywords.txt");
for ($i=1; $i <= $NofKeywords; $i++) {
    echo $Keyword[$i-1] . "<br />";
}

for ($j=1; $j <= $NofKeywords; $j++) {
    $BBCSearch = "get_iplayer " . "'" . trim($Keyword[$j-1]) . "'" . " --metadata-only --metada
ta=generic";
    echo $BBCSearch;
    $FileName = "BBC" . $j . ".bat";
    $SearchFile = fopen("$FileName", "w") or die("Unable to open file!");
    fwrite($SearchFile, "cd C:\Program Files\get_iplayer" . "\r\n");
    fwrite($SearchFile, $BBCSearch);
    fclose($SearchFile);
}

for($k=1; $k <= $NofKeywords; $k++) {
    exec('C:\xampp\htdocs\BBC' . $k . '.bat');
}

/**Specify the Number of Result's Files**
$FileNumber = (count(scandir('C:\Users\lza045\Desktop\iPlayer Recordings')) - 2);
$files1 = scandir('C:\Users\lza045\Desktop\iPlayer Recordings');

/**Read and Extract the Retrieved Metadata**
for($i=1; $i <= $FileNumber; $i++) {
    ${'file' . $i} = $files1[$i+1];
    ${'ADDRESS' . $i} = "C:/Users/lza045/Desktop/iPlayer Recordings/" . $files1[$i+1];
    $StringValue = (string){${'ADDRESS' . $i}};
    ${'xml' . $i}=simplexml_load_file(${'ADDRESS' . $i}) or die("Error: Cannot create object");

    ${'BBCTitle' . $i}=${'xml' . $i}->title;
    ${'BBCCategory' . $i}=${'xml' . $i}->category;
    ${'BBCDescription' . $i}=${'xml' . $i}->desc;
    ${'BBCWebAddress' . $i}=${'xml' . $i}->player;
    ${'BBCDuration' . $i}=${'xml' . $i}->duration;
    ${'BBCDate' . $i}=${'xml' . $i}->firstbcastdate;
    ${'BBCFormat' . $i}=${'xml' . $i}->modes;
    ${'BBCTitleKeyword' . $i}=${'xml' . $i}->nameshort;
}

```

Fig. 5.5 Part of BBC-iPlayer Interface Code

All of these search agent modules (PHP code) work simultaneously with the same search keywords being sent by the search agent leader to them to search the connected online sources for any relevant programmes. Then, after searching the connected online content sources and translating and saving the results in files as shown in Fig. 5.2, the search agent leader reads those files to extract the required information. The function of the search agent leader has been created in Matlab to read the files that are generated by each PHP search agent module and then create the programme matrix. Figure 5.2 shows the XML metadata that is output from the associated search agent and then figure 5.6 shows how this is presented after passing through the search agent leader.

| 3 |
|---|
| HAPPY FEET |
| INTO THE WORLD OF THE EMPEROR PENGUINS, WHO FIND THEIR SOUL MATES THROUGH SONG... |
| ANIMATION, COMEDY |
| Null |
| 1.865194 |
| HAPPY FEET |
| 2006-11-16 |
| Null |
| ENGLISH |
| AUSTRALIA |
| RELEASED |
| Null |
| Null |
| Null |
| Null |
| Null |
| Null |
| Null |
| KENNEDY MILLER PRODUCTIONS, ANIMAL LOGIC |
| Null |
| Null |
| Null |
| http://www.warnerbros.com/happy-feet |

Fig. 5.6 Programme Matrix

Note that the missing programme features are replaced with “Null” in the programme matrix. Each programme that results from the searching operation is represented by a matrix as in figure 5.6 regardless of which online source they come from. Thereafter, all of these programmes are forward to the parser (which is also implemented within Matlab). It consists of two parts: the combiner and parser. The combiner is used to combine all the programme matrices into one matrix where each programme is represented by a column in this matrix. This matrix is then sent to the parser to remove any duplicates by checking for and removing similar columns. The last step of the searching operation is to save the final matrix which contains all the programme metadata obtained from the search operation (which is called the content matrix) in the content database server. Figure 5.7 shows a part of Matlab code that responsible for combining the new content that resulted from a search operation and checking and removing the duplicates before saving the final results in the

content matrix. The full Matlab code for the combiner and parser are listed in Appendix ((VI-Implicit Function) Create VH Searching Content MAT).

```

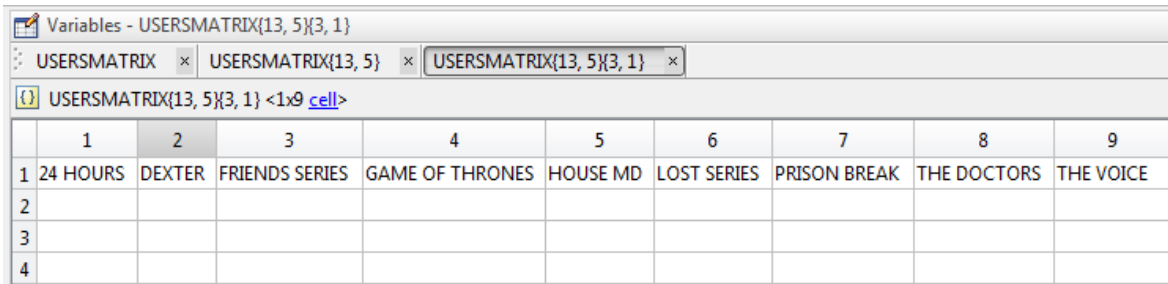
% Add the New Content to the User Content Matrix
%-----
if row == 1
    USERSCONTENTMAT{1,UserColumn}(:,1) = TmdbContent(:,h);
else
    USERSCONTENTMAT{1,UserColumn}(:,column+1) = TmdbContent(:,h);
end
save ('F:\Matlab training files\USERSCONTENTMAT.mat', 'USERSCONTENTMAT');
%Compare the added Content with the Exist Content to check the duplicates
%-----
[Newrow,Newcolumn]=size(USERSCONTENTMAT{1,UserColumn});
for r=1:Newcolumn-1;           %Find the Similar content from the Content Matrix
    for s=r+1:Newcolumn;
        ProgramSimilarity = strcmp(USERSCONTENTMAT{1,UserColumn}(:,r),USERSCONTENTMAT{1,UserColumn}(:,s));
        SimilarityResult = nnz(ProgramSimilarity);
        if SimilarityResult == 21           %this represent the similar programmes
            SimilarContent(r,s) = s;
        end
        if SimilarityResult == 22           %this represent the similar programmes
            SimilarContent(r,s) = s;
        end
        if SimilarityResult < 21           %this represent the non-similar programmes
            SimilarContent(r,s) = 0;
        end
    end
end
% Remove the Duplicates
%-----
for f=1:Newcolumn;           %Find the Column Number of the similar content
    FinalRemove(1,f) = nnz(SimilarContent(:,f));
end

Similar=find(FinalRemove>=1);
SimilarList=Similar';
USERSCONTENTMAT{1,UserColumn}(:,SimilarList)=[];
save ('F:\Matlab training files\USERSCONTENTMAT.mat', 'USERSCONTENTMAT');

```

Fig. 5.7 Part of Matlab Code Represents the Combiner and Parser

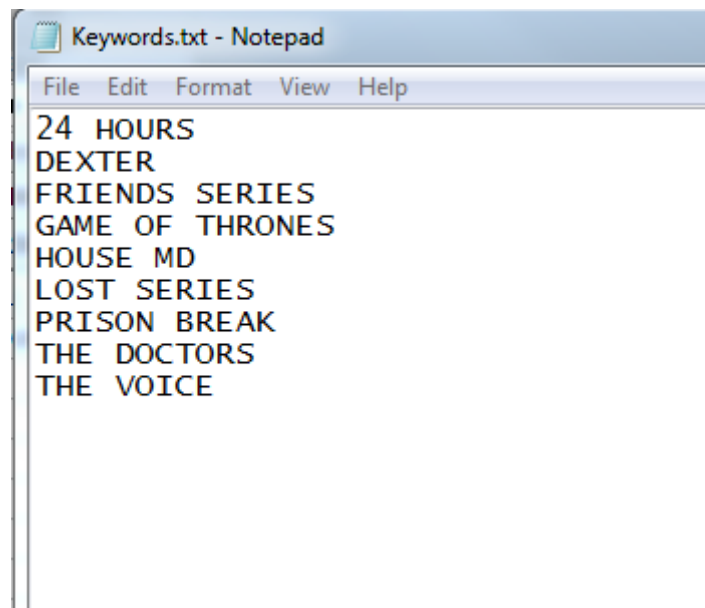
The search agent leader receives keywords from the EPG system and forwards them to each search agent as a text file. Figure 5.8 a and b show these search keywords at the input to and output from search agent leader step respectively.



The screenshot shows the MATLAB workspace with a variable named 'USERSMATRIX{13, 5}{3, 1}' containing a 1x9 cell array. The array contains the following search keywords: 24 HOURS, DEXTER, FRIENDS SERIES, GAME OF THRONES, HOUSE MD, LOST SERIES, PRISON BREAK, THE DOCTORS, and THE VOICE.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|----------|--------|----------------|-----------------|----------|-------------|--------------|-------------|-----------|
| 1 | 24 HOURS | DEXTER | FRIENDS SERIES | GAME OF THRONES | HOUSE MD | LOST SERIES | PRISON BREAK | THE DOCTORS | THE VOICE |
| 2 | | | | | | | | | |
| 3 | | | | | | | | | |
| 4 | | | | | | | | | |

- a. Search Keywords Which Saved in User Matrix (Context-aware Module)



- b. Search Keywords After Passing Through Search Agent Leader

Fig. 5.8 Search Keywords before and after Search Agent Leader

Figure 5.9 shows how search agent leader code reads and forwards the search keywords within Matlab (as matrix) and then these keywords are written in a text file and saved to be used by all of the search agents modules simultaneously. Full Matlab code is listed in Appendix ((VI-Implicit) VH Searching) for the whole process of searching.

```
%Receive the Search Keywords From the EPG System
%-----

load ('F:\Matlab training files\USERSMATRIX.mat');
SearchKeyword=USERSMATRIX{13,UserColumn}{3,1}
NofKeywords=numel(SearchKeyword)

PHPFile = fopen('C:\xampp\htdocs\NofKeywords.txt', 'w');
fprintf(PHPFile, '%d', ExactNofKeywords);
fclose(PHPFile);

%Write the same Keywords in Text File and Forward them to the Search Agents
%-----

if strcmp(SearchKeyword{1,1}, 'Null')==0
Yfile = fopen('C:\xampp\htdocs\Keywords.txt', 'w');
for i=1:NofKeywords
    if strcmp(SearchKeyword{1,i}, 'Null')==1
        end
    if strcmp(SearchKeyword{1,i}, 'Null')==0
        fprintf(Yfile, '%s\r\n', SearchKeyword{1,i});
    end
end
fclose(Yfile);
```

Fig. 5.9 Part of Search Agent Leader Code to Read and Write the Search Keywords in Text File

5.3 Recommendation Module

As described in the previous chapter, multiple recommendation schemes have been employed in the EPG to generate a more precise and appropriate set of recommendations for the users. Therefore, a hybrid recommending method comprising content-based, collaborative filtering and context-aware techniques has been used to overcome the drawbacks of using just one recommendation process. Moreover, social networks have been exploited to support the recommendation operation with information related to the preferences of a user's friends. This module within the EPG has been developed using PHP to communicate with social networks through their API's and Matlab codes to create the analysing and recommending blocks. The three lists that are created by this module are discussed in detail in the following subsections.

5.3.1 Viewing History Based (Content-Based) Recommendation

A user's viewing history is used as the basis of a new search operation to find additional programmes that are similar to those that have already been watched by the user. The matching block (as shown in figure 4.5) has been developed using Matlab where a similarity calculation is done between each programme in the searched list and each programme in the viewing history. Here, in order to calculate the similarity, it is important to mention that the similarity is calculated between cells in each matrix that contains both words and numbers. Therefore, a simple similarity measure has been implemented by calculating the intersection divided by the union of the cells' content. This measure is called the Jaccard coefficient measure. However, the similarity has been calculated between two columns (searched programme column from the content matrix and viewing history programme column from user matrix) as shown in equations 5.1 and 5.2:

$$Sim(A, B) = \sum_{j \in I} SL(A_i, B_j) \quad \dots\dots\dots \text{Equation 5.1}$$

$$SL(A, B) = (\sum_{i \in K} (a_i, A) \cap \sum_{j \in L} (b_j, B)) / (\sum_{i \in K} (a_i, A) \cup \sum_{j \in L} (b_j, B)) \quad \dots\dots \text{Equation 5.2}$$

Where A is the viewing history programme features matrix which contains i features, B is the searched programme features matrix which contains j features, Sim(A, B) is the similarity between them and SL is the similarity level between each corresponding feature in both matrices. The keywords in each feature are represented by a and b where A has K keywords a, and B has L keywords b. However, because the intersection and union between each feature results in a group of words, SL will be the number of words which result from the intersection divided by the number of words which result from the union. SL range is bounded between [0, 1]. After calculating the similarity between each corresponding feature in both matrices SL, the similarity value between the matrices, is calculated by finding the summation of all SL's of the different features. The purpose of this operation is to recommend the programmes that have highly similar features with the user's viewing history. Based on experimentation, a threshold has been set for this at 0.7.

Figure 5.10 shows a part of recommendation process Matlab code represents the similarity calculations based on Jaccard Coefficient measure.

```

for Feature=1:21;
    if strcmp(USERSMATRIX{3,UserColumn}{1, HOUR}{Feature, hourProgram},'Null')==0 %this will give you matrix of comparison the
        FirstTerm = regexp(USERSMATRIX{3,UserColumn}{1, HOUR}{Feature, hourProgram}, '\s*:\s*', 'split');
        ItemWords = numel(FirstTerm);
        SecondTerm = regexp(USERSCONTENTMAT{1,UserColumn}{Feature, Program}, '\s*:\s*', 'split');
        SimilarWords = intersect(SecondTerm, FirstTerm);
        Words = numel(SimilarWords);
        ratio = Words/ItemWords;
        if ratio>0.7
            result2(Feature,1)=1;
        else
            result2(Feature,1)=0;
        end
    end
    if strcmp(USERSMATRIX{3,UserColumn}{1, HOUR}{Feature, hourProgram},'Null')==1 % we put Null to mention that no information
        result2(Feature,1)=1;
    end
end
else
    for Feature=1:21;
        result2(Feature,1)=0;
    end
end
result1 = [result2];
Similarity1(Program, hourProgram) = nnz(result1) %Find the programmes that have non-zero similarity value

```

Fig. 5.10 Part of Recommendation Process Matlab Code Represents the Similarity Calculations Based on Jaccard Coefficient Measure

The figure shows how the system calculates the similarity between the searched programmes and the programmes listed for each hour within the viewing history. The similarity calculated for each programme feature (the 21 features) and at the end of this operation the programmes that have similarity above the threshold will be selected and saved in the recommended list for that particular hour. The resulting list of programmes is ranked according to the similarity value where the greater similarity value produces a higher priority in the displayed list. Figure 5.11 shows a part of user matrix which represents the recommended list that has been created for each hour based on what the user has previously watched in that hour. Each cell which is shown actually contains a matrix representing the related programme features.

For the hours that don't contain any watched programmes in the viewing history, the EPG system will recommend some programmes at that hour from the recommended programmes that stored in the user content matrix in the content database server. For test purposes, the number of random programmes has been limited to 8. Thereafter, the recommended list of programmes is saved in the user matrix within the context-aware module where it is ready to be displayed to the user when they log onto the EPG. The operations described in this section are repeated every day to reflect updates in the user's recent viewing habits.

5.3.2 Users Clusters

The second recommendation technique that is used in the EPG system is collaborative filtering. As mentioned previously, 36 programme categories have been used to classify the programmes into different groups. Matlab code has been developed to cluster users into these groups based on the categories of the programmes which they have already watched. A K-means algorithm is employed for this purpose using a function that is available within Matlab. The system reads the categories of the watched programmes from the user matrix for each user and then determines the number of clusters. This number could be specified according to any criteria that the developed selects. For test purpose, equation. 5.3 has been used for determining the number of clusters based on the number of users. Note that, this equation is not a standard equation, but it is created according to the needs of our test. However, number of clusters could be specified based on any another criteria such as number of categories.

$$\text{Number of Clusters} = \text{Integer} \left(\frac{\text{Number of Users}}{3} \right) \quad \dots\dots\dots \text{Equation 5.3}$$

Thereafter, the system applies the k-means algorithm and groups users into clusters. Each cluster contains the users who have a close viewing pattern (they have common watched programmes' categories). Then the system reads the viewing history of those users from the user matrix and creates a group viewing history. Figure 5.12 shows a part of Matlab code that used for user clustering. Here, at the beginning of this code, the number of clusters (groups) is specified based on the equation 5.3. Thereafter, the categories matrix is

created and K-means algorithm is applied. Other tasks are implemented as shown in the figure such as determining the number of users in each group and creating the viewing history of each group. The full Matlab code is listed in Appendix ((VI-Implicit Collaborative Filtering)).

```

%Specifying the number of groups based on the number of users in the system
groups = round(n/3)

% Create the Categories Matrix to apply the K-means algorithm
for i=1:n;
    for j=1:o;
        TestTerm = regexp(USERSMATRIX{13,i}{2,1}', '\.', 'split');
        [rTerm, cTerm] = size(USERSMATRIX{13,i}{2,1}');
        CFMatrix(i,j)=0;
        for k=1:rTerm;
            if (nnz(strcmp(TestTerm{k,1}, CategoriesMatrix{j,1}))>0)
                CFMatrix(i,j)=1;
            end
        end
    end
end

Groups = kmeans(CFMatrix, groups)

%Determine how many user in each group and specify each user to retrieve
%the necessary information from his profile
for loop1=1:groups;
    UserGroups(loop1,1)=find(Groups==loop1)
    NoUserGroup(loop1,1)=numel(UserGroups(loop1,1))
end

%Create the viewing history of each group
for loop2=1:groups;
    GroupViewingHistory{1,loop2}={};
    for counter2=1:NoUserGroup(loop2,1);
        if (strcmp((USERSMATRIX{13,UserGroups{loop2,1}(counter2,1)}{1,1}), 'Null')==0)
            GroupViewingHistory{1,loop2} = union(GroupViewingHistory{1,loop2}, USERSMATRIX{13,UserGroups{loop2,1}(counter2,1)}{1,:});
        end
    end
    [GVHrow, GVHcolumn] = size(GroupViewingHistory{1,loop2});
    if GVHcolumn>GVHrow
        GroupViewingHistory{1,loop2} = GroupViewingHistory{1,loop2}';
    end
end
end

```

Fig. 5.12 Part of Matlab Code That Used for User Clustering

Subsequently, the user within each cluster will be provided with a list of programmes that have been watched by the other users in the same cluster but not yet watched by the specific user. Moreover, the analyzer will determine the popular programmes through determining the number of users who have watched the programmes in each cluster. This step is used to rank the programmes within the final EPG where the programmes that have more viewers will have a higher ranking. The recommended list of programmes is then

checked against the content database server to retrieve the details of those programmes. If there are any missing programmes, those programme titles are forwarded to the search agent leader as shown in Fig. 5.8.a to search for them in the online content sources. Thereafter, the programmes are added to the content matrix of the specific user. The list of recommended programmes for each user is saved in the user matrix which is displayed to the user when they log on to the EPG. This operation is implemented every day. Figure 5.13 shows a part of the user matrix that represents the recommended list of programmes based on the user clustering method.

| USERSMATRIX | |
|--------------------------------|--|
| USERSMATRIX{15, 7} <23x2 cell> | |
| 1 | |
| 1 | THE VOICE (Watched By (2) Subscribers) |
| 2 | TOP 10 MOST AMAZING AUDITIONS THE VOICE US OF ALL TIME - THE BEST OF THE VOICE US (Watched By (2) Subscribers) |
| 3 | 2016 NATIONAL GEOGRAPHIC TRAVEL PHOTOGRAPHER OF THE YEAR WINNERS (Watched By (1) Subscribers) |
| 4 | ADEL IMAM - SHAHED MA SHAFSH HAGA - Ø²Ø§Ø¬Ù„ Ø¥Ù…Ø§Ù… - Ø´Ø§Ù±Ø¬Ù…Ø§ Ø´Ù Ø´Ø-Ø§Ø¬-Ø© (Watched By (1) S |
| 5 | BARCELONA VS REAL MADRID 1-2 FULL HD 1080P Å— 02/04/2016 (Watched By (1) Subscribers) |
| 6 | BBC NEWSLINE (Watched By (1) Subscribers) |
| 7 | BEST OF JOEY IN FRIENDS SEASON 5.WMV (Watched By (1) Subscribers) |
| 8 | CAST AWAY (Watched By (1) Subscribers) |
| 9 | CAST AWAY TRAILER (Watched By (1) Subscribers) |
| 10 | HORROR (Watched By (1) Subscribers) |
| 11 | LIONEL MESSI - ALL 21 GOALS VS REAL MADRID (HD) (Watched By (1) Subscribers) |
| 12 | NATIONAL GEOGRAPHIC WILD THAILAND - NAT GEO WILD HD (NATURE DOCUMENTARY) (Watched By (1) Subscribers) |
| 13 | NATIONAL GEOGRAPHIC'S DOLPHINS: THE WILD SIDE (Watched By (1) Subscribers) |
| 14 | NATIONAL GEOGRAPHIC: EMPERORS OF THE ICE (Watched By (1) Subscribers) |
| 15 | REAL IELTS SPEAKING BAND 7 (Watched By (1) Subscribers) |
| 16 | RIO 2016 TOP 10 FAILS COMPILATION (SUMMER OLYMPICS) (Watched By (1) Subscribers) |
| 17 | RONALDINHO VS REAL MADRID 2005/2006 (AWAY) (Watched By (1) Subscribers) |
| 18 | THE LORD OF THE RINGS: THE RETURN OF THE KING (Watched By (1) Subscribers) |
| 19 | TOM AND JERRY (Watched By (1) Subscribers) |
| 20 | TOM AND JERRY - EPISODE 60 - SLICKED-UP PUP (1951) (Watched By (1) Subscribers) |

Fig. 5.13 Part of User Matrix which Represents the Recommended List of Programmes based on the Users Clusters Method

5.3.3 Social Networks Contribution

Social networks are exploited by retrieving information from the user's account and those of their friends once each has agreed to join the application located within the social network. Facebook has been used in the EPG as an example social network for evaluation purpose. An application has been developed in Facebook using PHP and Facebook SDK in order to retrieve the required information from each user's account and their friends' accounts as shown in figure 5.14. Facebook SDK for PHP is a library with features that enable PHP developers to integrate Facebook login in to their application and to communicate with the Graph API. The full application code is included in Appendix V.

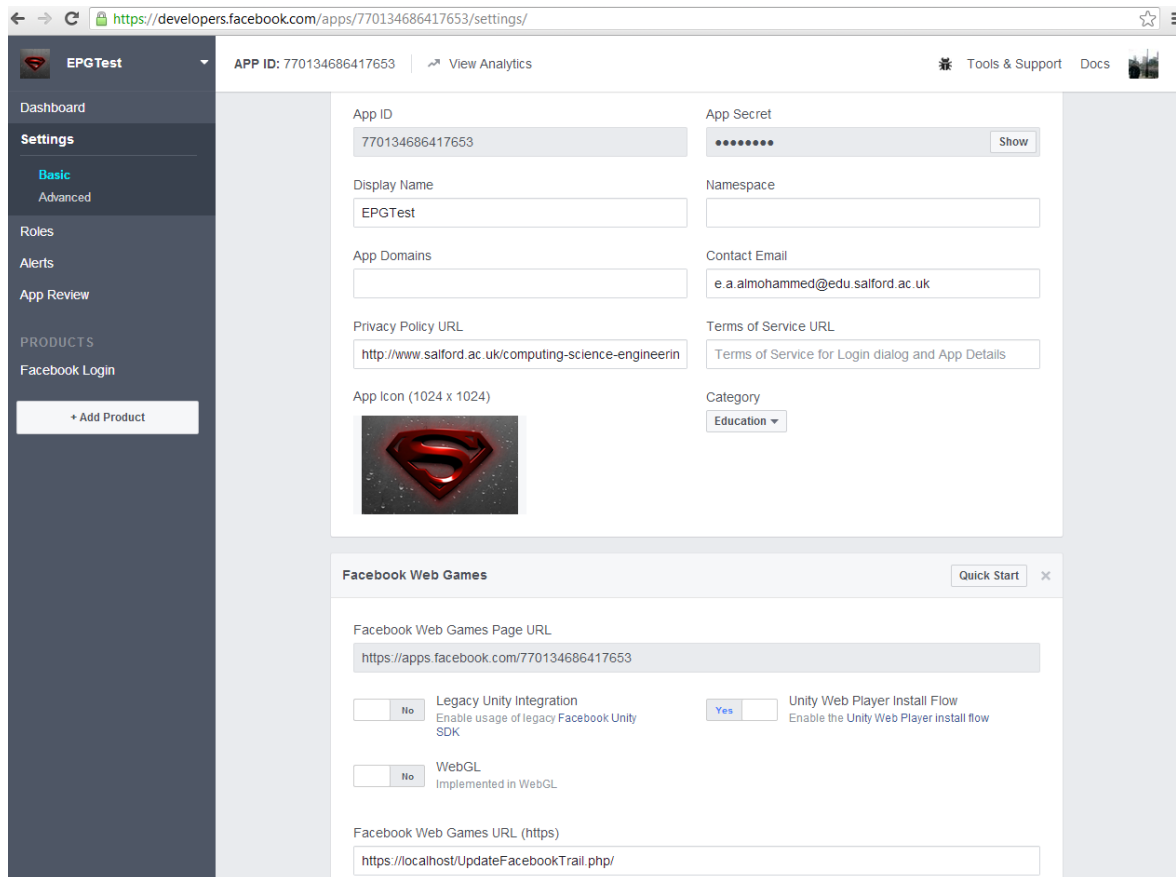


Fig. 5.14 Facebook Application

The retrieved information includes the list of friends who are also members in the EPG system and information about the ‘liked’, ‘watched’ and ‘want to watch’ movies and TV shows which have selected by the users and their friends. Figure 5.15 shows a part of Facebook application that used to retrieve the information about the ‘liked’, ‘watched’ and ‘want to watch’ movies and TV shows for each user.

```

/**Retrieve Watched Videos**
//-----
    $getvideo1 = (new FacebookRequest(
    $session,
    'GET',
    '/' . $userID . '/video.watches'
    ))->execute()->getGraphObject()->asArray();

    $WatchedNumber = count($getvideo1['data']);
    $i = 0;
    foreach ($getvideo1['data'] as $key) {
    $i++;
    ${'WatchedCategory' . $i} = $key->application->name;
    foreach ($key->data as $key1) {
    ${'WatchedTitle' . $i} = $key1->title;
    }
    }

/**Retrieve Want to Watch Videos**
//-----
    $getvideo2 = (new FacebookRequest(
    $session,
    'GET',
    '/' . $userID . '/video.wants_to_watch'
    ))->execute()->getGraphObject()->asArray();
    $WantToWatchNumber = count($getvideo2['data']);
    $a = 0;
    foreach ($getvideo2['data'] as $key) {
    $a++;
    ${'WantedCategory' . $a} = $key->application->name;
    foreach ($key->data as $key1) {
    ${'WantedTitle' . $a} = $key1->title;
    }
    }

/**Retrieve Liked Videos**
//-----
    $getvideo3 = (new FacebookRequest(
    $session,
    'GET',
    '/' . $userID . '/likes'
    ))->execute()->getGraphObject()->asArray();
    $LikesNumber = count($getvideo3['data']);

    $b = 0;
    foreach ($getvideo3['data'] as $key) {
    $b++;
    ${'LikesCategory' . $b} = $key->category;
    ${'LikesTitle' . $b} = $key->name;
    }
    }

```

Fig. 5.15 Part of Facebook Application that used to retrieve the Information about the ‘liked’, ‘watched’ and ‘want to watch’ Movies and TV Shows for each user

Facebook provides this information through its API in JSON format and part of this shown in figure 5.16 which represents the watched videos by a user.

```
Array ( [data] => Array ( [0] => stdClass Object ( [end_time] => 2015-05-22T10:34:42+0000 [start_time] => 2015-05-22T10:34:42+0000
[type] => video.watches [application] => stdClass Object ( [link] => https://www.facebook.com/games/?app_id=475761232465217 [name]
=> Movies [id] => 475761232465217 ) [comments] => stdClass Object ( [data] => Array ( ) [can_comment] => 1 [comment_order] =>
chronological [count] => 0 ) [from] => stdClass Object ( [name] => Emad Jasim [id] => 10205779120271322 ) [likes] => stdClass Object
( [data] => Array ( ) [can_like] => 1 [count] => 0 [user_likes] => ) [no_feed_story] => [publish_time] => 2015-05-22T10:34:42+0000
[data] => stdClass Object ( [movie] => stdClass Object ( [id] => 126064364200766 [title] => The Expendables 2 [type] => video.movie
[url] => https://www.facebook.com/TheExpendables2Movie/ ) ) [id] => 10205843156432186 ) [1] => stdClass Object ( [end_time] =>
2015-05-15T18:33:40+0000 [start_time] => 2015-05-15T18:33:40+0000 [type] => video.watches [application] => stdClass Object ( [link]
=> https://www.facebook.com/games/?app_id=297380653703981 [name] => TV Shows [id] => 297380653703981 ) [comments] =>
stdClass Object ( [data] => Array ( ) [can_comment] => 1 [comment_order] => chronological [count] => 0 ) [from] => stdClass Object
( [name] => Emad Jasim [id] => 10205779120271322 ) [likes] => stdClass Object ( [data] => Array ( ) [can_like] => 1 [count] => 0
[user_likes] => ) [no_feed_story] => [publish_time] => 2015-05-15T18:33:40+0000 [data] => stdClass Object ( [tv_show] => stdClass
Object ( [id] => 22577904575 [title] => FRIENDS (TV Show) [type] => video.tv_show [url] => https://www.facebook.com/friends.tv/ ) )
[id] => 10205800812413612 ) [2] => stdClass Object ( [end_time] => 2015-05-15T18:33:33+0000 [start_time] => 2015-05-
15T18:33:33+0000 [type] => video.watches [application] => stdClass Object ( [link] => https://www.facebook.com/games/?
app_id=297380653703981 [name] => TV Shows [id] => 297380653703981 ) [comments] => stdClass Object ( [data] => Array ( )
[can_comment] => 1 [comment_order] => chronological [count] => 0 ) [from] => stdClass Object ( [name] => Emad Jasim [id] =>
10205779120271322 ) [likes] => stdClass Object ( [data] => Array ( ) [can_like] => 1 [count] => 0 [user_likes] => ) [no_feed_story] =>
[publish_time] => 2015-05-15T18:33:33+0000 [data] => stdClass Object ( [tv_show] => stdClass Object ( [id] => 131583793581459 [title]
=> Tom and Jerry [type] => video.tv_show [url] => https://www.facebook.com/TomandJerry/ ) ) [id] => 10205800811973601 ) [3] =>
stdClass Object ( [end_time] => 2015-05-15T18:33:25+0000 [start_time] => 2015-05-15T18:33:25+0000 [type] => video.watches
[application] => stdClass Object ( [link] => https://www.facebook.com/games/?app_id=297380653703981 [name] => TV Shows [id] =>
297380653703981 ) [comments] => stdClass Object ( [data] => Array ( ) [can_comment] => 1 [comment_order] => chronological [count]
=> 0 ) [from] => stdClass Object ( [name] => Emad Jasim [id] => 10205779120271322 ) [likes] => stdClass Object ( [data] => Array ( )
[can_like] => 1 [count] => 0 [user_likes] => ) [no_feed_story] => [publish_time] => 2015-05-15T18:33:25+0000 [data] => stdClass Object
( [tv_show] => stdClass Object ( [id] => 110475388978628 [title] => The Walking Dead [type] => video.tv_show [url] =>
https://www.facebook.com/TheWalkingDeadAMC/ ) ) [id] => 10205800811613592 ) [4] => stdClass Object ( [end_time] => 2015-05-
15T17:33:02+0000 [start_time] => 2015-05-15T17:33:02+0000 [type] => video.watches [application] => stdClass Object ( [link] =>
https://www.facebook.com/games/?app_id=475761232465217 [name] => Movies [id] => 475761232465217 ) [comments] => stdClass
Object ( [data] => Array ( ) [can_comment] => 1 [comment_order] => chronological [count] => 0 ) [from] => stdClass Object ( [name] =>
Emad Jasim [id] => 10205779120271322 ) [likes] => stdClass Object ( [data] => Array ( ) [can_like] => 1 [count] => 0 [user_likes] => )
[no_feed_story] => [publish_time] => 2015-05-15T17:33:02+0000 [data] => stdClass Object ( [movie] => stdClass Object ( [id] =>
212216417436 [title] => The Lord of the Rings Trilogy [type] => video.movie [url] => https://www.facebook.com/lordoftheringstrilogy/ ) )
[id] => 10205800443724395 ) [5] => stdClass Object ( [end_time] => 2015-05-08T11:26:17+0000 [start_time] => 2015-05-
08T11:26:17+0000 [type] => video.watches [application] => stdClass Object ( [link] => https://www.facebook.com/games/?
app_id=475761232465217 [name] => Movies [id] => 475761232465217 ) [comments] => stdClass Object ( [data] => Array ( )
[can_comment] => 1 [comment_order] => chronological [count] => 0 ) [from] => stdClass Object ( [name] => Emad Jasim [id] =>
10205779120271322 ) [likes] => stdClass Object ( [data] => Array ( ) [can_like] => 1 [count] => 0 [user_likes] => ) [no_feed_story] =>
[publish_time] => 2015-05-08T11:26:17+0000 [data] => stdClass Object ( [movie] => stdClass Object ( [id] => 12393266550 [title] =>
The Lion King [type] => video.movie [url] => https://www.facebook.com/DisneyTheLionKing/ ) ) [id] => 10205763320516338 ) )
[paging] => stdClass Object ( [cursors] => stdClass Object ( [before] => MTAyMDU0NDMxNTY0MzIxODYzZD [after] =>
MTAyMDU0NDUzNmZmMjA1MTYzZmZkZD ) ) )
```

Figure 5.16 Data Retrieved from Facebook (in JSON) about Watched Movies and TV Shows for Particular User

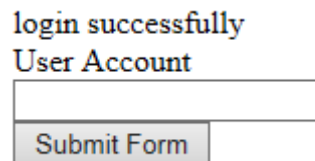
This information includes the titles and categories of 6 movies and TV shows that this particular user selected them as ‘watched’ where the ‘Array’ (which represents the retrieved information) started from ‘[0]’ in the first line of the figure and ended with ‘[5]’ in the ninth line from the bottom. This data is analyzed and translated into XML as shown in figure 5.17 which shows the same data in figure 5.16 but after extracting the needed information. For example, the first watched programme (with number ‘[0]’) has been

defined as 'Movies' (the category) at the end of the second line '[name]=Movies' while the title of the movie has been written at the end of sixth line '[title]=Expendables 2'.

```
<WatchedNumber>6</WatchedNumber>
<WatchedVideos>
<WatchedTitle1>The Expendables 2</WatchedTitle1>
<WatchedCategory1>Movies</WatchedCategory1>
<WatchedTitle2>FRIENDS (TV Show)</WatchedTitle2>
<WatchedCategory2>TV Shows</WatchedCategory2>
<WatchedTitle3>Tom and Jerry</WatchedTitle3>
<WatchedCategory3>TV Shows</WatchedCategory3>
<WatchedTitle4>The Walking Dead</WatchedTitle4>
<WatchedCategory4>TV Shows</WatchedCategory4>
<WatchedTitle5>The Lord of the Rings Trilogy</WatchedTitle5>
<WatchedCategory5>Movies</WatchedCategory5>
<WatchedTitle6>The Lion King</WatchedTitle6>
<WatchedCategory6>Movies</WatchedCategory6>
</WatchedVideos>
```

Fig. 5.17 Information retrieved from Facebook after Translation Operation

All of the information that is related to the user's preferences is gathered in a single file as shown in Fig. 5.18. The name of this file is the same name that is used when the user logs in to the EPG system. Moreover, another file is created which contains the users' friends and their ID's but only the friends who are also members of the EPG system and who have joined the Facebook EPG application.



login successfully
User Account

Figure 5.18 Joining the Application in Facebook

Thereafter, the Interpreter, which is also implemented in Matlab code, reads these files and creates a matrix containing the user preferences and friends' list where this information is

saved in the context-aware module. Subsequently, the system reads information about the preferences and recommendation of the user's friends which is located in the user matrix of each one of those friends. Then, the system analyzes this information to find the commonalities and differences between the user and his friends and the common programmes between them. Common programmes are given higher rank and priority. All of the programme titles are collected into one list and checked through the availability block against the content database server. The programmes that exist in the content database server will be labeled with a specific flag while the programmes that do not exist have to be forwarded to the search agent leader to search the online content sources for relevant programmes. If some of the programmes are not found in the online content sources during the search operation, those programmes will be searched for on the next day. At the end of this operation, a list of programmes will be created and ranked based on the programme's popularity between friends. This is then saved in the user matrix of that particular user. The Matlab code which implements all these tasks is listed in Appendices ((VI-Implicit Function) Facebook Info) and ((VI-Implicit Function) Facebook Matrix). Figure 5.19 shows part of the user matrix that includes the recommended list of programmes after the analyzing and ranking operation.

All of the programmes in the recommended list are linked to their details in the content database server. By selecting any programme, the system will retrieve the programme's details from the content database server based on an associated pointer in the user matrix. All the retrieving and analyzing operations that relevant with integration the social networks in recommendation process is performed when the users logs off from the EPG. When the user logs on, the recommended list of programmes will be displayed for them.

5.3.4 Filtering Operation

The final operation that is implemented on the recommended lists of programmes before displaying them to the user is filtering according to the user's viewing environment. This operation ensures that recommended programmes match the time of day, the user's device type and the available network bandwidth. As mentioned previously, this information is provided to the system through an application installed on the user device. However, for the prototype system this information is provided via a text file which the EPG system reads when the user logs onto the EPG.

| | 1 |
|----|--|
| 1 | 'Frozen (Recommended by2 Friends)' |
| 2 | 'Gone in 60 Seconds (Recommended by2 Friends)' |
| 3 | 'The Walking Dead (Recommended by2 Friends)' |
| 4 | 'Twilight Saga: Eclipse (Recommended by2 Friends)' |
| 5 | 'Underworld (Recommended by2 Friends)' |
| 6 | '13 Real Asian Horror Stories (Recommended by1 Friends)' |
| 7 | 'Alvin and the Chipmunks: Chipwrecked (Recommended by1 Friends)' |
| 8 | 'Avatar (Recommended by1 Friends)' |
| 9 | 'Bad Girls Club (season 14) (Recommended by1 Friends)' |
| 10 | 'Batman: The Dark Knight (Recommended by1 Friends)' |
| 11 | 'Bee Movie (Recommended by1 Friends)' |
| 12 | 'Beverly Hills Cop II (Recommended by1 Friends)' |
| 13 | 'Big Hero 6 (Recommended by1 Friends)' |
| 14 | 'Bourne Ultimatum (film) (Recommended by1 Friends)' |
| 15 | 'Brave (Recommended by1 Friends)' |
| 16 | 'Braveheart (Recommended by1 Friends)' |
| 17 | 'Charlie and the Chocolate Factory (Recommended by1 Friends)' |
| 18 | 'Crank (Recommended by1 Friends)' |
| 19 | 'Die Hard 4 (Recommended by1 Friends)' |
| 20 | 'Disney Alice (Recommended by1 Friends)' |

Fig. 5.19 Facebook Recommended List of Programmes after Analysing and Ranking Operation

Specifying the time of day has been employed to ensure that the recommended list of programmes is available for that particular hour. For test purposes, in order to adapt to device type and network bandwidth, tables have been created to determine the resolution of the recommended programmes based on the device type and available bandwidth. The resolution of programmes has been classified into five types as shown in Table 5.1 (Barb, 2012).

| No. | Term | Definition |
|------------|-------------|--|
| 1 | Flash | Refers to Low Definition Videos |
| 2 | STD | Refers to Standard Definition Videos (480p) |
| 3 | HD | Refers to High Definition Videos (720p) |
| 4 | HDX | Refers to High Definition Videos (1080p) |
| 5 | 3D | Refers to 3 Dimensional High Definition Videos |

Table 5.1 Types of Videos Resolution that employed in the Prototype

Those types have been exploited in filtering the recommended programmes according to the device type as shown in Table 5.2 and available bandwidth as shown in Table 5.3. The recommendation module reads the recommended lists of programmes from the user matrix and the environmental features from the context-aware module and then executes a matching operation between the available environmental features and the resolution cell in the programme features column (see figure 5.6).

| No. | Device Type | Terms Included |
|------------|--------------------|-------------------------|
| 1 | Smartphone | Flash, STD, HD |
| 2 | Tablet | Flash, STD, HD |
| 3 | Laptop and PC | Flash, STD, HD, HDX |
| 4 | SD Smart TV | Flash, STD |
| 5 | HD Smart TV | Flash, STD, HD |
| 6 | Full HD Smart TV | Flash, STD, HD, HDX |
| 7 | 3D Smart TV | Flash, STD, HD, HDX, 3D |

Table 5.2 Device Type versus Resolution Table

| No. | Available Network Bandwidth (Mbps) | Terms Included |
|-----|------------------------------------|-------------------------|
| 1 | more than 0.3 | Flash |
| 2 | 1-2 | Flash, STD |
| 3 | 2-4 | Flash, STD, HD |
| 4 | 4-9 | Flash, STD, HD, HDX |
| 5 | Greater than 9 | Flash, STD, HD, HDX, 3D |

Table 5.3 Available Network Bandwidth versus Resolution Table

Based on this, the system recommends the programmes that match the device type and available network capacity and filters out the others. The matching operation implemented in Matlab directly by using a string comparison method between two cells in both matrices. Figure 5.20 shows the part of Matlab code that is used for filtering the recommended programmes based on available network bandwidth. Note that the first part of this code that showed in figure 5.20 is used to recognizing the programmes that match the network bandwidth while the second part is used to remove the non-match programmes from the list. For more details related with the full Matlab code that is used for the overall filtering operation see Appendix ((VI) User Interface).

Then, the system forwards the final list to the designed online EPG to be displayed to the users. This operation is the only operation that is implemented when the user logs onto the EPG and hence, decrease the time required to display the recommended programmes to the users.

Note that, adaptive bitrate streaming techniques are becoming more common to adapt to changing network conditions and provide high quality playback with fewer re-buffering stages. An example is MPEG-DASH where the multimedia file is partitioned into one or more segments and delivered to the client over the internet in real-time using HTTP where each segment can be encoded at a different resolution and bit rate. The goal of MPEG-DASH is therefore to deliver the best quality content with the fewest dropouts and least possible buffering.

```

%Titles and Addresses of Recommended Programmes that should be displayed to each user
%-----
for prog=1:VHrow;
    [VHrow1,VHcolumn1] = size(USERSMATRIX{11, userColumn}{prog,WatchingTime});
    if VHrow1>0
        if strcmp(USERSMATRIX{11,userColumn}{prog,WatchingTime}{1,1}, 'Null')==0
            if Rate<0.3
                VHTitle{prog,1} = USERSMATRIX{11, userColumn}{prog,WatchingTime}{1,1};
                VHAddress{prog,1} = 'Sorry No Enough Bandwidth';
            else
                if numel(cell2mat(regex(USERSMATRIX{11, userColumn}{prog,WatchingTime}{8,1}, modes')))>1
                    VHTitle{prog,1} = USERSMATRIX{11, userColumn}{prog,WatchingTime}{1,1};
                    VHAddress{prog,1} = USERSMATRIX{11, userColumn}{prog,WatchingTime}{22,1};
                elseif regexp(USERSMATRIX{11, userColumn}{prog,WatchingTime}{8,1}, 'Null')>0
                    VHTitle{prog,1} = USERSMATRIX{11, userColumn}{prog,WatchingTime}{1,1};
                    VHAddress{prog,1} = USERSMATRIX{11, userColumn}{prog,WatchingTime}{22,1};
                else
                    VHTitle{prog,1} = USERSMATRIX{11, userColumn}{prog,WatchingTime}{1,1};
                    VHAddress{prog,1} = 'Not Available Due to Low Bandwidth';
                end
            end
        end
    end
end
end
end

%Filtering Operation to remove the programs that are not match the available bandwidth
%-----
USERSMATRIX{17, userColumn} = {};
FilterVHTitle = {};
FilterVHAddress = {};

for filterprog=1:VHrow;
    [VHrow1,VHcolumn1] = size(USERSMATRIX{11, userColumn}{filterprog,WatchingTime});
    if VHrow1>0
        [filterrow,filtercolumn] = size(USERSMATRIX{17,userColumn});
        if strcmp(VHAddress{filterprog,1}, 'Not Available Due to Low Bandwidth')==0
            USERSMATRIX{17,userColumn}{filterrow+1,1} = USERSMATRIX{11,userColumn}{filterprog,WatchingTime};
            save ('F:\Matlab training files\USERSMATRIX.mat', 'USERSMATRIX');
            FilterVHTitle{filterrow+1,1} = USERSMATRIX{17,userColumn}{filterrow+1,1}{1,1};
            FilterVHAddress{filterrow+1,1} = USERSMATRIX{17,userColumn}{filterrow+1,1}{22,1};
        end
    end
end

%The end of Filtering Operation
%-----

```

Fig. 5.20 Part of Matlab Code that is used for filtering the Recommended Programmes based on Available Network Bandwidth

YouTube and Netflix already support MPEG-DASH, but not all of the other online sources do. However, for those online sources that do support MPEG-DASH, the programme's metadata file contains information that refers to its usage as shown in figure 5.21 which shows a manifest file which defines the adaptive DASH-MPD within a "Media Presentation Description". In this case, the search agent translates the resolution in the metadata file to "Null" in the unified format, which is an equivalent case to that where there is no information about the resolution. The system in these cases provides the

programme directly without applying a filtering operation (YouTube, 2014, Mozilla, 2017 & Jan, 2011).

```
<?xml version="1.0" encoding="UTF-8"?>
<MPD xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
  xmlns="urn:mpeg:dash:schema:mpd:2011"
  xsi:schemaLocation="urn:mpeg:dash:schema:mpd:2011 DASH-MPD.xsd"
  type="dynamic"
  profiles="urn:mpeg:dash:profile:isoff-live:2011"
  minimumUpdatePeriod="PT60S"
  minBufferTime="PT12S"
  availabilityStartTime="2016-04-13T20:52:58" >
<Period start="PT0S" id="1">
  <AdaptationSet mimeType="video/webm">
    <ContentComponent contentType="video" id="1"/>
    <SegmentTemplate timescale="1000"
      duration="2000"
      startNumber="1"
      initialization="data:video/mp4;base64,AAAAGGZ0eXBpc...AAA"
      media="/dash_upload?cid=xxxx-xxxx-xxxx-xxxx&copy=0&file=media-$Number%09d$.webm"/>
    <Representation id="1" width="1920" height="1080">
      <SubRepresentation contentComponent="1"/>
    </Representation>
  </AdaptationSet>
</Period>
</MPD>
```

Fig. 5.21 YouTube Example for encoding MPEG-DASH stream showing Media Presentation Description (MPD) (YouTube)

However, selecting these viewing environmental features for the filtering operation is not finally where any other features could be added to this operation to be as filtering criteria. These features are used in the implementation of the system to demonstrate that the system is able to filter the final list of programmes based on the chosen viewing environmental features.

5.4 Context-aware Module

The context-aware module stores information that is relevant to each user. This module has been implemented in Matlab to create the user matrix and reads the viewing environmental features. The user matrix includes the viewing history, information extracted from social networks, the various lists of recommended programmes and the dynamic information which is relevant to the user's viewing environment.

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 | 21 | 22 | 23 | 24 | 25 |
|---|------|------|------|------|------|------|------|------|------|------|-------------|------|-------------|-------------|------|-------------|------|------|------|-------------|------|-------------|------|-------------|------|
| 1 | Null | Null | Null | Null | Null | Null | Null | Null | Null | Null | <23x2 cell> | Null | <23x3 cell> | <23x4 cell> | Null | <23x5 cell> | Null | Null | Null | <23x4 cell> | Null | <23x3 cell> | Null | <23x2 cell> | Null |
| 2 | | | | | | | | | | | | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | | | | | | | | | | | | | |

Fig. 5.23 User Viewing History

Note that hours that don't include any watched programmes are referred to as 'Null' while the hours that include watched programmes contain records such as <23x2 cell>. This record represents the programmes that were watched in that particular hour and are defined by a column comprising 23 rows which represent the 21 programme features plus the address of the programme plus the time of viewing. The second number represents the number of programmes that were watched during that particular hour. Therefore, <32x2 cell> means that two programmes were watched in that hour. When the user watches a new programme through the EPG, the system sends the programme metadata to be saved in the user's viewing history and the time it was watched. Figure 5.24 shows a part of the Matlab code that used to save the watched programme in user viewing history. The full Matlab code is listed in Appendix (VI).

```

%Saving the Watched Programme in User Viewing History
%-----
for H=1:22;
    %Saving the Watch Programme in the Viewing History
    %-----
    USERSMATRIX{3,userColumn}{1,WatchingTime}{H,column+1} = SelectedProgram(H,1)

    %Saving Relevant Information that will be used later in recommending Programmes based on
    %Collaborative Filtering Technique
    %-----
    if H==1
        USERSMATRIX{4,userColumn}{1,TVHcolumn+1} = SelectedProgram(H,1)
    end
    if H==3
        USERSMATRIX{4,userColumn}{2,TVHcolumn+1} = SelectedProgram(H,1)
        USERSMATRIX{4,userColumn}{3,TVHcolumn+1} = [3 userColumn; 1 WatchingTime; 1 column+1]
    end
    if H==6
        USERSMATRIX{4,userColumn}{4,TVHcolumn+1} = SelectedProgram(H,1)
    end
    save ('F:\Matlab training files\USERSMATRIX.mat', 'USERSMATRIX');
end
%Saving the Viewing Time
%-----
USERSMATRIX{3,userColumn}{1,WatchingTime}{23,column+1} = datestr(now);
save ('F:\Matlab training files\USERSMATRIX.mat', 'USERSMATRIX');
end
%Updating the list of programme's genres for the user for clustering process
%-----
if strcmp(USERSMATRIX{13,userColumn}{1,1}, 'Null')==1
    USERSMATRIX{13,userColumn}{1,1} = cellstr(SelectedProgram(1,1))
    if strcmp(USERSMATRIX{13,userColumn}{2,1}, 'Null')==1
        USERSMATRIX{13,userColumn}{2,1} = cellstr(SelectedProgram(3,1))
    else
        USERSMATRIX{13,userColumn}{2,1} = union(USERSMATRIX{13,userColumn}{2,1}, cellstr(SelectedProgram(3,1)))
    end
    USERSMATRIX{13,userColumn}{3,1} = cellstr(SelectedProgram(6,1))
else
    USERSMATRIX{13,userColumn}{1,1} = union(USERSMATRIX{13,userColumn}{1,1}, cellstr(SelectedProgram(1,1)))
    USERSMATRIX{13,userColumn}{2,1} = union(USERSMATRIX{13,userColumn}{2,1}, cellstr(SelectedProgram(3,1)))
    USERSMATRIX{13,userColumn}{3,1} = union(USERSMATRIX{13,userColumn}{3,1}, cellstr(SelectedProgram(6,1)))
end
save ('F:\Matlab training files\USERSMATRIX.mat', 'USERSMATRIX');

```

Fig. 5.24 Part of the Matlab Code that used to save the Watched Programme in User Viewing History

This Matlab code shown in figure 5.24 is responsible for saving the metadata parameters of the watched programme where each programme contains 21 features plus the address. The saving operation is implemented in more than one place where the programme features will be used later in different operations. For example, the second group of commands are responsible for saving the title, genre, pointer (for location) and the search keyword that relevant to this programme. This information is used in creating the viewing history of the groups that resulted from collaborative filtering technique. The third group of commands is responsible for updating the list of programme's genres for that particular user in order to use this list later in the user clustering process.

This information is then used by the matching operation block in recommending programmes based on the user's viewing history. Moreover, the user matrix includes information obtained from social networks comprises the social network ID, the user's preferred programmes and lists of programmes from their friends. Each part of this information is saved in a separate cell in the user matrix. Additionally, the user matrix includes cells for watched programme categories and the recommended list of programmes based on collaborative filtering technique. The user matrix includes cells for the results of the analysing operations and other services such as pre-paid online sources.

Finally, the user matrix includes a cell for the dynamic information relating with the user's viewing environmental features. Figure 5.25 shows an overview for the user matrix for several users.

5.5 Content Database Server

The content database server is represented by a matrix in Matlab containing the programmes that have been searched. These programmes have resulted from the different search operation (viewing history, user clustering and social networks) where the programme metadata has been translated, collated and saved as columns each comprising the programme's features. This matrix is known as the content matrix and consists of columns where each column is linked to a specific user. As a result, the number of columns in the content matrix is equal to the number of columns in the user matrix. Figure 5.26 shows the content matrix for several users. Each column contains the programmes that have been searched and retrieved from the online content sources. Figure 5.27 shows some of the programmes that are located in one cell (column) of the content matrix.

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
|----|-------------------|--------------------|-------------------|------------------|-------------------|--------------------|--------------------|--------------------|--------------------|--------------------|---------------|
| 1 | Emad Abdulrazzaq' | 'Ahmed Alsahlani' | 'Alaa Fahh' | 'Hasanin Shaker' | 'Emad Mohammed' | 'Ahmed Altameemi' | 'Mustafa Alalaf' | 'Osamah Salim' | Ahmed Jumaa' | 'Asaad Qasim' | 'Durai' |
| 2 | 1051671294862992' | 10204115097822998' | '950264318342606' | 912966715467133' | '118398768588593' | '1148939678480832' | '1168198076535649' | '1123840110971875' | 10209351905730772' | '1327431880604368' | '1872364022' |
| 3 | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> | <1x24 cell> |
| 4 | <4x4 cell> | <4x8 cell> | <4x12 cell> | <4x23 cell> | <4x11 cell> | <4x4 cell> | <4x10 cell> | <4x12 cell> | <4x11 cell> | <4x14 cell> | <4x13 cell> |
| 5 | <2x3 cell> | <14x3 cell> | <8x3 cell> | <11x3 cell> | <1x3 cell> | <10x3 cell> | <8x3 cell> | <2x3 cell> | <7x3 cell> | <10x3 cell> | <9x3 cell> |
| 6 | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> | <1x4 cell> |
| 7 | <18x3 cell> | <132x3 cell> | <84x3 cell> | <99x3 cell> | [] | <87x3 cell> | <104x3 cell> | <19x3 cell> | <7x3 cell> | <101x3 cell> | <92x3 cell> |
| 8 | <11x3 cell> | <16x3 cell> | <14x3 cell> | <24x3 cell> | <13x3 cell> | <29x3 cell> | <11x3 cell> | <7x3 cell> | <8x3 cell> | <13x3 cell> | <9x3 cell> |
| 9 | <22x106 cell> | <22x252 cell> | <22x242 cell> | <22x336 cell> | <22x20 cell> | <22x280 cell> | <22x224 cell> | <22x86 cell> | <22x251 cell> | <22x227 cell> | <22x202 cell> |
| 10 | <29x3 cell> | <148x3 cell> | <98x3 cell> | <123x3 cell> | <13x3 cell> | <116x3 cell> | <115x3 cell> | <26x3 cell> | <86x3 cell> | <114x3 cell> | <101x3 cell> |
| 11 | <13x24 cell> | <24x24 cell> | <15x24 cell> | <26x24 cell> | <16x24 cell> | <13x24 cell> | <19x24 cell> | <17x24 cell> | <11x24 cell> | <22x24 cell> | <10x24 cell> |
| 12 | YES' | YES' | YES' | YES' | YES' | YES' | YES' | YES' | YES' | YES' | YES' |
| 13 | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> | <3x1 cell> |
| 14 | <23x2 cell> | <9x2 cell> | <25x2 cell> | <36x2 cell> | <25x2 cell> | <23x2 cell> | <41x2 cell> | <26x2 cell> | <45x2 cell> | <22x2 cell> | <38x2 cell> |
| 15 | <23x2 cell> | <9x2 cell> | <25x2 cell> | <36x2 cell> | <25x2 cell> | <23x2 cell> | <41x2 cell> | <26x2 cell> | <45x2 cell> | <22x2 cell> | <38x2 cell> |
| 16 | <22x23 cell> | <22x9 cell> | <22x25 cell> | <22x36 cell> | <22x25 cell> | <22x23 cell> | <22x41 cell> | <22x26 cell> | <22x45 cell> | <22x22 cell> | <22x38 cell> |

Fig. 5.25 User Matrix for Several Users

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | 12 | | | | |
|---|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------------|---------|--|--|--|
| 1 | <22x480 cell> | <22x139 cell> | <22x399 cell> | <22x326 cell> | <22x457 cell> | <22x102 cell> | <22x379 cell> | <22x317 cell> | <22x144 cell> | <22x309 cell> | <22x319 cell> | <22x280 cell> | <22x... | | | |
| 2 | | | | | | | | | | | | | | | | |
| 3 | | | | | | | | | | | | | | | | |
| 4 | | | | | | | | | | | | | | | | |

Fig. 5.26 Content Matrix for Several Users

The system checks those programmes before each search operation to find the relevant programmes to each search keyword. Keywords that don't match programmes in the content matrix are forwarded to the search agent leader to search the online content sources for relevant programmes. Furthermore, the recommendation module reads the content matrix when creating the recommended list of programmes based on the viewing history, where similarity calculations are applied between each programmes in the content matrix and each programme in user viewing history. Figure 5.28 shows a part of Matlab code that is responsible of checking the user content matrix for any relevant programmes to social network's recommended list of programmes. The entire Matlab code is listed in Appendix ((VI-Implicit Function) Facebook Info).

This code calculates the similarity between each programme title in the list that created based on Facebook contribution with each programme title in the content matrix. The list that created based on Facebook contains two groups of programmes which are programmes that are selected by friends and programmes that are selected by the user themselves. That is represented in the first group of commands as shown in the figure 5.28. The programmes that have been found in the content matrix tagged as number '1' while the programmes that have not been found tagged as number '0'. Thereafter, the titles of programmes that are tagged as '0' are collated in one file and forwarded to the search agent leader to search the online content sources as shown in the second group of commands.

| | 450 | 451 | 452 | 453 | 454 | 455 | 456 | 457 |
|----|---|--|--|---|---|---|---|---|
| 1 | GAME OF THRONES: COMPLETE HISTORY... GAME OF THRONES: A DAY IN THE LIFE | GAME OF THRONES: COMPLETE HISTORY AND... LOST CONCEPTS SERIES: PRESLEY & CAS... | GAME OF THRONES: COMPLETE HISTORY AND... BEFORE THEY WERE LEGENDS: ELVIS PRES... | PRISON BREAK: THE FINAL BREAK | PRISON BREAK | THE FINISH: DOCTORS REBOOT | THE VOICE | |
| 2 | GAME OF THRONES: COMPLETE HISTORY... GILPHEE THE EPIC SCALE OF GAME OF T... | GAME OF THRONES: COMPLETE HISTORY AND... ANIMATION, SCIENTIFIC FICTION, ACTION/AD... | GAME OF THRONES: COMPLETE HISTORY AND... ANIMATION, SCIENTIFIC FICTION, ACTION/AD... | THE MOVIE COVERS THE EVENTS... BARTON MACLAINE STARS AS A T... | A STAR STUDDO SPECIAL WRITTEN AND... BOSS (VOICES) IS THE STOR | | | |
| 3 | ANIMATION, SCIENTIFIC FICTION, ACTO... | DOCUMENTARY | DOCUMENTARY | ACTION/ADVENTURE, DRAMA, T... | ACTION/ADVENTURE, DRAMA | | | |
| 4 | 26MINUTE | 26MINUTE | 64MINUTE | 60MINUTE | 89MINUTE | 72MINUTE | 30MINUTE | 100MINUTE |
| 5 | 0.109883 | 0.005101 | 0.001001 | 0.077332 | 0.277332 | 0.000143 | 0.330282 | 0.002321 |
| 6 | GAME OF THRONES | GAME OF THRONES | HOUSE MD | LOST SERIES | PRISON BREAK | PRISON BREAK | THE DOCTORS | THE VOICE |
| 7 | 2015-02-08 | 2015-02-08 | 2011-05-13 | 2008-04-06 | 2009-05-26 | 1998-07-12 | 2013-11-23 | 2008-07-11 |
| 8 | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH |
| 9 | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH | ENGLISH |
| 10 | RUSSIA | UNITED STATES OF AMERICA | UNITED STATES OF AMERICA | UNITED STATES OF AMERICA | UNITED STATES OF AMERICA | UNITED STATES OF AMERICA | UNITED STATES OF AMERICA | PHILIPPINES |
| 11 | RELEASED | RELEASED | RELEASED | RELEASED | RELEASED | RELEASED | RELEASED | RELEASED |
| 12 | Null | Null | Null | Null | Null | Null | Null | Null |
| 13 | Null | Null | Null | Null | Null | Null | Null | Null |
| 14 | Null | Null | Null | Null | Null | Null | Null | Null |
| 15 | Null | Null | Null | Null | Null | Null | Null | Null |
| 16 | Null | Null | Null | Null | Null | Null | Null | Null |
| 17 | Null | Null | Null | Null | Null | Null | Null | Null |
| 18 | Null | Null | Null | Null | Null | Null | Null | Null |
| 19 | Null | Null | Null | Null | Null | Null | Null | Null |
| 20 | Null | Null | Null | Null | Null | Null | Null | Null |
| 21 | Null | Null | Null | Null | Null | Null | Null | Null |
| 22 | https://www.fantv/movies/269623 | https://www.fantv/movies/340200 | https://www.fantv/movies/269623 | https://www.fantv/movies/317502 | https://www.fantv/movies/116241 | https://www.fantv/movies/242693 | https://www.fantv/movies/246202 | https://www.fantv/movies/ |
| 23 | | | | | | | | |
| 24 | | | | | | | | |

Fig. 5.27 Programmes that located in One of the Cells (Columns) of Content Matrix

```

%Search the Content Matrix to find any programme similar to the items found in the facebook lists
%-----
if strcmp(USERSMATRIX{2,UserColumn},'NULL')==0
FirstFBProgramsMatrix1{1,UserColumn} = 'NULL';
for line=7:8;
[row1,column1] = size(USERSMATRIX{line,UserColumn});
[row2,column2] = size(USERSCONTENTMAT{1,UserColumn});
    for i=1:row1;
        USERSMATRIX{line,UserColumn}{i,3}=0;
        for l=1:column2;
            FirstTerm = regexp(upper(USERSMATRIX{line,UserColumn}{i,1}), '\s*:\s', 'split');
            ItemWords = numel(FirstTerm);
            SecondTerm = regexp(USERSCONTENTMAT{1,UserColumn}{1,l}, '\s*:\s', 'split');
            SimilarWords = intersect(SecondTerm, FirstTerm);
            Words = numel(SimilarWords);
            ratio = Words/ItemWords;
            if ratio>0.65
                USERSMATRIX{line,UserColumn}{i,3}=1;
                [row3,column3]=size(FirstFBProgramsMatrix1(:,UserColumn));
                if strcmp(FirstFBProgramsMatrix1{1,UserColumn},'NULL')==1
                    FirstFBProgramsMatrix1{1,UserColumn}=USERSCONTENTMAT{1,UserColumn}{1,l};
                end
                FirstFBProgramsMatrix1{row3+1,UserColumn}=USERSCONTENTMAT{1,UserColumn}{1,l};
            end
        end
    end
end
end
save ('F:\Matlab training files\USERSMATRIX.mat', 'USERSMATRIX');
save ('F:\Matlab training files\USERSCONTENTMAT.mat', 'USERSCONTENTMAT');

%Search the Online Content Sources for programmes that have not been found in the Content Matrix
%-----
NofMovies1 = numel(find([USERSMATRIX{7,UserColumn}{:,3}]==0))
MissingPro1 = find([USERSMATRIX{7,UserColumn}{:,3}]==0)
SearchList1 = USERSMATRIX{7,UserColumn}(MissingPro1,1)
NofMovies2 = numel(find([USERSMATRIX{8,UserColumn}{:,3}]==0))
MissingPro2 = find([USERSMATRIX{8,UserColumn}{:,3}]==0)
SearchList2 = USERSMATRIX{8,UserColumn}(MissingPro2,1)
NofMovies = NofMovies1+NofMovies2
SearchList = union(SearchList1, SearchList2)
PHPFile = fopen('C:\xampp\htdocs\NofKeywords.txt', 'w');
fprintf(PHPFile, '%d', NofMovies);
fclose(PHPFile);
Yfile = fopen('C:\xampp\htdocs\Keywords.txt', 'w');
for i=1:NofMovies
    fprintf(Yfile, '%s\r\n', SearchList{i,1});
end

```

Fig. 5.28 Part of Matlab Code that is responsible of Checking the User Content Matrix for any relevant Programmes to Social Network's Recommended List of Programmes

To better understand the data flow through the EPG, two scenarios are presented. The first one which is shown in figure 5.29, is when the user joins the system for the first time while the second, which is shown in figure 5.30, is for creating the recommended lists of programmes based on all of the recommendation schemes. Figure 5.29 Shows the first scenario where the number on the arrows are described after the figure.

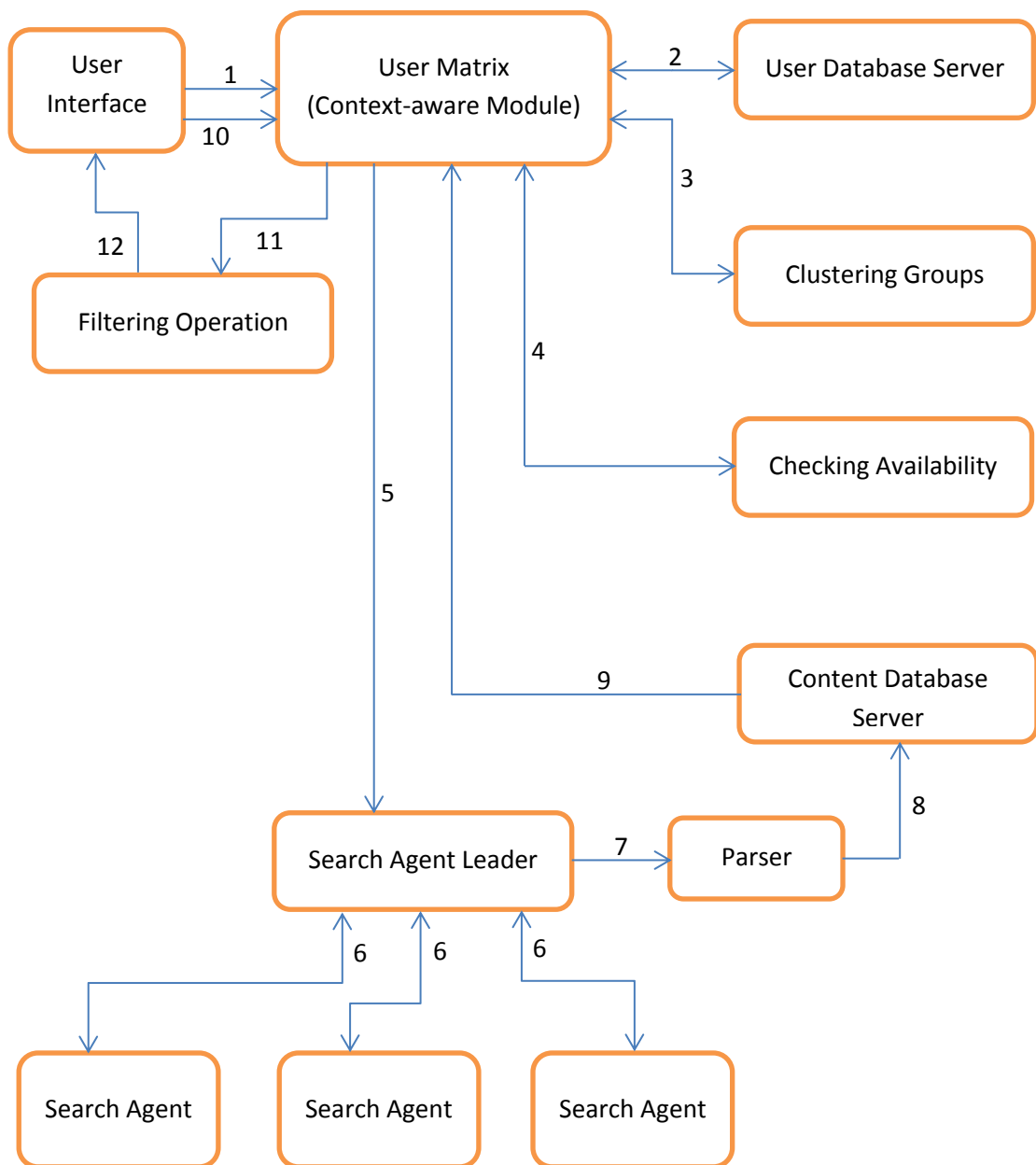


Fig. 5.29 Data Flow Diagram when the User joins the System for the First Time

1. Users register within the EPG system and select the preferred categories.
2. The context-aware module checks the user database server to find if the user has a Facebook account or not. If the user has Facebook account, the data will be converted to a matrix format and saved into the user matrix. This data includes the

information that retrieved from the user Facebook account when they joined the Facebook application.

3. Implementing collaborative filtering technique to find the user group and create the recommended list of programmes and save it in the user matrix.
4. Checking the availability of programmes in both lists (Facebook list and CF list) and saving the result of checking in the user matrix.
5. Sending the programmes that are not found in the content database server to the search agent leader to search the online content sources.
6. Searching and translating the retrieved programmes' metadata and forwarding the result to the search agent leader.
7. Combining the result into one matrix and removing the duplicated.
8. Saving the result in the content database server.
9. Sending a list of programmes that have been found in the online content sources and keeping a list of the missed programmes in order to search of them in the next day.
10. Users log on to their EPG and the viewing environmental features are forwarded to the context-aware module.
11. Filtering the recommended list of programmes according to the viewing environmental features.
12. Displaying the results to the user.

Similarly, figure 5.30 Shows the second scenario where the numbers on arrows are illustrated as shown below:

1. The context-aware module checks the user database server to find if the user has a Facebook account or not. If the user has Facebook account, the data will be converted to a matrix format and saved into the user matrix. This data includes the information that retrieved from the user Facebook account when they joined the Facebook application.
2. Implementing collaborative filtering technique to find the user group and create the recommended list of programmes and save it in the user matrix.
3. Checking the availability of programmes in both lists (Facebook list and CF list) and save the result of checking in the user matrix.

4. Sending the search keywords to the search agent leader which include the viewing history programmes, Facebook and CF programmes but only those which are not exist in the content database server.
5. Searching and translating the retrieved programmes' metadata and forwarding the result to the search agent leader.

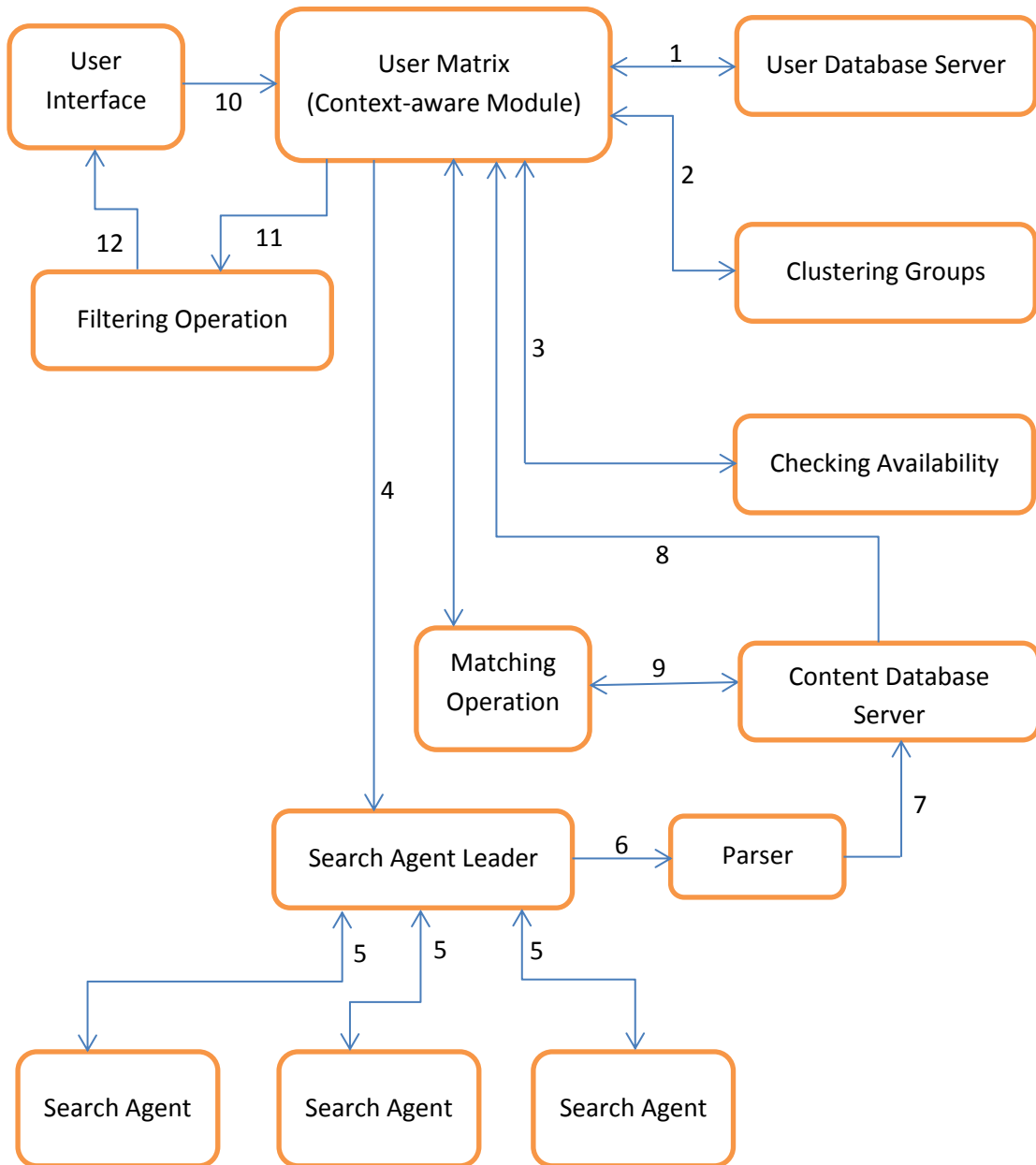


Fig. 5.30 Data Flow Diagram when the System creating the Recommended Lists of Programmes based on all of the Recommendation Schemes

6. Combining the result into one matrix and removing the duplicated.
7. Saving the result in the content database server.
8. Sending a list of programmes that have been found in the online content sources (for Facebook and CF only) and keeping a list of the missed programmes in order to search of them in the next day.
9. Implementing a matching operation between the programmes in the viewing history and the programmes in the content database server to create lists of programmes for each hour.
10. Users log on to their EPG and the viewing environmental features are forwarded to the context-aware module.
11. Filtering the recommended list of programmes according to the viewing environmental features.
12. Displaying the results to the user.

Finally, table 5.4 shows the overall system timings. Note that, for data purging, records in the user viewing history and the content database server which are older than one month are removed on a daily basis. This setting has been chosen for test purposes but it can be changed to suit the operational environment. Data purging is implemented by checking the watched time that is recorded with the programme features in the programme matrix in the user viewing history. For the content database server, the keyword of the deleted programme will be matched with the programmes in the content database server, then any programme has this keyword will be deleted as well.

Note that, if the deleted programmes were relevant to another keyword (such as if this keyword are common between two hours in the viewing history and the programme in that another hour is not older than one month), the programmes that deleted from the content database server will be found again in the next searching operation where all the keywords that haven't matched programmes in the content database server will be searched in the online content sources every day.

| Timelines of Operations | Operation |
|-------------------------|--|
| Real Time | <ol style="list-style-type: none"> 1. Update user viewing history and viewing pattern (programme categories). 2. Search the online content sources for particular keyword. 3. Filtering the recommended lists based on viewing environmental features. |
| Daily | <ol style="list-style-type: none"> 1. Searching online content sources based on recommendation schemes. 2. Clustering users into groups based on viewing pattern. 3. Update social network data for each user. 4. Recommending list of programmes based on all the recommendation schemes. 5. Data purging. |

Table 5.4 Timelines of the System Operations

5.6 Online Electronic Programme Guide

In order to evaluate the system a user interface has been created to allow users to use the services that are offered by the EPG. This allows users to create their own accounts using the same name that they use for accessing Facebook. The prototype EPG has three areas to display the recommended lists of programmes based on the three recommendation techniques. Moreover, it contains ON and OFF buttons to turn on and turn off the EPG. When the user turns the EPG on, the recommended lists of programmes will be displayed. By choosing any one of the recommended programmes, details of that programme will be displayed which include the programme's description, genre, length and production year. Additionally, the address of the selected programme is added to the displayed details in order to check the original online source.

Another display area has been added to the interface for doing a direct keyword search. Two buttons have been added for this purpose; the first one is used to input the search keyword, where this keyword will be sent to the context-aware module program and from

there the search agent leader will search the online content sources. Thereafter, by pressing the second button, the system will display the list of programmes that are recommended. Figure 5.31 shows a snapshot of the designed EPG.

The interface also allows users to watch the selected programmes, where a button has been located in the interface for this purpose. This button opens an Internet browser with the address of the selected programme. The system then automatically sends the programme's features matrix of this watched programme to the context-aware module so that the user viewing history can be updated.

In order to enhance the recommendation process and avoid the cold-start problem, the user interface has been provided with a list of categories. The user is promoted to select their preferred categories during the registration stage. This helps the system to provide a list of recommended programmes based on the selected categories (Collaborative Filtering Technique) even when the user hasn't watched any programme and therefore has no viewing history. Furthermore, the users can add additional categories at any time when they turn on the EPG.

Because some of the online content sources are pre-paid or subscription sources, the user interface has been provided with an option to either includes the pre-paid sources and sources that need subscription or not. This is done through providing a choice within the designed interface and the user can select 'Yes' to include those sources or select 'NO' to avoid the search results that come from those sources during the registration stage. The full configuration software of the designed EPG is located in Appendix VI which combines all the functions within the proposed system.

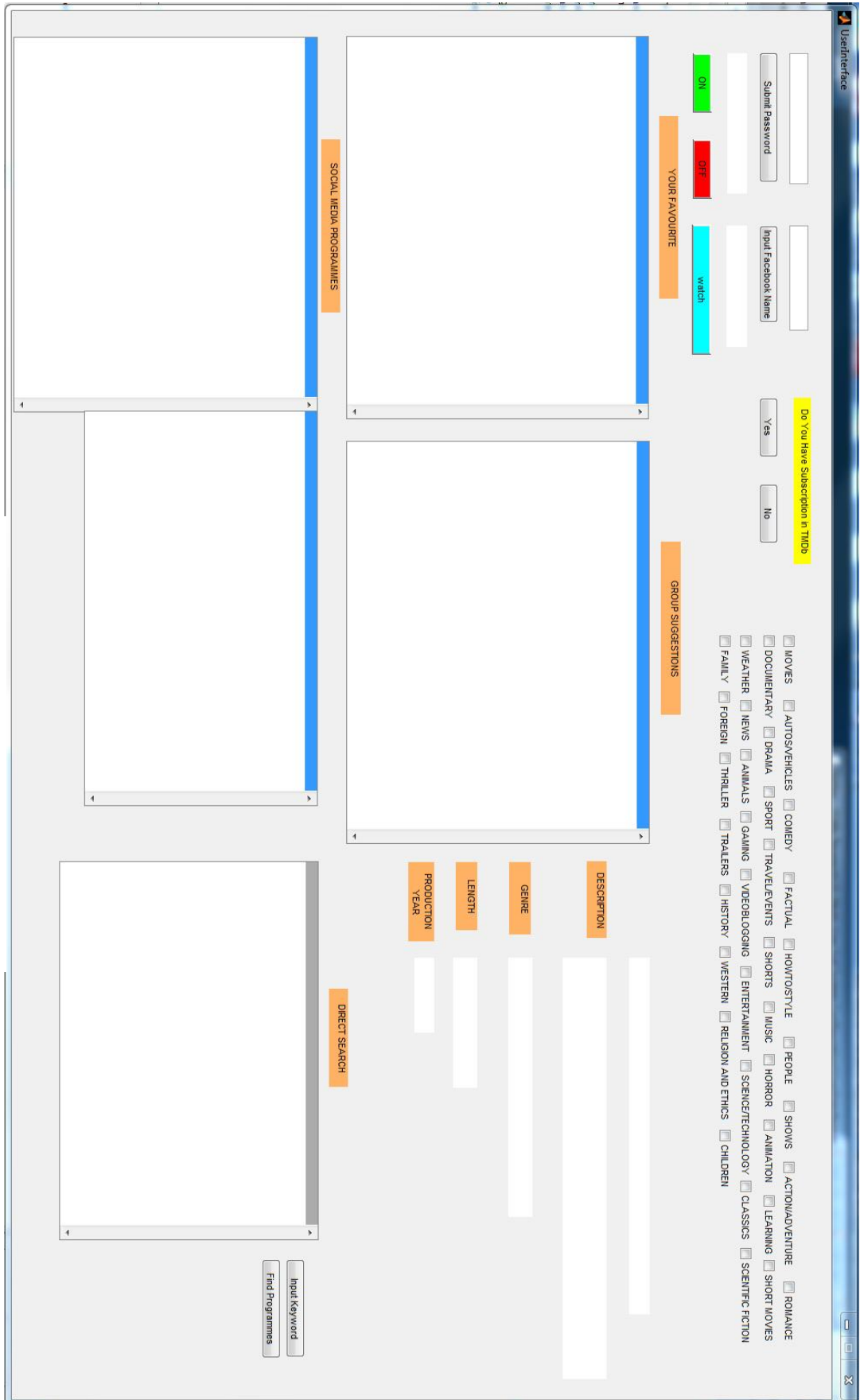


Fig. 5.31 Snapshot of the Designed EPG

5.7 Summary

In this chapter, the implementation of the proposed EPG system has been described. The different functional blocks have been created using PHP and Matlab. According to the current implementation, the EPG system has the following features:

1. The content search module is able to retrieve the programmes' metadata from all the connected online video content sources which are YouTube, BBC iPlayer, Dailymotion and The Movie Database.
2. All of the metadata formats that are retrieved from the different online content sources are converted to a unified format, which include 21 programme features, that is used within the remainder of the system.
3. Multiple recommendation schemes have been used to enhance the recommendation process which are content-based, collaborative filtering and context-awareness.
4. Integration of social networks has been realised through an interface to Facebook from which a list of programmes is obtained based on the users' preferences and their friends and is used to enhance the recommendation operation.
5. The system is generic and flexible to any expansion in the number of online video content sources where the core system architecture remains the same as new online content sources are added. The only bespoke module is the search module which has to be unique for each new online content source.
6. The system is provided as a cloud based service in which all recommendation operations are carried out in a centralised server with the results being stored in that server and made available to the user when they log onto the EPG.
7. A filtering stage has been created to filter the recommended programmes based on the resolution of user viewing device and network bandwidth.

The next chapter will discuss the evaluation of the EPG functions based on the implementation which has been described in this chapter.

CHAPTER SIX

SYSTEM EVALUATION

Chapter Six

6.1 Introduction

The previous chapter illustrated the details of the EPG system implementation. This chapter describes the processes through which the system was tested and evaluated. In order to validate the proposed architecture, the EPG system has been evaluated through assessing its performance according to the main functions that were presented in chapter three. As mentioned in the previous chapter, in order to evaluate the proposed system, a basic user interface was designed to allow the users use the system and its services.

6.2 User Trial Scenario

In order to evaluate the designed system, 20 users were chosen to participate in a user trial by creating accounts in the EPG. The same name used for the EPG account is also used when joining the Facebook application. Once an account has been created, each user then selects their preferred programme categories and chooses whether they want to view pre-paid online content within the recommended programmes or not.

Each one of these users had a Facebook account and all of them joined the associated Facebook EPG application as shown in Figure 6.1. Because the Facebook application is not published for all Facebook users (Application under test), these participants are added to the application as developers to give them the ability to see the application in their profile and then join it. However, to join the Facebook application, they needed to separately login to their Facebook accounts. Figure 6.2 shows the application within the list of user's application in Facebook account.

By joining the Facebook application, the user is asked to give permission to retrieve information relevant to their public profile, friends, likes and video activity as shown in Figure 6.3. When the user accepts this request, the application asks the user to submit their name that will be used in the EPG system as shown in Figure 5.18.

The next step of testing was to build the users' viewing history by watching different programmes at different times of the day. This meant that each user had to login to their account at different hours and watch programmes through the EPG system. This step was performed on different days.

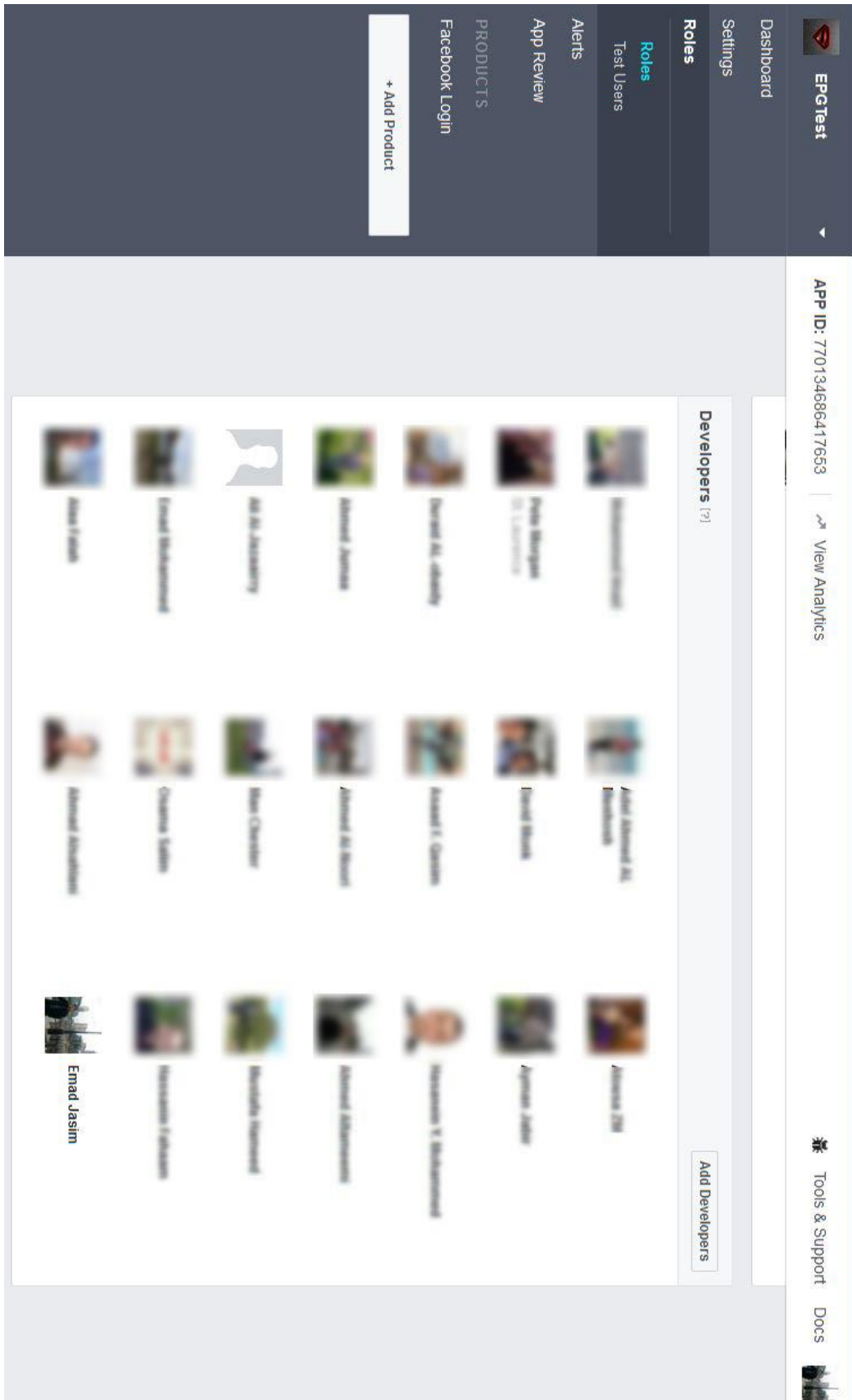


Fig. 6.1 The Users who joined the Facebook Application

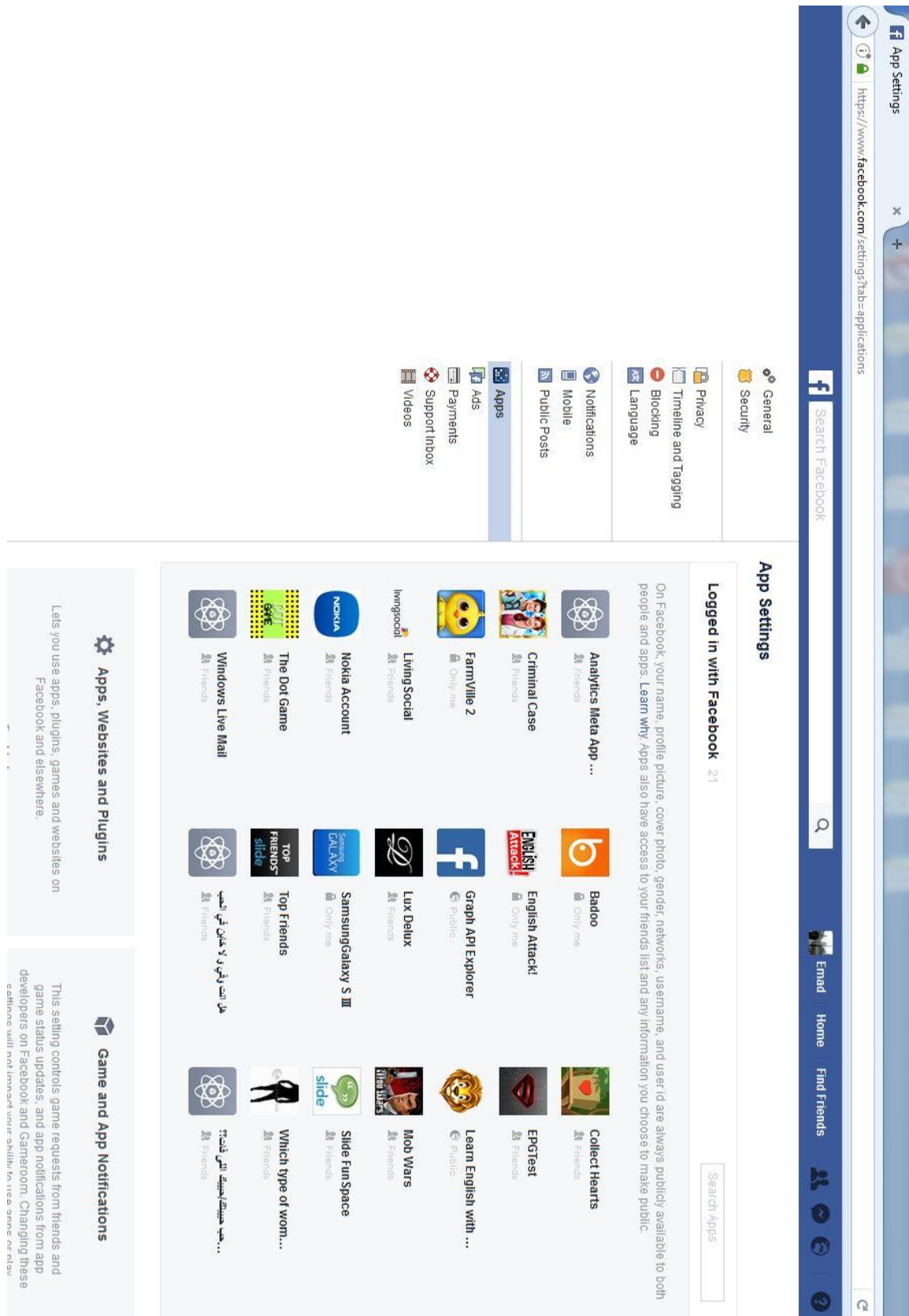


Fig. 6.2 The EPG Application (EPGTest) within the List of Application in a Facebook User’s Profile

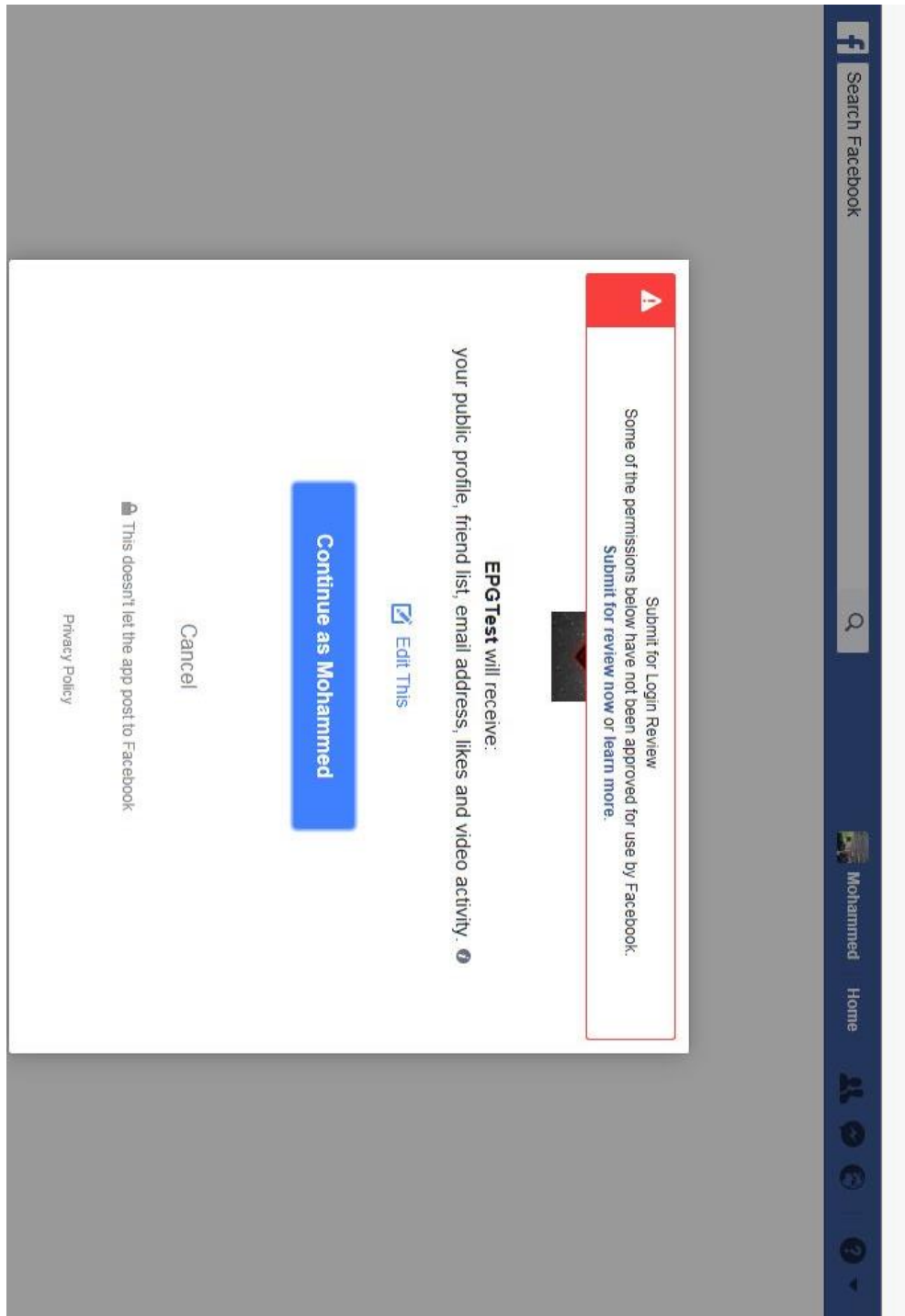


Fig. 6.3 Facebook Application Request for Information Retrieving Permission

Thereafter, the users were asked to evaluate the system's recommendation by rating the list of recommended programmes that the EPG produced. This rating was then used to evaluate the recommendation process and the effect of integrating the social networks in enhancing the recommendation operation.

6.3 Evaluating the Function of Multiple Online Content Sources

As mentioned previously, the designed system is connected to BBC-iPlayer, YouTube, Dailymotion and The Movie Database as example of online content sources. The system retrieves the programmes' metadata from those sources in the searching operation whether in order to create the recommended lists of programmes based on the system operations or through a direct keyword search initiated by the user. By evaluating the results that were obtained from user trials, the list of recommended programmes has been checked to prove whether or not the system was able to retrieve the programmes' metadata from all the connected online sources. Figure 6.4.a shows on the left, the recommended list of programmes which consists of 21 programmes where the second programme in the list has been chosen and the details of that programme have been displayed on the right hand side. As shown in the figure 6.4.a, the first line in the details shows the address of that programme which in this case is YouTube. Figure 6.4.b shows the third programme in the same list being selected where the address in that case refers to Dailymotion. Similarly, figure 6.4.c shows another programme in the same list refers to BBC-iPlayer and figure 6.4.d shows that another programme belongs to The Movie database. In addition to the source of the programme, other details are presented which include description, genre, length and production year.

It is clear that the system can retrieve programme metadata from all of the connected online content sources even though these sources offer different interfaces and use different metadata formats. Hence, the EPG system can translate these different metadata formats into a single unified format allowing all programmes to be displayed to the user in the same way irrespective of where they came from. However, the set of metadata (description, genre, length and production year) that is displayed here for test purposes and clearly, more or less information can be presented in the final EPG.

On the other hand, the searching time depends on the number of search keywords sent to the search agent leader. Keywords arise from several processes but all relate to programmes, which are not currently stored within the EPG's content database server. According to the design of the search agent leader, these keywords are then sent to each search agent at the same time. A search agent interfaces to a content provider via an API through which a search is initiated. Where the API only permits a single keyword search, then the maximum total search time will scale as a function of the number of keywords.

As an example, if the number of search keywords is doubled, then the worst case time needed for the whole searching operation will also be doubled. Hence, in the worst case the total searching time scales linearly with the number of search keywords.

Practically, the time performance of the EPG depends on many different parameters such as the server hardware, the operating system and the programming language used to implement the EPG system. In addition, for the searching module, the required time also depends on the performance of the content provider's own systems, the efficiency of their API and the bandwidth of the connection, which links them to the EPG.

Moreover, increasing the number of users will increase the overall time performance of the EPG system, where the processing time will increase linearly with the number of users. Whilst, the search operation associated with a single user is determined by the number of search keywords, clearly with more users, the total time devoted to search operations by the EPG system, as a whole, must increase. However, the vast majority of this search time is allocated whilst the user is logged off from the EPG hence, masking it from them. Increasing the number of users also increases the amount of stored data in the system within both the user matrix and content matrix. A calculation is presented in the summary section of this chapter that provides an approximate value for how this storage requirement grows as a function of the number of users.

YOUR FAVOURITE

THE AVENGERS - HULK SMASH
MINIONS THE AVENGERS ADVENTURES (SUPERHEROES IDOL) - FUNNY CARTOON (HD)
AVENGERS 2 FULL HD MOVIE 2015
MARVEL'S THE AVENGERS- AGE OF ULTRON 2015 - TRALER #3
ROBIN HOOD: COMHSTRISA CHOLLE: CEANNSACHADH NA COLLE
ROBIN HOOD: COMHSTRISA CHOLLE: LEASAMAN RÁ-OGHAL
ROBIN HOOD: COMHSTRISA CHOLLE: AN SUANÉALACHE
ROBIN HOOD: COMHSTRISA CHOLLE: AN UGHICHE-BRÁIGE
ANIMF (SA BRANCH) - ROBIN HOOD TAX
ROBIN HOOD: COMHSTRISA CHOLLE: AN FEAR FUDADAN
ROBIN HOOD: COMHSTRISA CHOLLE: TÁR NA TÁSGADH
ROBIN HOOD: COMHSTRISA CHOLLE: MANGHSTRA NAIPPAD
ROBIN HOOD: COMHSTRISA CHOLLE: CLADHEAMH NA RÁ-GHREAN
THE AVENGERS
THE AVENGERS
AVENGERS: AGE OF ULTRON
THE STORY OF BEAN
ROBIN HOOD: COMHSTRISA CHOLLE: CROIS DHÁ BALT NA CARAD
ROBIN HOOD: COMHSTRISA CHOLLE: AN T-ÁIR NEO-FHACSNINÉACH
AVENGERS INFINITY WAR CAPTAIN MARVEL BREAKDOWN
AVENGERS: AGE OF ULTRON - CLIP IRON MAN VS. ULTRON (HD) - SUBTITULADO POR
AVENGERS: AGE OF ULTRON TRAILERS 5

GROUP SUGGESTIONS

SAUNA ADELE SADIO FEAT BLOODLINE (OFFICIAL HD MUSIC VIDEO) (Watched By (1) Subscribers)
ADELE - HELLO (Watched By (1) Subscribers)
ADELE - SOMEONE LIKE YOU (GALSTONBURY 2016) (Watched By (1) Subscribers)
FAMILY GUY - THE SIMPSONS (24 HOUR) LA'VE (Watched By (1) Subscribers)
HOW I MET YOUR MOTHER SEASON 9 EPISODE 6 7 8 9 - THE POKER GAME (Watched By (1) Subscribers)
INDIANA JONES AND THE TEMPLE OF DOOM (1984) - HARRISON FORD, KATE CAPSHAW INSIDE THE MIND OF A FEMINIST - A NATIONAL GEOGRAPHIC SPECIAL (Watched By (1) Subscribers)
KETCHUP! HOW IT'S MADE (Watched By (1) Subscribers)
LEFT MUSIC (Watched By (1) Subscribers)
SHAUN THE SHEEP: SERIES 2: CHEETAH CHEATER (Watched By (1) Subscribers)
SPIDER-MAN (Watched By (1) Subscribers)
THE BIG BANG THEORY: IT ALL STARTED WITH A BIG BANG (Watched By (1) Subscribers)
THE BOY OFFICIAL TRALER #1 (2016) LAUREN COHAN HORROR MOVIE HD (Watched By (1) Subscribers)
THE SIMPSONS - BART'S BIRTH (Watched By (1) Subscribers)
TOM AND JERRY, 63 EPISODE - THE FLYING CAT (1952) (Watched By (1) Subscribers)
TOM AND JERRY, 64 EPISODE - THE DUCK DOCTOR (1952) (Watched By (1) Subscribers)
TOP 10 HOW I MET YOUR MOTHER EPISODES (Watched By (1) Subscribers)
TRANSFORMERS: AGE OF EXTINCTION (Watched By (1) Subscribers)
UNDERWORLD: EVOLUTION (Watched By (1) Subscribers)

<https://www.youtube.com/watch?v=r-1ohtTfgLQ>

| | |
|-----------------|---|
| DESCRIPTION | MINIONS THE AVENGERS ADVENTURES (SUPERHEROES IDOL) - FUNNY CARTOON AVENGERS MINIONS: JUDGES: NICK FURY, STAN LEE AND JOSS WHEDC |
| GENRE | ENTERTAINMENT |
| LENGTH | 15MINUTE |
| PRODUCTION YEAR | Null |

6.4.a. Programme Belongs to YouTube Web

YOUR FAVOURITE

THE AVENGERS - HULK SMASH
MINIONS THE AVENGERS ADVENTURES (SUPERHEROES IDOL) - FUNNY CARTOON (HD)
AVENGERS 2 FULL HD MOVIE 2015
MARVEL'S THE AVENGERS- AGE OF ULTRON 2015 - TRALER #3
ROBIN HOOD: COMHSTRISA CHOLLE: CEANNSACHADH NA COLLE
ROBIN HOOD: COMHSTRISA CHOLLE: LEASAMAN RÁ-OGHAL
ROBIN HOOD: COMHSTRISA CHOLLE: AN SUANÉALACHE
ROBIN HOOD: COMHSTRISA CHOLLE: AN UGHICHE-BRÁIGE
ANIMF (SA BRANCH) - ROBIN HOOD TAX
ROBIN HOOD: COMHSTRISA CHOLLE: AN FEAR FUDADAN
ROBIN HOOD: COMHSTRISA CHOLLE: TÁR NA TÁSGADH
ROBIN HOOD: COMHSTRISA CHOLLE: MANGHSTRA NAIPPAD
ROBIN HOOD: COMHSTRISA CHOLLE: CLADHEAMH NA RÁ-GHREAN
THE AVENGERS
THE AVENGERS
AVENGERS: AGE OF ULTRON
THE STORY OF BEAN
ROBIN HOOD: COMHSTRISA CHOLLE: CROIS DHÁ BALT NA CARAD
ROBIN HOOD: COMHSTRISA CHOLLE: AN T-ÁIR NEO-FHACSNINÉACH
AVENGERS INFINITY WAR CAPTAIN MARVEL BREAKDOWN
AVENGERS: AGE OF ULTRON - CLIP IRON MAN VS. ULTRON (HD) - SUBTITULADO POR
AVENGERS: AGE OF ULTRON TRAILERS 5

GROUP SUGGESTIONS

SAUNA ADELE SADIO FEAT BLOODLINE (OFFICIAL HD MUSIC VIDEO) (Watched By (1) Subscribers)
ADELE - HELLO (Watched By (1) Subscribers)
ADELE - SOMEONE LIKE YOU (GALSTONBURY 2016) (Watched By (1) Subscribers)
FAMILY GUY - THE SIMPSONS (24 HOUR) LA'VE (Watched By (1) Subscribers)
HOW I MET YOUR MOTHER SEASON 9 EPISODE 6 7 8 9 - THE POKER GAME (Watched By (1) Subscribers)
INDIANA JONES AND THE TEMPLE OF DOOM (1984) - HARRISON FORD, KATE CAPSHAW INSIDE THE MIND OF A FEMINIST - A NATIONAL GEOGRAPHIC SPECIAL (Watched By (1) Subscribers)
KETCHUP! HOW IT'S MADE (Watched By (1) Subscribers)
LEFT MUSIC (Watched By (1) Subscribers)
SHAUN THE SHEEP: SERIES 2: CHEETAH CHEATER (Watched By (1) Subscribers)
SPIDER-MAN (Watched By (1) Subscribers)
THE BIG BANG THEORY: IT ALL STARTED WITH A BIG BANG (Watched By (1) Subscribers)
THE BOY OFFICIAL TRALER #1 (2016) LAUREN COHAN HORROR MOVIE HD (Watched By (1) Subscribers)
THE SIMPSONS - BART'S BIRTH (Watched By (1) Subscribers)
TOM AND JERRY, 63 EPISODE - THE FLYING CAT (1952) (Watched By (1) Subscribers)
TOM AND JERRY, 64 EPISODE - THE DUCK DOCTOR (1952) (Watched By (1) Subscribers)
TOP 10 HOW I MET YOUR MOTHER EPISODES (Watched By (1) Subscribers)
TRANSFORMERS: AGE OF EXTINCTION (Watched By (1) Subscribers)
UNDERWORLD: EVOLUTION (Watched By (1) Subscribers)

<http://www.dailymotion.com/video/x2psjuv>

| | |
|-----------------|-------------------------------|
| DESCRIPTION | AVENGERS 2 FULL HD MOVIE 2015 |
| GENRE | MOVIES |
| LENGTH | 120MINUTE |
| PRODUCTION YEAR | Null |

6.4.b. Programme Belongs to Dailymotion Web

[GROUP SUGGESTIONS](#)

THE AVENGERS - HULK SMASH
MINIONS THE AVENGERS ADVENTURES (SUPERHEROES DOL) - FUNNY CARTOON (HD)
AVENGERS 2: FULL HD MOVIE 2015
MARVEL'S THE AVENGERS- AGE OF ULTRON 2015 - TRAILER #3
ROBIN HOOD: COMHSTRI SA CHOLLE: CEANNSACADH NA COLLE
ROBIN HOOD: COMHSTRI SA CHOLLE: LEASANAI RA-OGHAL
ROBIN HOOD: COMHSTRI SA CHOLLE: AN SUAIMIALAICHE
ROBIN HOOD: COMHSTRI SA CHOLLE: AN LEHCHE-ERÁIGE
ANIN (SA BRANCH) - ROBIN HOOD TAX
ROBIN HOOD: COMHSTRI SA CHOLLE: AM FEAR FIADAIN
ROBIN HOOD: COMHSTRI SA CHOLLE: TÁIR NA TÁSGADH
ROBIN HOOD: COMHSTRI SA CHOLLE: MAIGHSTR NAH PUPAD
ROBIN HOOD: COMHSTRI SA CHOLLE: CLADHEAHIH NAN RA-SHREAN
THE AVENGERS
AVENGERS: AGE OF ULTRON
THE STORY OF BEAN
ROBIN HOOD: COMHSTRI SA CHOLLE: CROIS DHÁ'BALT NA CARAD
ROBIN HOOD: COMHSTRI SA CHOLLE: AN T-ÁIR NEO-FHACSNIEACH
AVENGERS: INFINITY WAR CAPTAIN MARVEL BREAKDOWN
AVENGERS: AGE OF ULTRON - CLIP IRON MAN VS. ULTRON (HD) - SUBTITULADO POR
AVENGERS: AGE OF ULTRON TRAILERS 5

[YOUR FAVOURITE](#)

THE AVENGERS - HULK SMASH
MINIONS THE AVENGERS ADVENTURES (SUPERHEROES DOL) - FUNNY CARTOON (HD)
AVENGERS 2: FULL HD MOVIE 2015
MARVEL'S THE AVENGERS- AGE OF ULTRON 2015 - TRAILER #3
ROBIN HOOD: COMHSTRI SA CHOLLE: CEANNSACADH NA COLLE
ROBIN HOOD: COMHSTRI SA CHOLLE: LEASANAI RA-OGHAL
ROBIN HOOD: COMHSTRI SA CHOLLE: AN SUAIMIALAICHE
ROBIN HOOD: COMHSTRI SA CHOLLE: AN LEHCHE-ERÁIGE
ANIN (SA BRANCH) - ROBIN HOOD TAX
ROBIN HOOD: COMHSTRI SA CHOLLE: AM FEAR FIADAIN
ROBIN HOOD: COMHSTRI SA CHOLLE: TÁIR NA TÁSGADH
ROBIN HOOD: COMHSTRI SA CHOLLE: MAIGHSTR NAH PUPAD
ROBIN HOOD: COMHSTRI SA CHOLLE: CLADHEAHIH NAN RA-SHREAN
THE AVENGERS
AVENGERS: AGE OF ULTRON
THE STORY OF BEAN
ROBIN HOOD: COMHSTRI SA CHOLLE: CROIS DHÁ'BALT NA CARAD
ROBIN HOOD: COMHSTRI SA CHOLLE: AN T-ÁIR NEO-FHACSNIEACH
AVENGERS: INFINITY WAR CAPTAIN MARVEL BREAKDOWN
AVENGERS: AGE OF ULTRON - CLIP IRON MAN VS. ULTRON (HD) - SUBTITULADO POR
AVENGERS: AGE OF ULTRON TRAILERS 5

[GROUP SUGGESTIONS](#)

SAUNA ADEEL SADIQ FEAT BLOODLINE (OFFICIAL HD MUSIC VIDEO) (Watched By (1))
ADEL - HELLO (Watched By (1) Subscribers)
ADEL - SOMEONE LIKE YOU (GUSTONBURY 2016) (Watched By (1) Subscribers)
FAMILY GUY - THE SIMPSONS (24 HOUR) L.A.'VE (Watched By (1) Subscribers)
HOW I MET YOUR MOTHER SEASON 9 EPISODE 5 & 7 & 9 - THE POKER GAME (Watched B
INDIANA JONES AND THE TEMPLE OF DOOM (1984) - HARRISON FORD, KATE CAPSHAW
INSIDE THE MIND OF A FEMINIST - A NATIONAL GEOGRAPHIC SPECIAL (Watched By (1)
KETCHUP / HOW IT'S MADE (Watched By (1) Subscribers)
LEFT MUSIC (Watched By (1) Subscribers)
SHAUN THE SHEEP: SERIES 2: CHEETAH CHEATER (Watched By (1) Subscribers)
SPIDER-MAN (Watched By (1) Subscribers)
THE BIG BANG THEORY: IT ALL STARTED WITH A BIG BANG (Watched By (1) Subscribers)
THE BOY OFFICIAL TRAILER #1 (2016) LAUREN COHAN HORROR MOVIE HD (Watched By
THE BIG BANG THEORY: BEST OF ALL SEASONS (Watched By (1) Subscribers)
THE SIMPSONS - BART'S BIRTH (Watched By (1) Subscribers)
TOM AND JERRY: 63 EPISODE - THE FLYING CAT (1952) (Watched By (1) Subscribers)
TOM AND JERRY: 64 EPISODE - THE DUCK DOCTOR (1952) (Watched By (1) Subscribers)
TRANSFORMERS: AGE OF EXTINCTION (Watched By (1) Subscribers)
UNDERWORLD: EVOLUTION (Watched By (1) Subscribers)

SAUNA ADEEL SADIQ FEAT BLOODLINE (OFFICIAL HD MUSIC VIDEO) (Watched By (1))
ADEL - HELLO (Watched By (1) Subscribers)
ADEL - SOMEONE LIKE YOU (GUSTONBURY 2016) (Watched By (1) Subscribers)
FAMILY GUY - THE SIMPSONS (24 HOUR) L.A.'VE (Watched By (1) Subscribers)
HOW I MET YOUR MOTHER SEASON 9 EPISODE 5 & 7 & 9 - THE POKER GAME (Watched B
INDIANA JONES AND THE TEMPLE OF DOOM (1984) - HARRISON FORD, KATE CAPSHAW
INSIDE THE MIND OF A FEMINIST - A NATIONAL GEOGRAPHIC SPECIAL (Watched By (1)
KETCHUP / HOW IT'S MADE (Watched By (1) Subscribers)
LEFT MUSIC (Watched By (1) Subscribers)
SHAUN THE SHEEP: SERIES 2: CHEETAH CHEATER (Watched By (1) Subscribers)
SPIDER-MAN (Watched By (1) Subscribers)
THE BIG BANG THEORY: IT ALL STARTED WITH A BIG BANG (Watched By (1) Subscribers)
THE BOY OFFICIAL TRAILER #1 (2016) LAUREN COHAN HORROR MOVIE HD (Watched By
THE BIG BANG THEORY: BEST OF ALL SEASONS (Watched By (1) Subscribers)
THE SIMPSONS - BART'S BIRTH (Watched By (1) Subscribers)
TOM AND JERRY: 63 EPISODE - THE FLYING CAT (1952) (Watched By (1) Subscribers)
TOM AND JERRY: 64 EPISODE - THE DUCK DOCTOR (1952) (Watched By (1) Subscribers)
TRANSFORMERS: AGE OF EXTINCTION (Watched By (1) Subscribers)
UNDERWORLD: EVOLUTION (Watched By (1) Subscribers)

<http://www.bbc.co.uk/iplayer/episode/b04ybnjn>

| | |
|------------------------|---|
| DESCRIPTION | THE SWINDLER PERDIGAN USES HIS HYPOCOTIC POWERS TO PUT ROBIN UN |
| GENRE | CHILDREN |
| LENGTH | 12MINUTE |
| PRODUCTION YEAR | 2015 |

6.4.c. Programme Belongs to BBC-iPlayer Web Site

[GROUP SUGGESTIONS](#)

THE AVENGERS - HULK SMASH
MINIONS THE AVENGERS ADVENTURES (SUPERHEROES DOL) - FUNNY CARTOON (HD)
AVENGERS 2: FULL HD MOVIE 2015
MARVEL'S THE AVENGERS- AGE OF ULTRON 2015 - TRAILER #3
ROBIN HOOD: COMHSTRI SA CHOLLE: CEANNSACADH NA COLLE
ROBIN HOOD: COMHSTRI SA CHOLLE: LEASANAI RA-OGHAL
ROBIN HOOD: COMHSTRI SA CHOLLE: AN SUAIMIALAICHE
ROBIN HOOD: COMHSTRI SA CHOLLE: AN LEHCHE-ERÁIGE
ANIN (SA BRANCH) - ROBIN HOOD TAX
ROBIN HOOD: COMHSTRI SA CHOLLE: AM FEAR FIADAIN
ROBIN HOOD: COMHSTRI SA CHOLLE: TÁIR NA TÁSGADH
ROBIN HOOD: COMHSTRI SA CHOLLE: MAIGHSTR NAH PUPAD
ROBIN HOOD: COMHSTRI SA CHOLLE: CLADHEAHIH NAN RA-SHREAN
THE AVENGERS
AVENGERS: AGE OF ULTRON
THE STORY OF BEAN
ROBIN HOOD: COMHSTRI SA CHOLLE: CROIS DHÁ'BALT NA CARAD
ROBIN HOOD: COMHSTRI SA CHOLLE: AN T-ÁIR NEO-FHACSNIEACH
AVENGERS: INFINITY WAR CAPTAIN MARVEL BREAKDOWN
AVENGERS: AGE OF ULTRON - CLIP IRON MAN VS. ULTRON (HD) - SUBTITULADO POR
AVENGERS: AGE OF ULTRON TRAILERS 5

[YOUR FAVOURITE](#)

THE AVENGERS - HULK SMASH
MINIONS THE AVENGERS ADVENTURES (SUPERHEROES DOL) - FUNNY CARTOON (HD)
AVENGERS 2: FULL HD MOVIE 2015
MARVEL'S THE AVENGERS- AGE OF ULTRON 2015 - TRAILER #3
ROBIN HOOD: COMHSTRI SA CHOLLE: CEANNSACADH NA COLLE
ROBIN HOOD: COMHSTRI SA CHOLLE: LEASANAI RA-OGHAL
ROBIN HOOD: COMHSTRI SA CHOLLE: AN SUAIMIALAICHE
ROBIN HOOD: COMHSTRI SA CHOLLE: AN LEHCHE-ERÁIGE
ANIN (SA BRANCH) - ROBIN HOOD TAX
ROBIN HOOD: COMHSTRI SA CHOLLE: AM FEAR FIADAIN
ROBIN HOOD: COMHSTRI SA CHOLLE: TÁIR NA TÁSGADH
ROBIN HOOD: COMHSTRI SA CHOLLE: MAIGHSTR NAH PUPAD
ROBIN HOOD: COMHSTRI SA CHOLLE: CLADHEAHIH NAN RA-SHREAN
THE AVENGERS
AVENGERS: AGE OF ULTRON
THE STORY OF BEAN
ROBIN HOOD: COMHSTRI SA CHOLLE: CROIS DHÁ'BALT NA CARAD
ROBIN HOOD: COMHSTRI SA CHOLLE: AN T-ÁIR NEO-FHACSNIEACH
AVENGERS: INFINITY WAR CAPTAIN MARVEL BREAKDOWN
AVENGERS: AGE OF ULTRON - CLIP IRON MAN VS. ULTRON (HD) - SUBTITULADO POR
AVENGERS: AGE OF ULTRON TRAILERS 5

[GROUP SUGGESTIONS](#)

SAUNA ADEEL SADIQ FEAT BLOODLINE (OFFICIAL HD MUSIC VIDEO) (Watched By (1))
ADEL - HELLO (Watched By (1) Subscribers)
ADEL - SOMEONE LIKE YOU (GUSTONBURY 2016) (Watched By (1) Subscribers)
FAMILY GUY - THE SIMPSONS (24 HOUR) L.A.'VE (Watched By (1) Subscribers)
HOW I MET YOUR MOTHER SEASON 9 EPISODE 5 & 7 & 9 - THE POKER GAME (Watched B
INDIANA JONES AND THE TEMPLE OF DOOM (1984) - HARRISON FORD, KATE CAPSHAW
INSIDE THE MIND OF A FEMINIST - A NATIONAL GEOGRAPHIC SPECIAL (Watched By (1)
KETCHUP / HOW IT'S MADE (Watched By (1) Subscribers)
LEFT MUSIC (Watched By (1) Subscribers)
SHAUN THE SHEEP: SERIES 2: CHEETAH CHEATER (Watched By (1) Subscribers)
SPIDER-MAN (Watched By (1) Subscribers)
THE BIG BANG THEORY: IT ALL STARTED WITH A BIG BANG (Watched By (1) Subscribers)
THE BOY OFFICIAL TRAILER #1 (2016) LAUREN COHAN HORROR MOVIE HD (Watched By
THE BIG BANG THEORY: BEST OF ALL SEASONS (Watched By (1) Subscribers)
THE SIMPSONS - BART'S BIRTH (Watched By (1) Subscribers)
TOM AND JERRY: 63 EPISODE - THE FLYING CAT (1952) (Watched By (1) Subscribers)
TOM AND JERRY: 64 EPISODE - THE DUCK DOCTOR (1952) (Watched By (1) Subscribers)
TRANSFORMERS: AGE OF EXTINCTION (Watched By (1) Subscribers)
UNDERWORLD: EVOLUTION (Watched By (1) Subscribers)

SAUNA ADEEL SADIQ FEAT BLOODLINE (OFFICIAL HD MUSIC VIDEO) (Watched By (1))
ADEL - HELLO (Watched By (1) Subscribers)
ADEL - SOMEONE LIKE YOU (GUSTONBURY 2016) (Watched By (1) Subscribers)
FAMILY GUY - THE SIMPSONS (24 HOUR) L.A.'VE (Watched By (1) Subscribers)
HOW I MET YOUR MOTHER SEASON 9 EPISODE 5 & 7 & 9 - THE POKER GAME (Watched B
INDIANA JONES AND THE TEMPLE OF DOOM (1984) - HARRISON FORD, KATE CAPSHAW
INSIDE THE MIND OF A FEMINIST - A NATIONAL GEOGRAPHIC SPECIAL (Watched By (1)
KETCHUP / HOW IT'S MADE (Watched By (1) Subscribers)
LEFT MUSIC (Watched By (1) Subscribers)
SHAUN THE SHEEP: SERIES 2: CHEETAH CHEATER (Watched By (1) Subscribers)
SPIDER-MAN (Watched By (1) Subscribers)
THE BIG BANG THEORY: IT ALL STARTED WITH A BIG BANG (Watched By (1) Subscribers)
THE BOY OFFICIAL TRAILER #1 (2016) LAUREN COHAN HORROR MOVIE HD (Watched By
THE BIG BANG THEORY: BEST OF ALL SEASONS (Watched By (1) Subscribers)
THE SIMPSONS - BART'S BIRTH (Watched By (1) Subscribers)
TOM AND JERRY: 63 EPISODE - THE FLYING CAT (1952) (Watched By (1) Subscribers)
TOM AND JERRY: 64 EPISODE - THE DUCK DOCTOR (1952) (Watched By (1) Subscribers)
TRANSFORMERS: AGE OF EXTINCTION (Watched By (1) Subscribers)
UNDERWORLD: EVOLUTION (Watched By (1) Subscribers)

<https://www.fan.tv/movies/24428>

| | |
|------------------------|---|
| DESCRIPTION | WHEN AN UNEXPECTED ENEMY EMERGES AND THREATENS GLOBAL SA NICK FURY, DIRECTOR OF THE INTERNATIONAL BEACKEEPING AGENCY, K FINDS HIMSELF IN NEED OF A TEAM TO PULL THE WORLD BACK FROM TH SPANNING THE GLOBE & FIGHTING SUPERHEROES FROM TH |
| GENRE | SCIENCE FICTION, ACTION/ADVENTURE, ACTION/ADVENTURE |
| LENGTH | 143MINUTE |
| PRODUCTION YEAR | 2012-04-25 |

6.4.d. Programme Belongs to The Movie Database Web Site

Fig. 6.4 List of Programmes retrieved from Multiple Online Content Sources

6.4 Evaluating the Recommendation Function

As mentioned in the previous chapter, three areas are provided in the user interface to display the three lists of programmes that are recommended to a user based on the different recommendation methods that are employed. These lists are the user favourite programmes which are recommended based on the user's viewing history, group suggestions which are recommended based on the user clusters method and social media programmes which are recommended based on a contribution from social networks and the user's friends. When the user log on to the EPG, these three lists of programmes are displayed to the user as shown in Figure 6.5. By selecting any one of the programmes in those lists, the programme details, extracted from programme's metadata, are displayed on the right side of the user interface.

6.4.1 Test Scenario

In order to evaluate the recommendation process and find out the effect of social networks in enhancing the recommendation operation, the users were asked to rate the recommended programmes. Each user had to complete a test sheet which is included in Appendix VII. When a user first uses the EPG there is no viewing history but each user joins the Facebook application and selects their preferred programme categories. Therefore at this point, the EPG system is only able to retrieve information from their Facebook user account and use the selected categories to specify the clusters that the user belongs to. This means that the user is presented with two lists of recommended programmes.

Users were then asked to rate the recommended programmes in these two lists and then to create a viewing history by watching several programmes at different times of the day. Then, the user logs out the EPG and this ends the second step. The third and fourth steps include rating the recommended programmes in all three recommended lists (the third list is based on the viewing history) and continue watching several more programmes at different times.

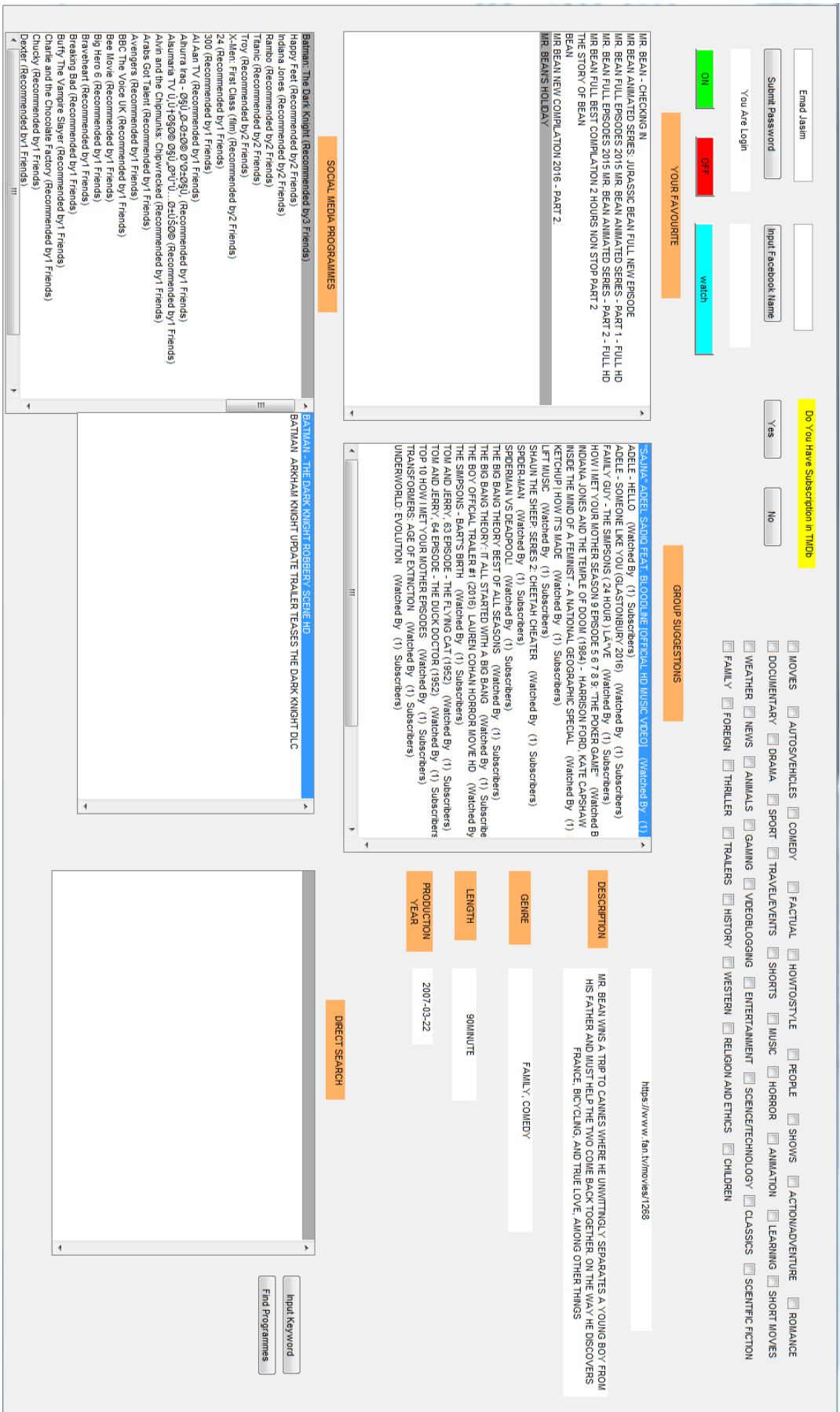


Fig. 6.5 User Interface Includes the Recommended Lists of Programmes

After each log out, the system implements all the searching and recommending operations. This includes retrieving data from Facebook, adding any new information to the user profile in the user matrix, repeating the users' clustering to accommodate newly watched programmes and searching for programmes relevant to the new watched programmes in the viewing history cell. Thereafter, the recommended list of programmes is prepared and saved in the user matrix to be ready for use when the user next logs in to the EPG.

6.4.2 Evaluation Metrics

In order to evaluate the recommendation process in any system, two standard metrics are always taken into consideration; the Precision and Recall. Additionally, Mean Absolute Error (MAE) is usually used to assess the prediction accuracy. Precision is defined as the measure of a recommendation system's ability to suggest content relevant to a user. It is represented by the ratio of relevant recommendations made to the user to the total number of recommendations made as given in Equation 6.1.

$$\textit{Precision} = \frac{\textit{Correctly Recommended Content}}{\textit{Total Recommended Content}} \dots\dots\dots\text{Equation 6.1}$$

Recall is defined as the measure of the recommendation system's ability to gather relevant content for a specific user. It is represented by the ratio of relevant recommendations made to the user to the content that are watched by the user. It can be calculated as given in Equation 6.2.

$$\textit{Recall} = \frac{\textit{Correctly Recommended Content}}{\textit{Relevant Content}} \dots\dots\dots\text{Equation 6.2}$$

However, precision is a more appropriate metric for the performance evaluation of recommendation systems rather than recall. That is because recall value increases with an increasing number of recommended programmes. Therefore, precision has been used in this research for evaluating the recommendation system performance.

On the other hand, MAE is a widely used metric to assess the difference between ratings and predictions in recommendation systems. It simply measures the deviation of predictions that are generated by the recommendation system from the real ratings which are specified by the user. MAE can be calculated as given in Equation 6.3.

$$MAE = \frac{1}{N} \sum_{i=1}^N |p_i - r_i| \quad \dots\dots\dots \text{Equation 6.3}$$

Where " N " is the total number of predictions, " p_i " is the predicted rating for item " i " and " r_i " is the user's real rating for item " i " (Al-Shamri & Bharadwaj, 2008; Barragáns-Martínez et al., 2010; Carrer-Neto et al., 2012; E. Kim et al., 2011; F. Liu & Lee, 2010; X. Liu & Aberer, 2013; Pham, Cao, Klamma, & Jarke, 2011; D. Shin et al., 2009; J. Xu et al., 2002; M. Xu et al., 2013).

6.4.3 First Step of User Trial (Cold-start Problem)

The first step is assessing the recommendations which are based solely on the users' clustering and social networks contributions. This step is considered as the solution of cold-start problem. The user needs to rate each programme in the recommended list by choosing 3, 2 or 1 which is corresponding to "like", "reasonable" or "dislike". The "like" programme means that this programme is a very interesting suggestion while "reasonable" means that this programme is an accepted suggestion but not very interesting. The "dislike" category means that this programme is not interesting and is considered as a bad suggestion for the user. Each user needs to assess up to 30 programmes in each recommended list of programmes. Both the "like" and "reasonable" categories will be considered as correct recommended programmes which will be used later in calculating the precision metric. On the other hand, the system rated the programmes that have high rank (have been watched or recommended by more than one subscriber) as rate '3' while the programmes that have been watched by only one subscriber as rate '2'. This operation has been applied for both of the recommended lists of programmes which arise from the user clusters (Group Suggestions) and social networks (Social Media Programmes) as shown in figure 6.5.

The data from the first user trial has been collected and analysed for both precision and MAE. The precision and MAE has been calculated for each user separately and then averaged for the whole group. Figure 6.6 shows the average precision value as a function of number of recommendations. The figure shows two curves; the blue curve represents the precision value versus the number of recommendations for users arising from only the clustering approach while the red curve represents the precision value for the same trial but when the recommendations are based on both clustering approach and social network contribution. Figure 6.7 shows the average MAE value as a function of number of recommendations. Similarly, the figure shows two curves; the blue curve represents the average MAE value for the user clustering approach only while the red curve represents the MAE for the same trial when both the user clustering approach and social networks have been taken into account.

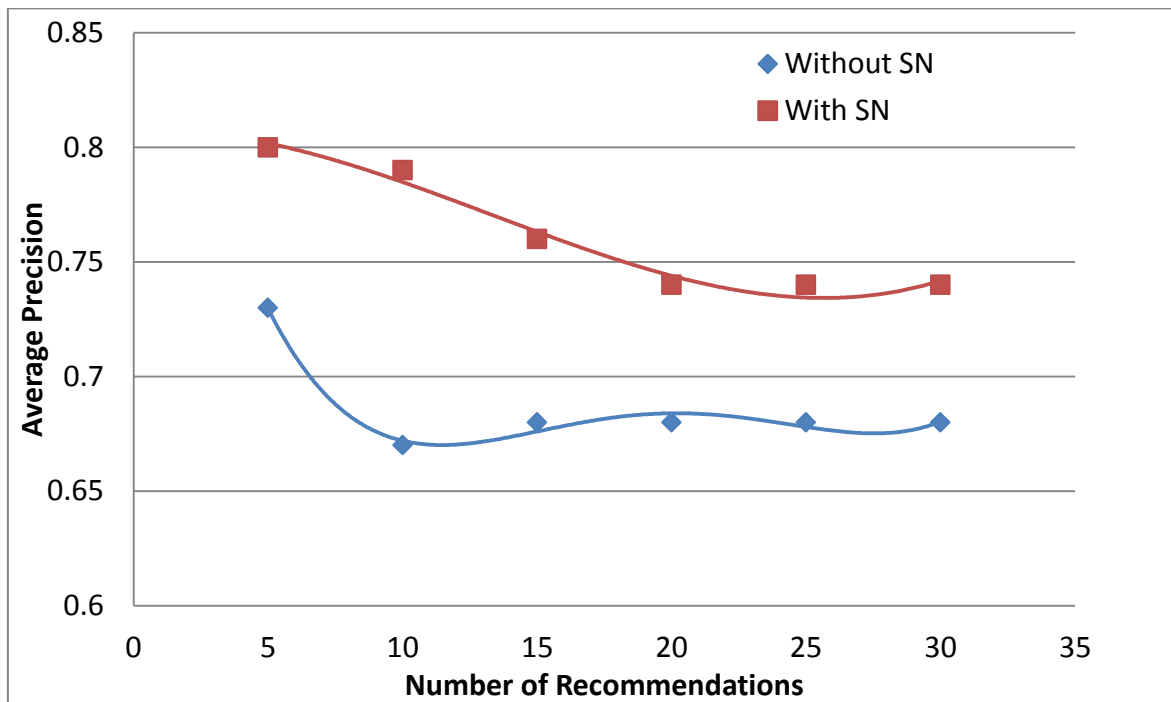


Fig. 6.6 Average Precision versus Number of Recommendations for Step One of Users' Trial

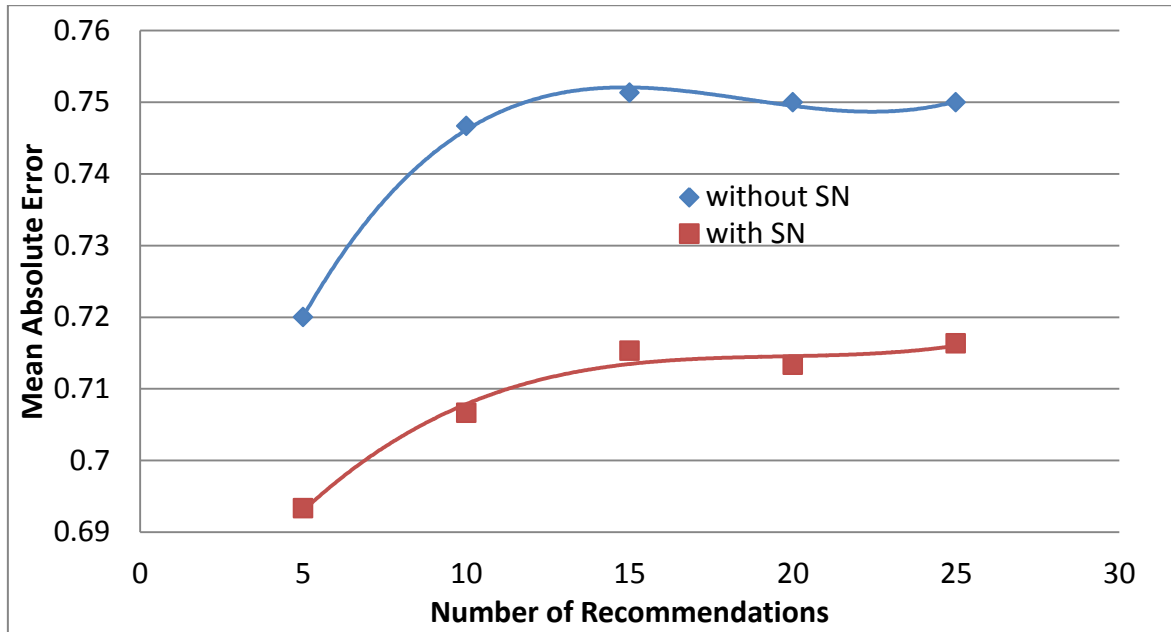


Fig. 6.7 Mean Absolute Error versus Number of Recommendations for Step One of Users' Trial

As shown in Figure 6.6, the average precision is calculated for different number of recommended programmes which varies from 5 to 30. The statistical data which was collected from this user trial demonstrates that the precision is increased when using social networks by approximately 6 to 12%. Additionally, the system presented an acceptable precision against the cold-start problem where no information was available about the user's viewing history. This can be claimed comparing with the previous research where the precision of their recommendation systems is around the resulted value in our user trials (Al-Shamri & Bharadwaj, 2008; Barragáns-Martínez et al., 2010; Carrer-Neto et al., 2012; E. Kim et al., 2011; F. Liu & Lee, 2010; X. Liu & Aberer, 2013; Pham et al., 2011; D. Shin et al., 2009; J. Xu et al., 2002; M. Xu et al., 2013). Figure 6.8 shows the average precision that calculated for different number of recommended programmes which presented by Kim and contains two cases which refer to the precision before and after the improvement of recommendation performance (E. Kim et al., 2011).

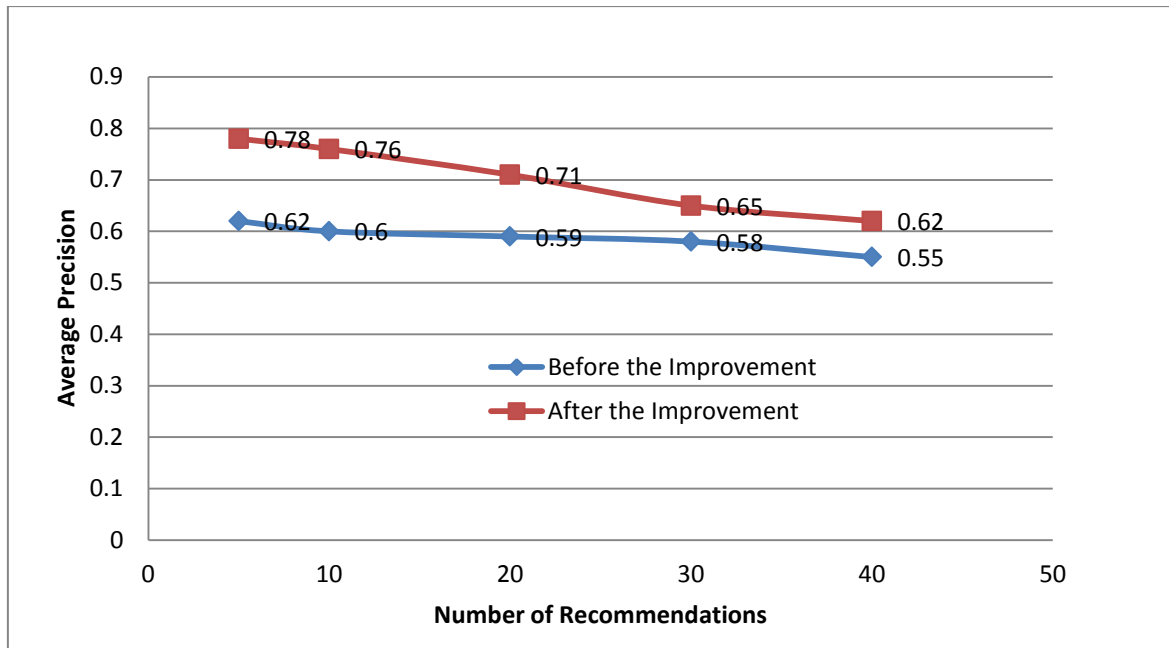


Fig. 6.8 Average Precision before and after Recommendation Performance Improvement
(E. Kim et al., 2011)

Figure 6.7 shows that the MAE of predicted recommendations decreases when the social network recommendations are included thereby improving the quality of the recommended programmes by approximately 3 to 4%.

However, it is clear that increasing the number of recommendations affects the precision and MAE values where both worsen. That is because programmes have been ranked based on their popularity. However, in the first phase, few programmes have been watched and so this limits those which have been ranked. Therefore, when the number of recommended programmes is increased, programmes which have a lower rank start to be included in the recommended list. Naturally, this affects the accuracy of the recommendation process. However, this problem should diminish by increasing the number of active users and watched programmes.

6.4.4 Second, Third and Fourth Step of Users' Trial (Viewing History)

In these steps, users need to repeat their assessment of the EPG recommended programmes and watch several more programmes in order to build up their viewing history. However, these steps now contain three recommended lists of programmes where a list of programmes based on a user's viewing history can now be added to the previous two lists.

Based on the collected data from these additional stages of user trial, precision and MAE values have been calculated for each user and then the average of all of users has been computed. Figures 6.9, 6.10 and 6.11 show the average precision value as a function of the number of recommendations for the second, third and fourth user trial respectively. In each case, two curves are plotted. The first is based on the user's viewing history and user clustering approach while the second includes the social network contribution. Figures 6.12, 6.13 and 6.14 show the MAE values against the number of recommendations for the three stages of user trial, with, similarly, two curves being plotted.

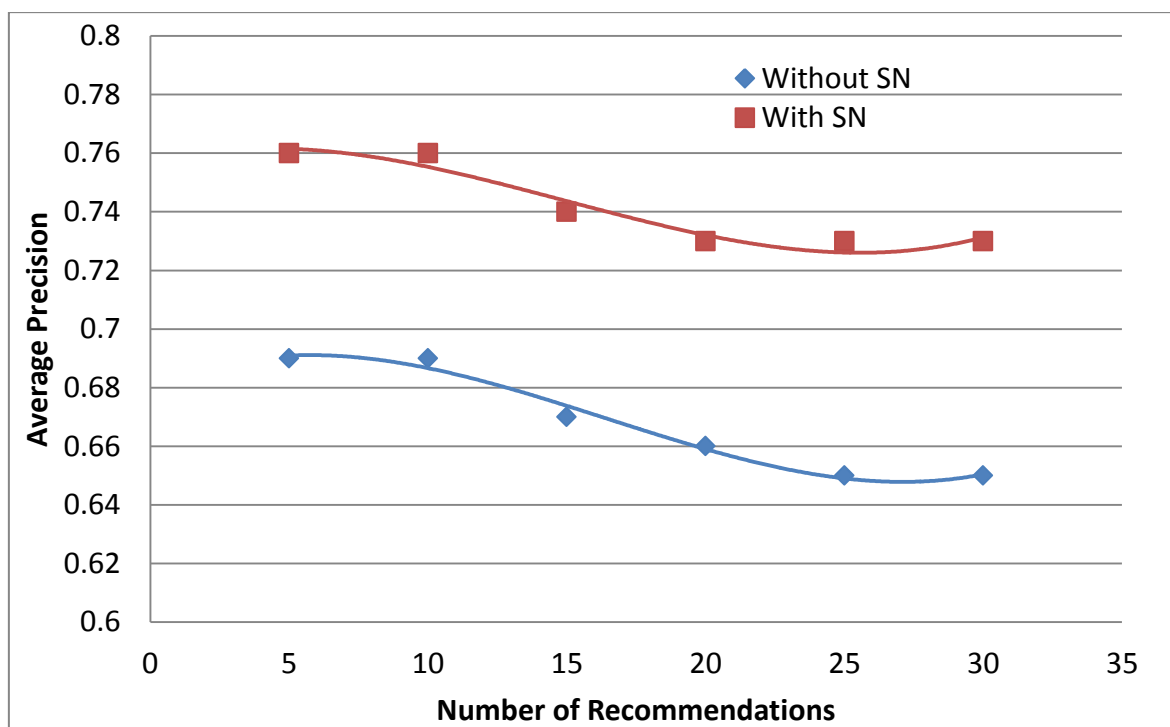


Fig. 6.9 Average Precision versus Number of Recommendations for Second Step of Users' Trial

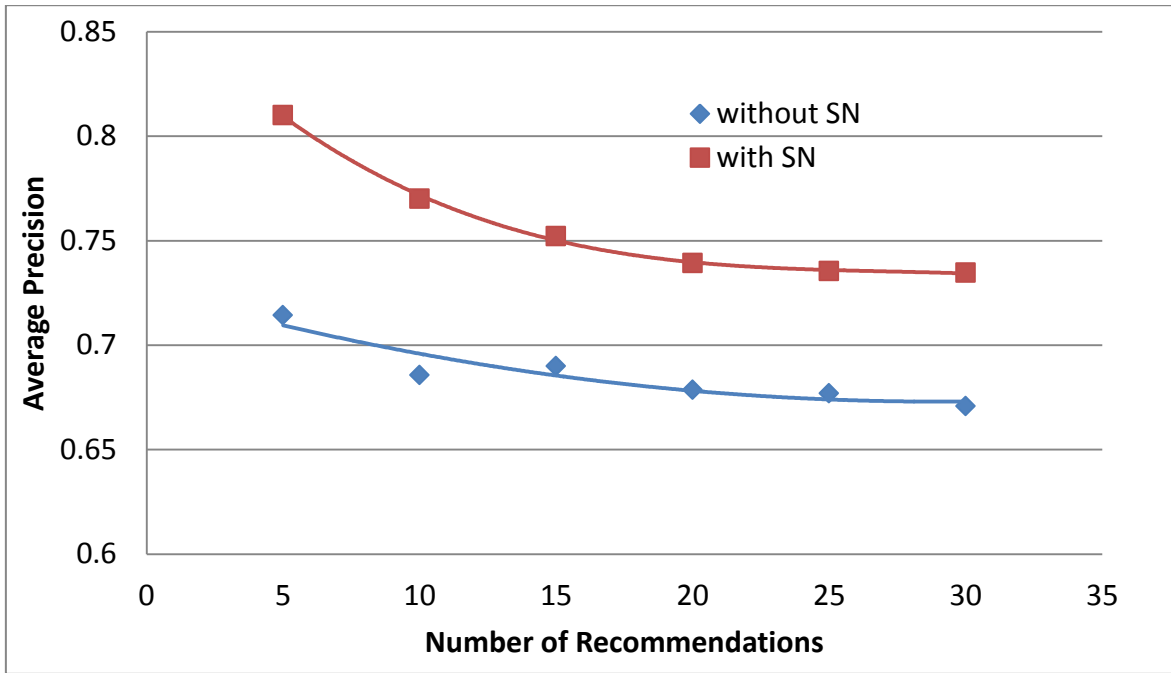


Fig. 6.10 Average Precision versus Number of Recommendations for Third Step of Users' Trial

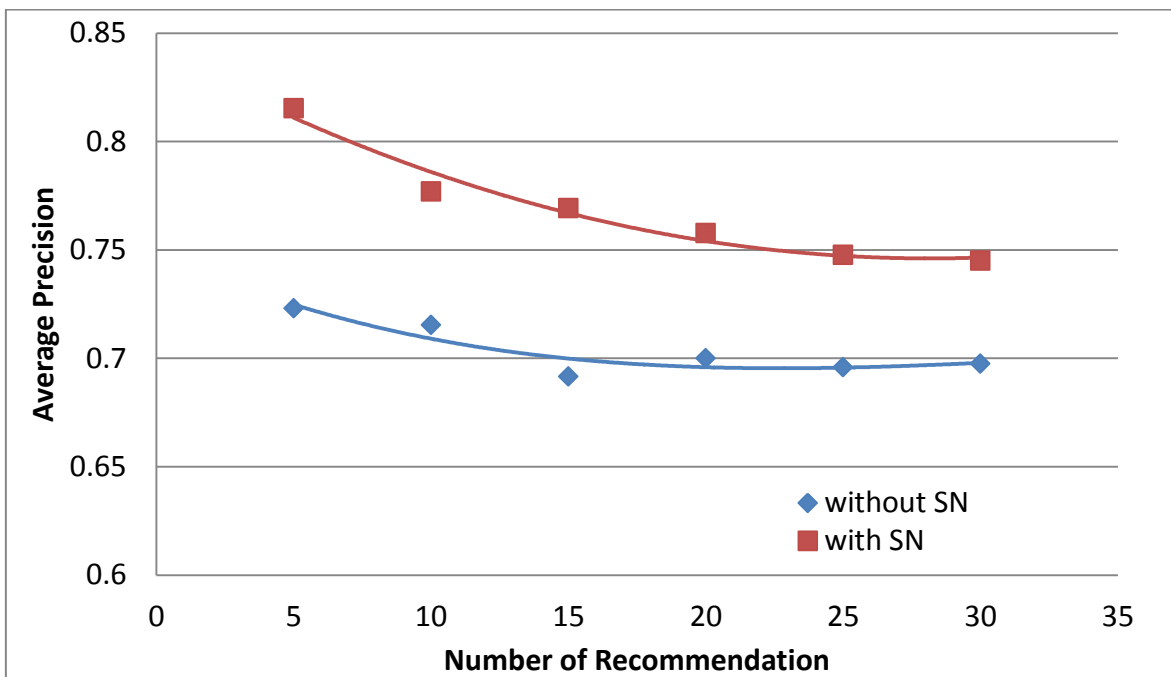


Fig. 6.11 Average Precision versus Number of Recommendations for Fourth Step of Users' Trial

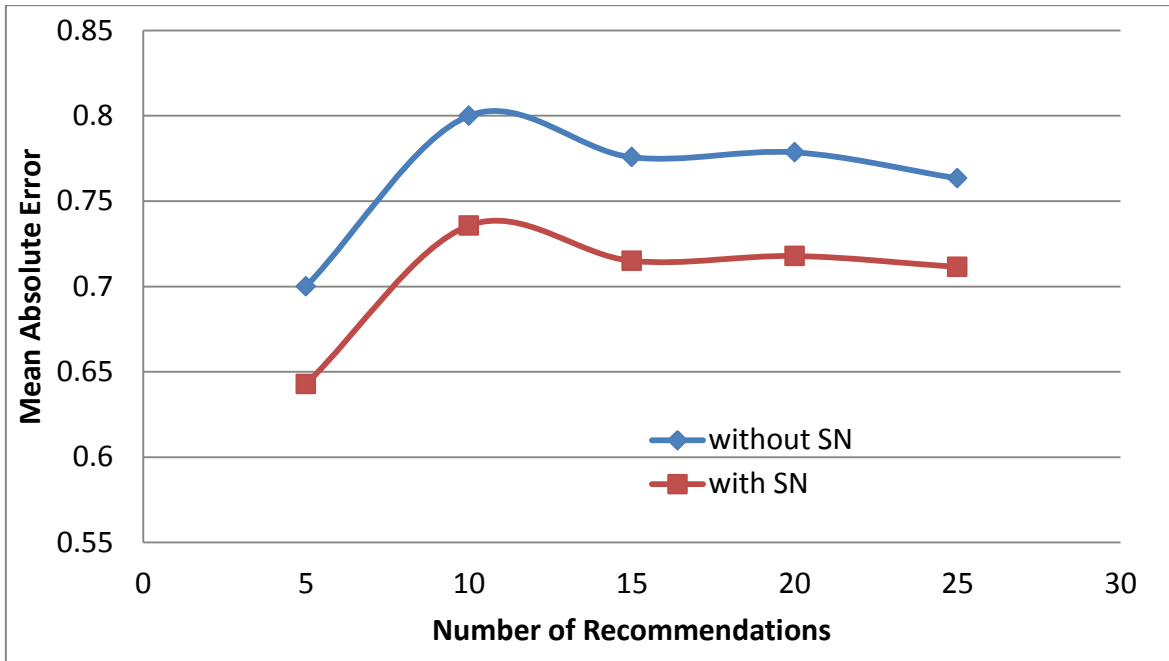


Fig. 6.12 Mean Absolute Error versus Number of Recommendations for Second Step of Users' Trial

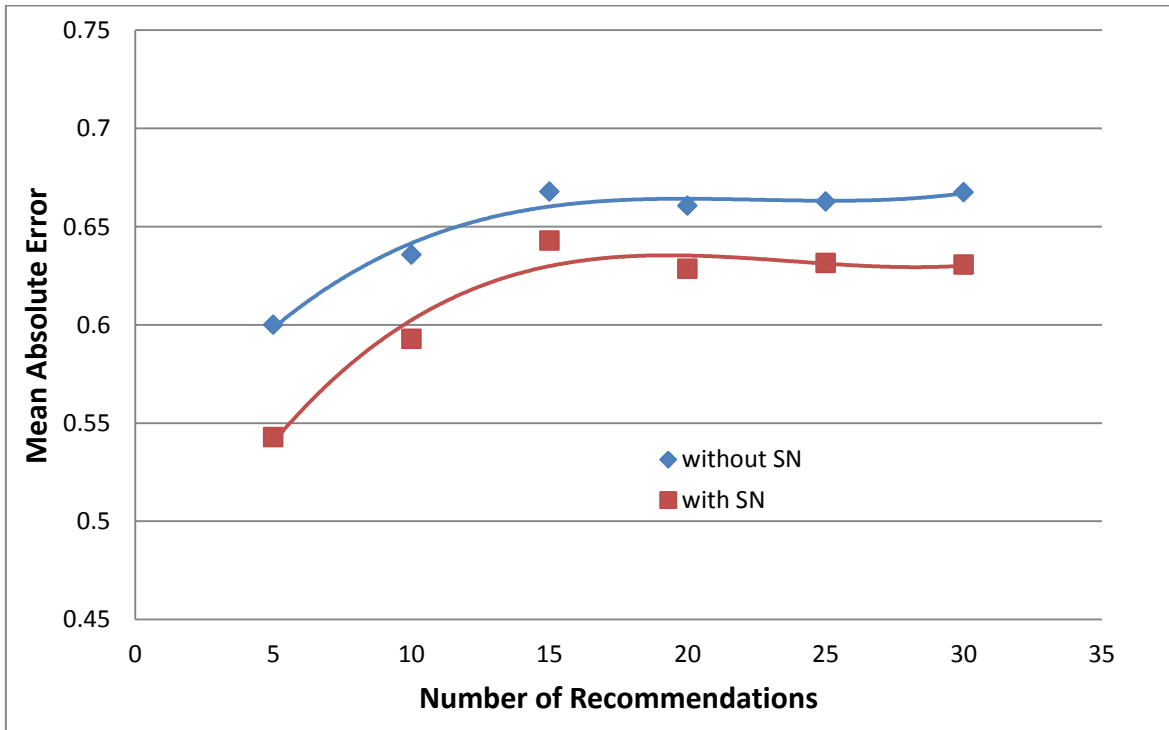


Fig. 6.13 Mean Absolute Error versus Number of Recommendations for Third Step of Users' Trial

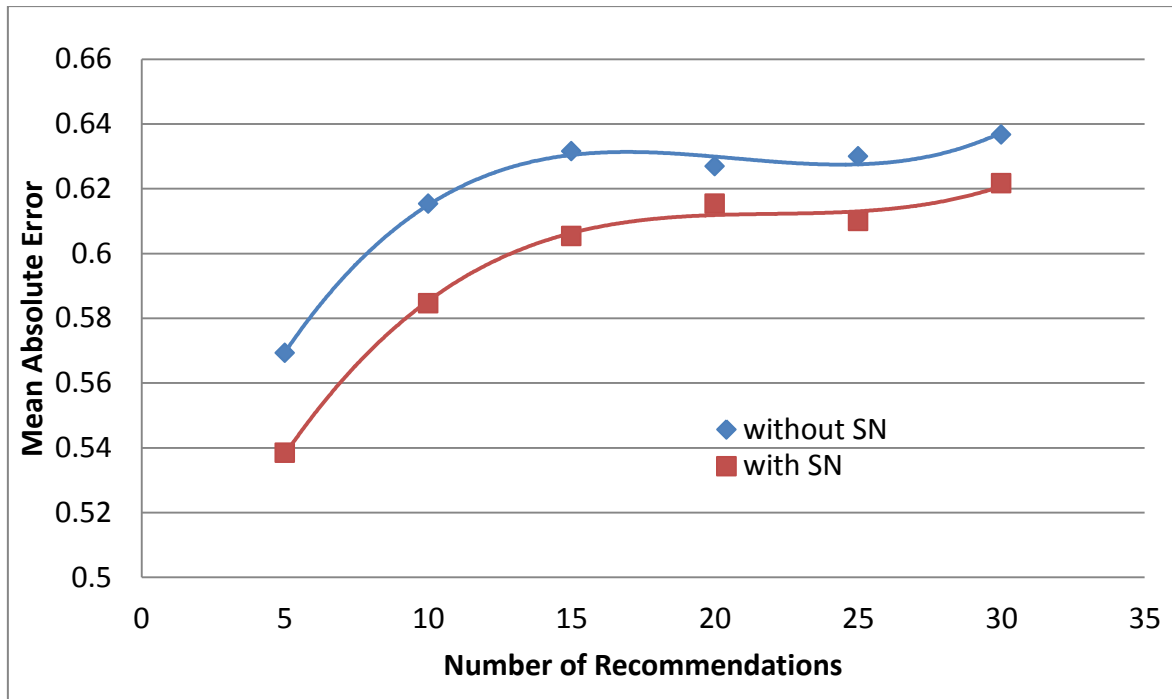


Fig. 6.14 Mean Absolute Error versus Number of Recommendations for Fourth Step of Users' Trial

Once again, the use of retrieved data from social networks improves the overall precision, showing an improvement of between 7% and 8% in the second stage, 7% and 10% in the third, and 5% and 10% in the fourth. Similarly, the MAE decreases with the use of social networks from 4 to 6% in the second stage, 3 to 6% in the third, and 1 to 3% in the fourth. As from the first stage of user trial, increasing the number of recommendations decreases the precision value and increases the MAE value because more lower rated programmes are included in the recommendation process.

Comparing figure 6.6 with figure 6.9, it is clear that the precision value of the second step (when the viewing history has been taken into account) is less than its value in the first step (without viewing history). However, it was expected that the precision would increase when the viewing history was included in the recommendation process. This didn't happen because, in the second stage, the users' viewing history contains only a few watched programmes which means that the system's recommendations are based on those programmes which makes it very difficult to predict the user's viewing pattern with so

little data. This affects the accuracy of the recommendation process based on the viewing history and thus the precision value drops when those statistics are included in the precision calculations.

Similarly, the MAE suffers from the same problem, where the MAE of the second step of user trial increases compared to the value from the first step. This can be seen when comparing figure 6.7 and figure 6.12. However, when the users' viewing history increases, this problem disappears. That is clear in figure 6.10 and 6.11 where the precision value is increased compared to figure 6.6. In the same way, by comparing figures 6.13 and 6.14 with figure 6.7, it can be seen that the MAE decreases again to 0.54 and 0.53 respectively.

In order to highlight the effect of user viewing history on the overall performance of the recommendation system, figure 6.15 shows the average precision value as a function of each stage of user trial. The average precision of each step has been calculated by summing the precision values of each step (as a function of the number of recommendations) and dividing that sum by the number of points plotted on the graph. It is clear that with an increasing viewing history, the average precision value increases and in the fourth step exceeds its value in the first step. However, it is expected to obtain more improvement in the precision value with continuing increase the users' viewing history. Moreover, it is also clear that integrating information retrieved from Facebook into the recommendation process improves the average precision value by 6.5 to 7.5%.

Similarly, figure 6.16 shows the average MAE as a function of each stage of user trial. The figure shows that including user viewing history has not improved the performance of the recommender system in the second step, however, with an increasing viewing history, the system offered a better performance in the third and fourth stages with the MAE value decreasing. Once again, the social networks have a positive effect on the system performance where the improvement in MAE value varies between 2 to 6%.

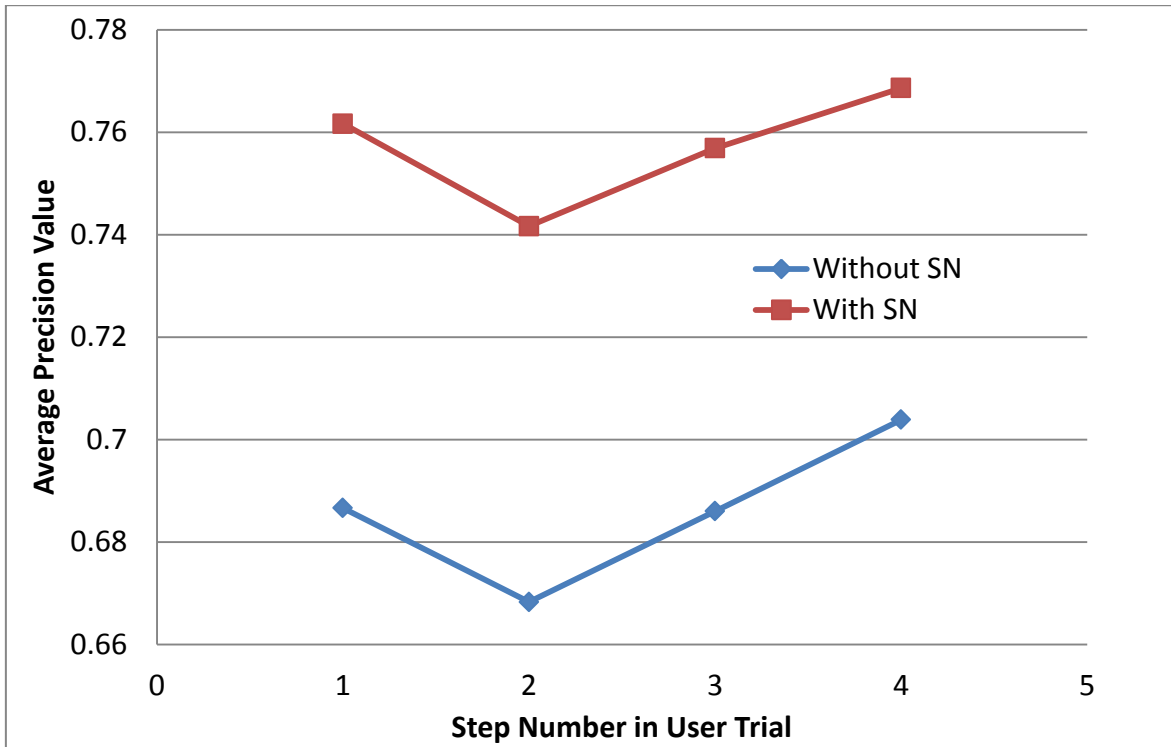


Fig. 6.15 Average Precision Value as a Function of User Trial’s Step

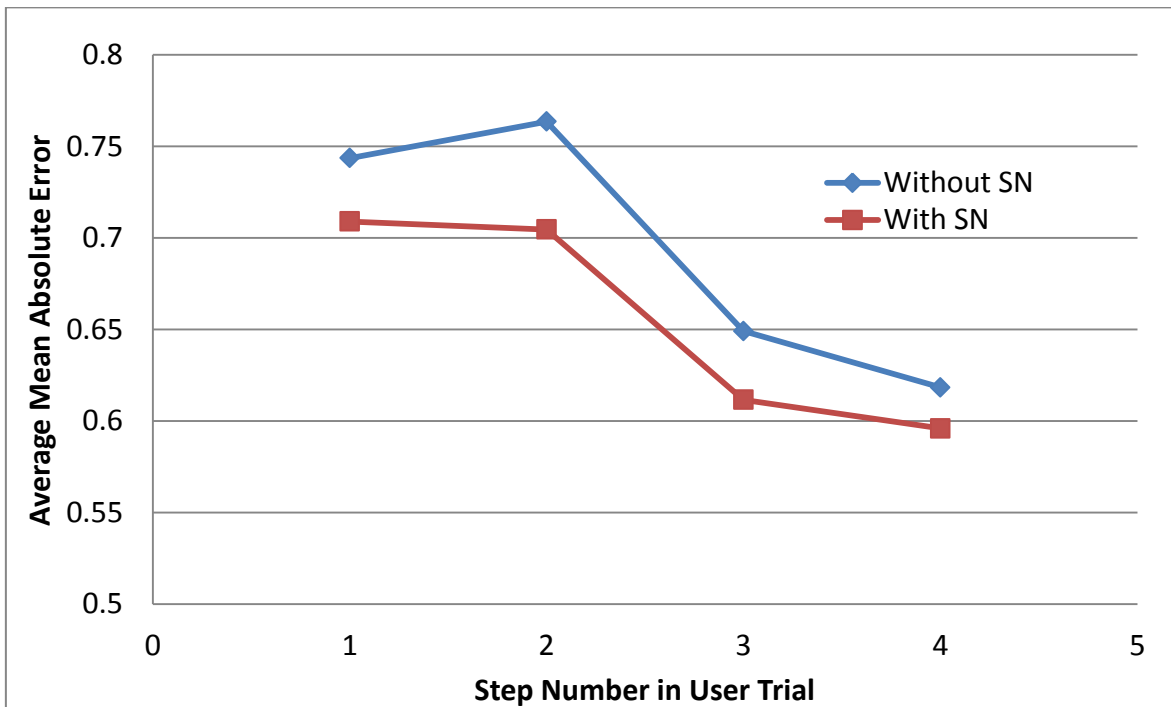
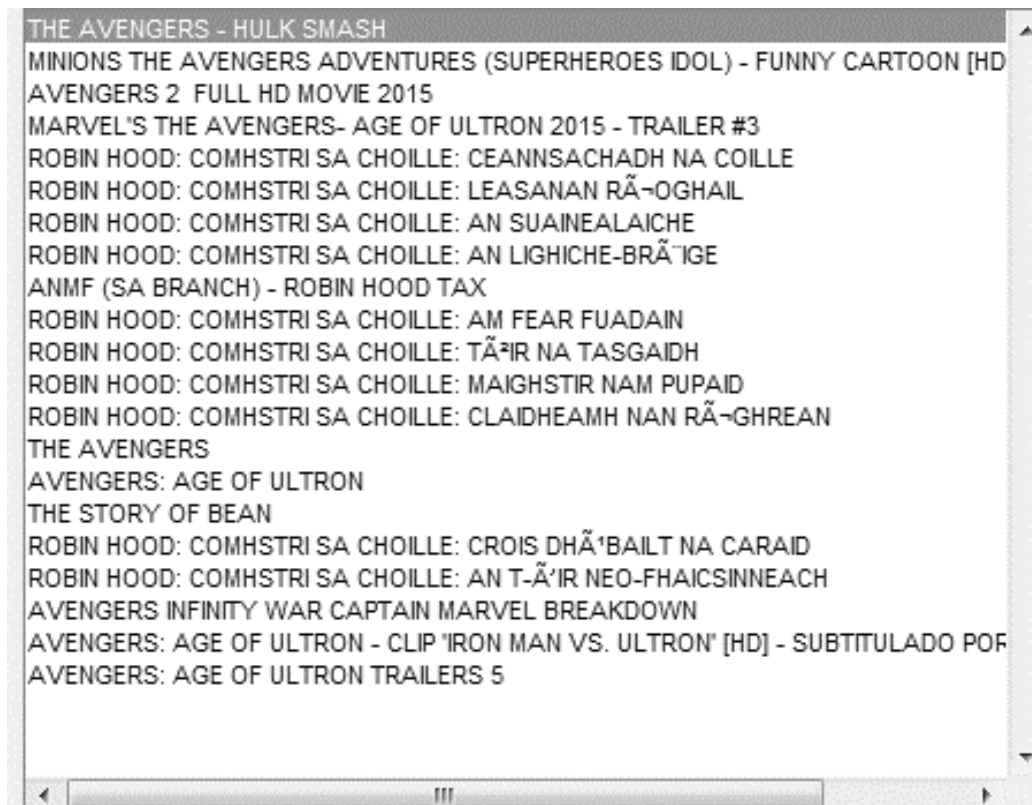


Fig. 6.16 Average Mean Absolute Error Value as a Function of User Trial’s Step

6.5 Filtering Operation

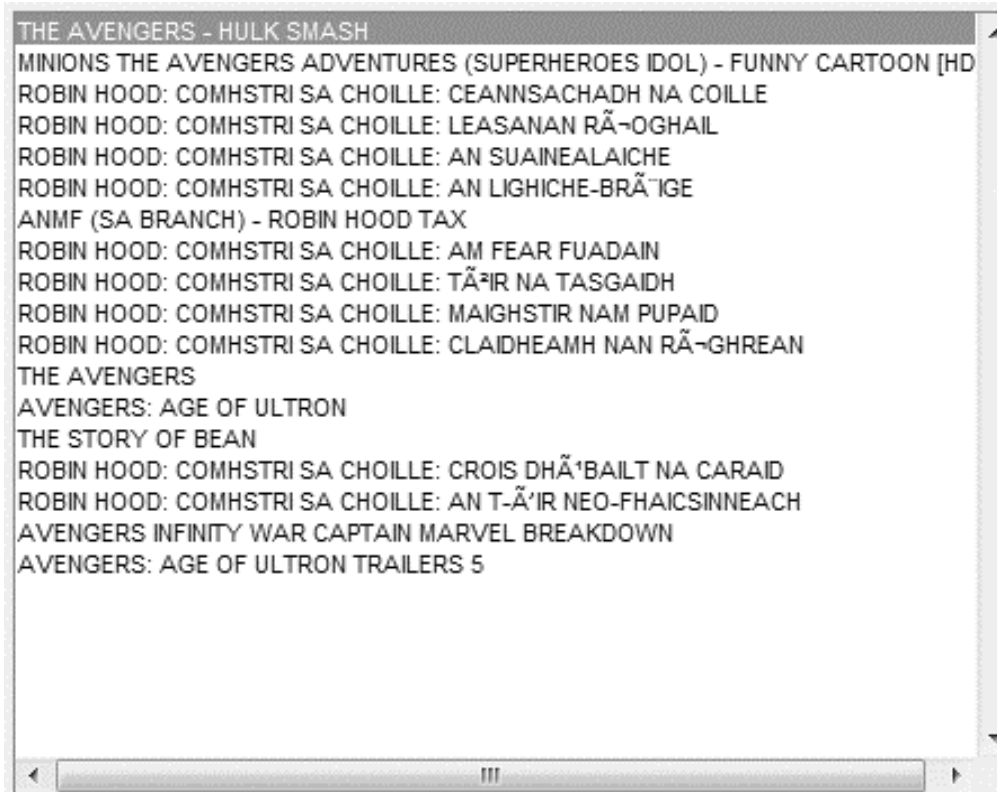
All three lists of recommended programmes are subjected to a filtering operation in order to ensure that programmes match the user's device and network connectivity. For the prototype system, information relevant to the user's device and available network bandwidth is emulated through the use of a text file. Figure 6.17 shows a recommended list of programmes that have been subjected to this filtering stage. In this case, three different values of bandwidth are chosen 10, 5 and 1.2 Mbps. As shown in figure 6.16.a, there is an initial recommended list comprising 21 programmes when the available network bandwidth is 10 Mbps.



6.17.a Available Bandwidth = 10 Mbps

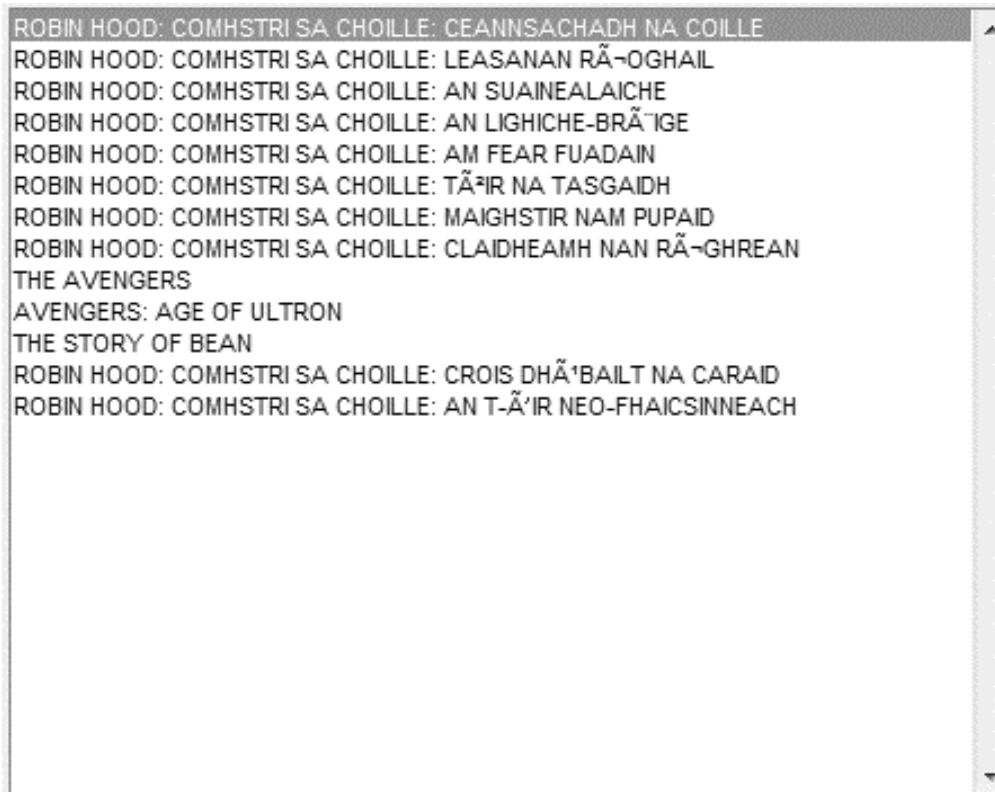
However, when the available network bandwidth is changed to 5 Mbps, figure 6.17.b shows the same recommended list of programmes which has now been reduced to 18 programmes. The reason for this is that there were three programmes within the list that

were defined as 3D movies which requires a higher bandwidth to play. When the bandwidth is further reduced to 1.2 Mbps figure 6.17.c shows the same list of recommended programmes has been reduced to 13 programmes. Here the system filtered out all High Definition and 3D movies from the list.



6.17.b Available Bandwidth = 5 Mbps

Consequently, it is clear that the system has used the user's viewing environmental features in recommending the programmes that match the current network connectivity and the final list of programmes has been subjected to a filtering operation based on the supported information. Because the capabilities of different viewing devices are also represented by a category number, as shown in the previous chapter in table 5.2, the filtering process works in exactly the same way when adapting the list of recommended programmes to match the user's device. This is illustrated in figure 6.18 where the same initial list of recommended programmes has now been filtered based on the fact that the user is viewing content on a smartphone. Here 3D programme content has been removed as it is deemed that this category of smartphone cannot display that type of content, even though there is sufficient network bandwidth available to stream it.



6.17.c Available Bandwidth = 1.2 Mbps

Fig. 6.17 Filtered Recommended List for Three Different Bandwidth Values

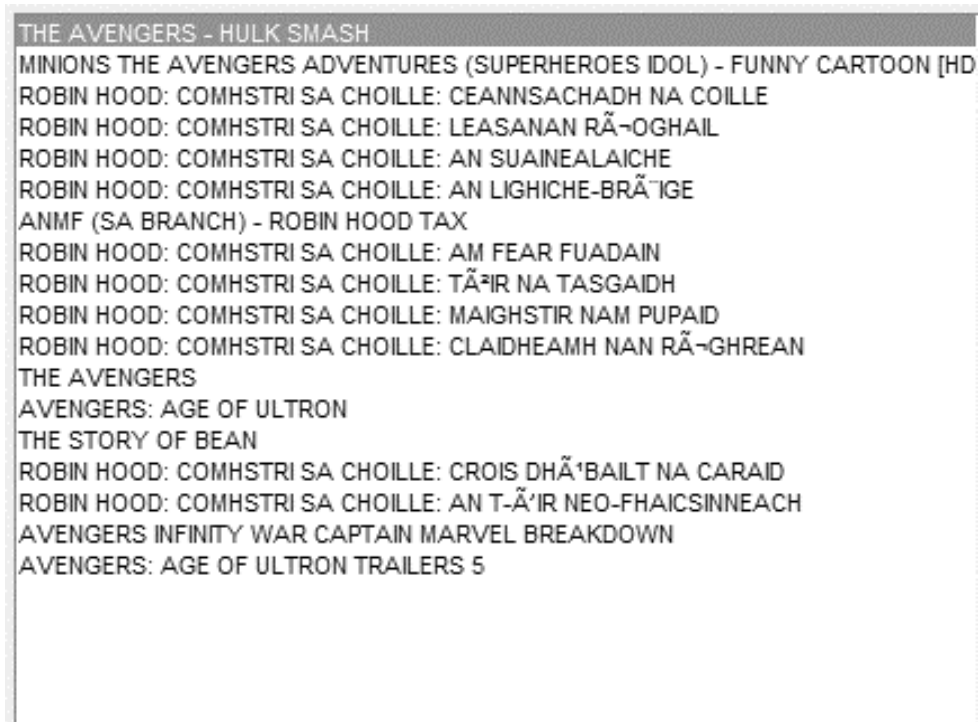


Fig. 6.18 Filtered Recommended List based on User's Device

6.6 Direct Search

Another feature that has been added to the EPG user interface is the direct search. This feature allows the user to perform a keyword search over all the connected online content sources. The user inputs the search keyword using the click “Input Keyword” button as shown in figure 6.19 where a new window will be open to enter the search keyword which could be the title or part thereof, or the category of the programme that the user wants to watch. After entering the keyword, the user needs to press “Submit Query” to send this keyword to the search agent leader and from there it will be sent to the different search engines which correspond to the connected online content sources. Thereafter, by pressing the “Find Programmes” button, a list of programmes will be displayed by the EPG as shown in figure 6.20. These results have been validated through comparing them with the search results that are obtained using the online content source directly where it was found that are same in both cases (with using the designed EPG and using the online content source itself).

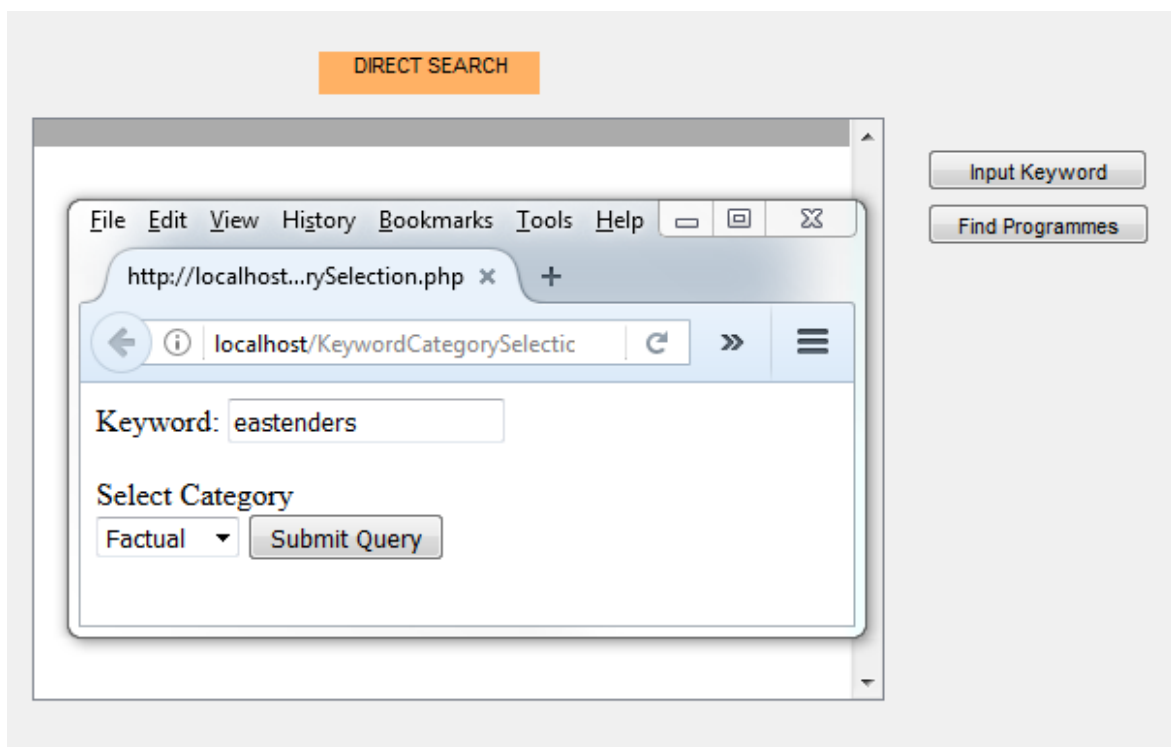


Fig. 6.19 Entering the Search Keyword in the Direct Search Operation

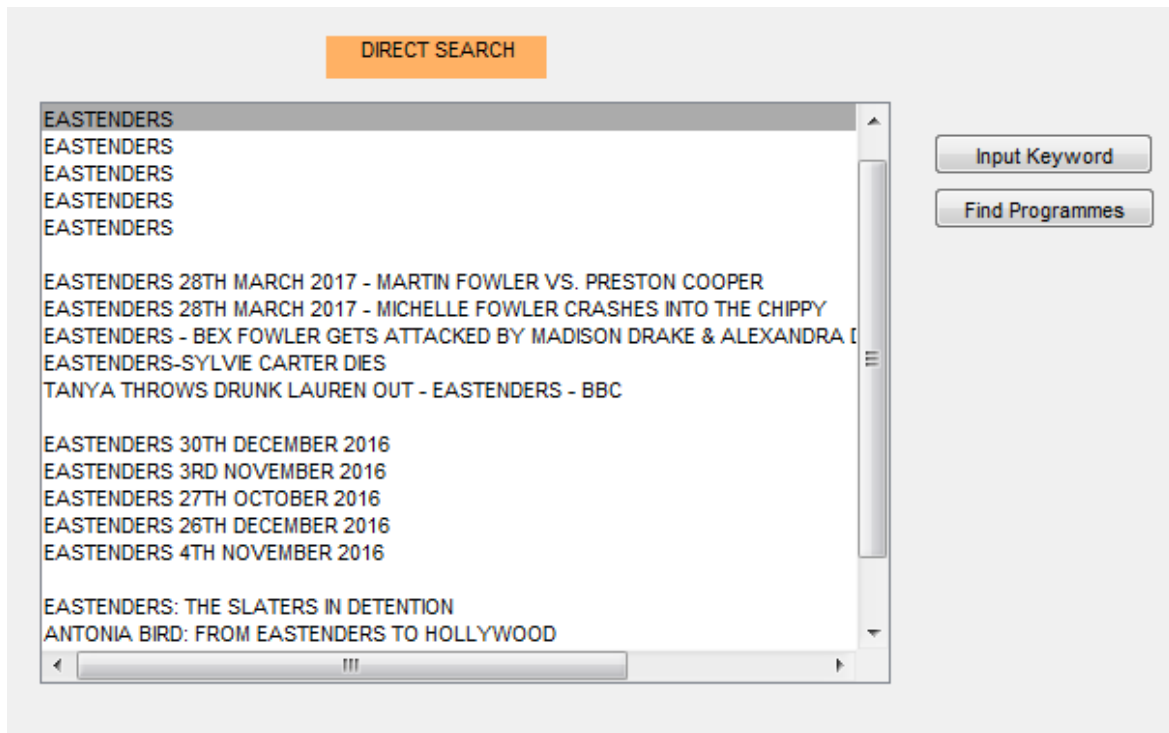


Fig. 6.20 Resulted List of Programmes from the Direct Search Operation

This list of programmes includes programmes from all of the connected online content sources. For test purpose, each group of programmes that was retrieved from a particular online content source has been separated from the other sources' programmes by a blank line. As an example, in figure 6.20, the first group belongs to BBC-iPlayer while the second group belongs to YouTube. The third group belongs to Dailymotion and finally the fourth group belongs to The Movie Database. By clicking on any one of those programmes, the details of the programme will be displayed on the user interface as shown in figure 6.21. If the user likes this programme, then by pressing the “watch” button, the user can directly watch the programme through a browser page and the details of the programme will be added to their viewing history for the hour within which it was watched. Figure 6.22 shows an example selected programme obtained after pressing the “Watch” button and it is clear from the web address that it is the same as in figure 6.21.

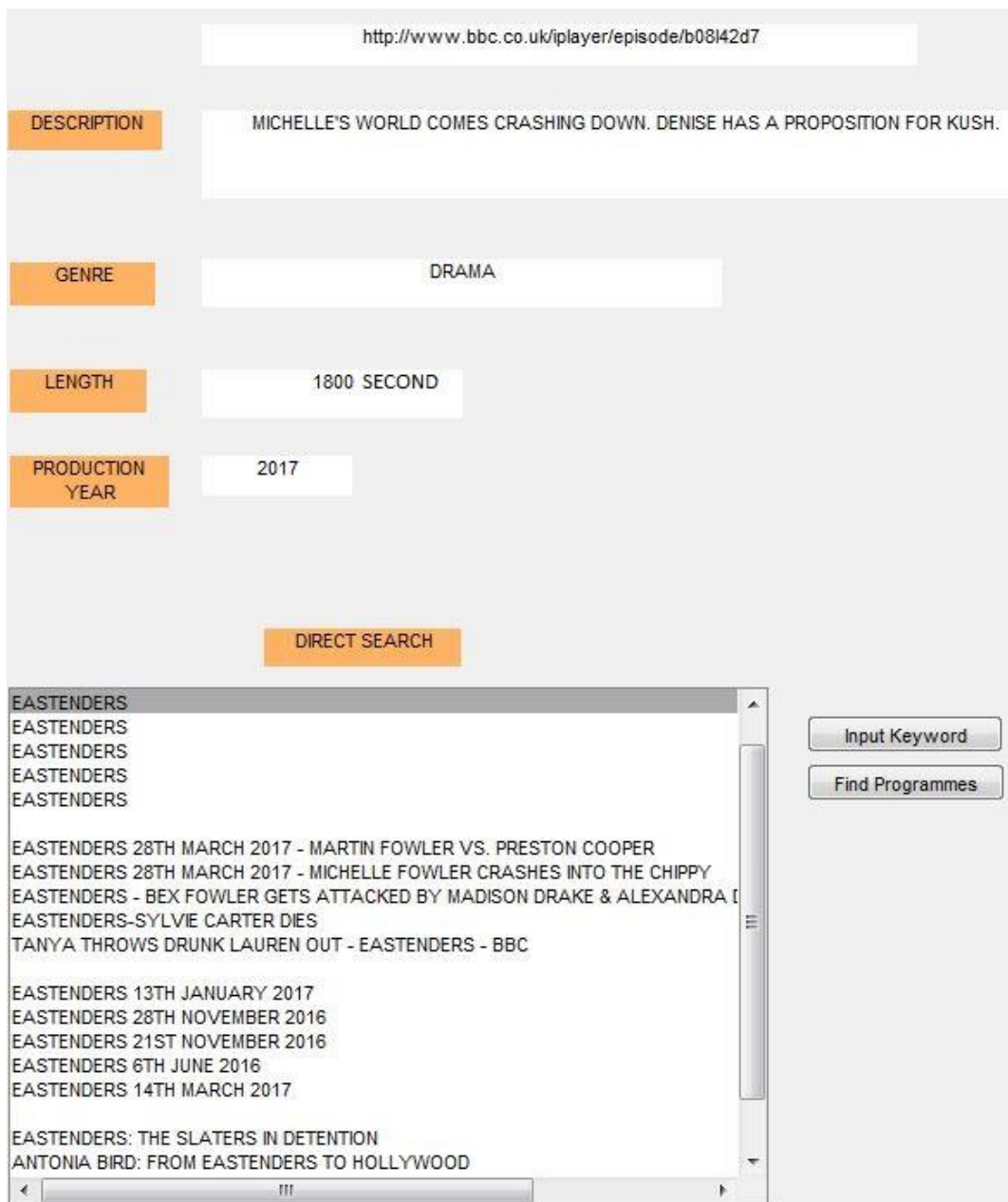


Fig. 6.21 The Details of Selected Programme in Direct Search Operation

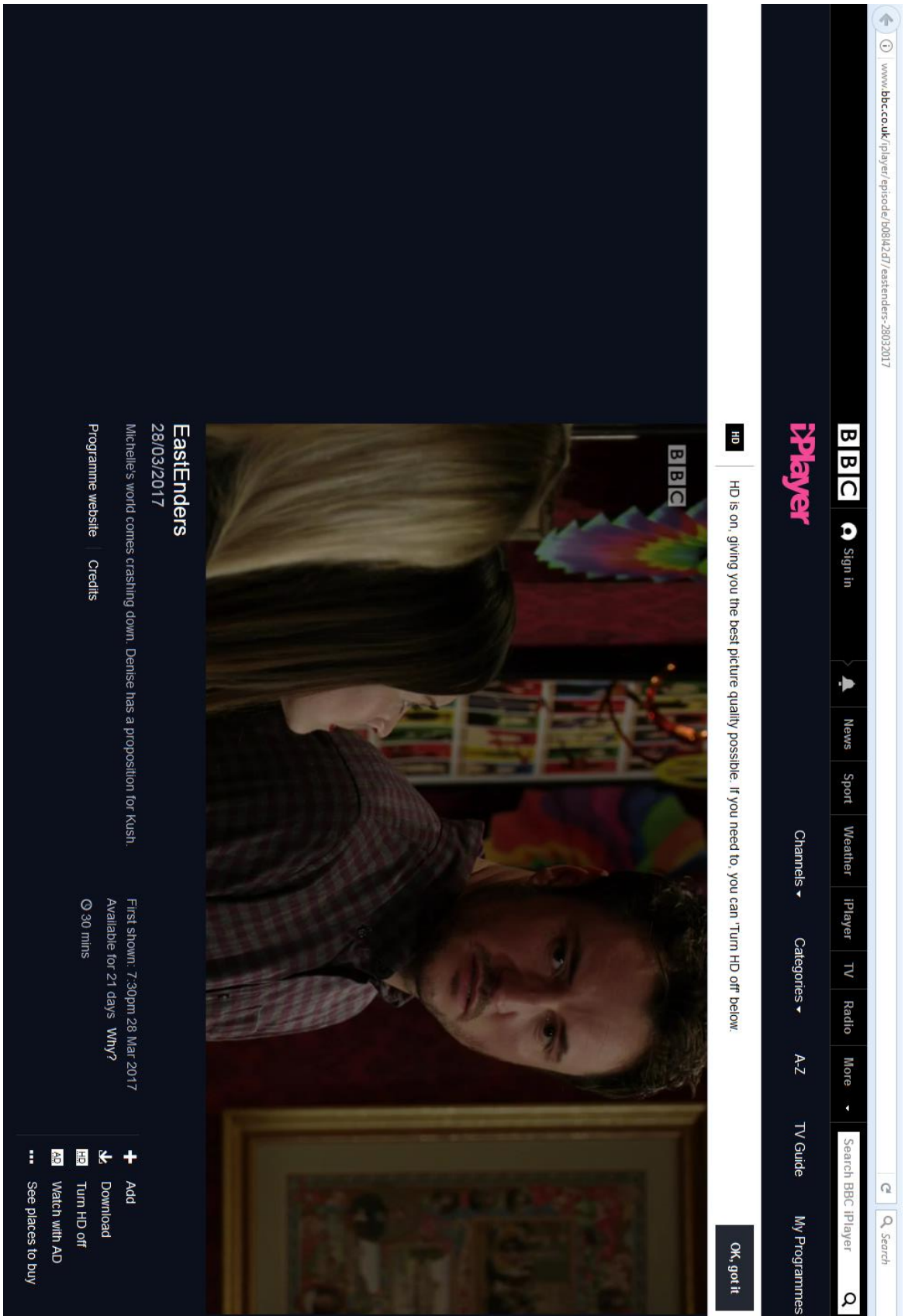


Fig. 6.22 Watching a Selected Programme from the displayed List

6.7 Summary

This chapter has discussed the validation of a prototype EPG system and the results obtained from a series of tests and user trials. Based on these results, the following features of the designed system have been proved:

1. The proposed architecture enables the system to retrieve programme metadata from all online content sources. This is evidenced from figures 6.4 and 6.20 which contain programmes from all of the four example online sources.
2. The designed system offered a cloud-based EPG system which can serve online viewing where the resulted EPG information can be displayed in a web page with the recommended programmes being supported with their web addresses to enable the user to watch those programmes directly. This is evidenced by figures 6.21 and 6.22.
3. The system applies multiple recommendation schemes to enhance the recommendation process and provide personalized EPG which can suggest programmes based on the user's viewing history (content-based approach), viewing pattern (collaborative filtering approach), viewing environmental features (context-awareness) and social networks information retrieved from the user account and their friends' accounts. This is evidenced by figures 6.6 to 6.14 and figures 6.17 and 6.18.
4. Social networks have been integrated and exploited to retrieve information from user accounts and those of their friends. By using liked, watched and want to watch categories the recommendation process has been enhanced and through the user trials it has been proven that it is possible to improve the precision of recommendations and decrease the Mean Absolute Error (MAE). Equally, reliance on the user's viewing history can improve the precision of recommendation process but it needs enough data to achieve that while the social suggestions can provide good source of recommendations even from the first time of using the system. this is evidenced by figures 6.6 to 6.16.
5. The system is able to interact with user's viewing environmental features and recommend only the programmes that match the user device type and available network bandwidth as shown in Figure 6.17 and 6.18. This point supports the ability of the system to function as a cloud-based EPG system where it can serve

different types of devices which have Internet connectivity such as personal computers, tablets, smartphones and the Smart TV.

6. It is worthy to mention that the storage area that has been used for both the user matrix and content matrix for 20 users was 4.453 MB (3.18 MB for User Matrix and 1.273 MB for Content Matrix). This will be valuable when the number of users increased where the type of information that stored using the unified format doesn't occupy a large area in the database server. In order to highlight the scalability of the system in this aspect, simple calculations will be done as shown below.

The storage size that required for the user matrix of 20 users is 3.18 MB which is approximately 3.2 MB. According to the size of viewing history for each user and the number of recommended programmes, an assumption will be made that with increasing the viewing history for each user, this size should be tripled which means that the storage size will be 9.6 MB or approximately 10 MB. If the relation between the increasing in number of users and required storage area is linear, the storage size that required for 1000 users is 500 MB and for 10000 users is 5000 MB (5 GB). However, even the relation is not linear, the required storage area will still reasonable. The same calculations could be applied for the content matrix which is smaller than the user matrix.

The next chapter will discuss the conclusions that have been determined based on the system design and the results evaluation chapters. Moreover, it will investigate the future work that could be undertaken for further improvement.

CHAPTER SEVEN

CONCLUSIONS AND FUTURE WORK

Chapter Seven

7.1 Conclusions

Whilst watching television continues to be popular but the mode of viewing has changed dramatically in recent years with a noticeable shift away from traditional linear viewing to catch-up and online. This has been further driven by the prevalence of laptops, tablets and smartphones. Similarly, both fixed and mobile broadband capacities have increased thus allowing high quality video streaming applications. With this trend towards online consumption of television and video content and the increasing number of online providers, there is a need for an EPG system which can communicate with all of the different online content sources. In that way, an EPG needs to become more focused on the programme rather than the concept of a channel.

In this thesis, a system architecture has been developed to serve online viewing. A prototype of the designed system has been implemented using a combination of Matlab and PHP and evaluated through a series of functional tests and user trials. According to the research objectives that have been mentioned in chapter 1 and described through research questions, the following points are concluded based on the system design and the results evaluation chapters.

1. How can a generic EPG system be designed that can interface with all the available online video content sources and include any new source without changing the system architecture?
 - A generic EPG system architecture has been presented to serve the online viewing. This system designed as a modular structure where each module represents a main functional part in the overall system. This system is extendible to include all the available online content sources without changing the system architecture. This system connected to four online content sources sequentially. It only needed to update a new search agent to interface with the new online content source and retrieve the needed information from it.
2. How can a searching system be designed to locate all sources of online video content given the diversity of such sources and the large variation in metadata which is used to describe programme content?

- The proposed architecture has the ability to connect with any online content source through its searching system. The design of the searching system offered this feature through providing multiple search agents controlled by search agent leader. This gives the system the ability to adapt with the different content sources by developing a search agent can adapt and communicate based on the requirements of that online content source and modify the results to be compatible with the system requirements when this results forwarded to the search agent leader.
3. Can a unified form of programme metadata be devised that works across all online video content sources?
 - A new unified format for programmes' metadata has been adopted in our proposed system which consists of 21 parameters and describes the main programme's features. This format has been used in all of the system's operations such as recommendation operations. The search engines extracted these parameters from the different online content sources' metadata and used them in creating the content matrix to describe the programmes features.
 4. How can social networks be exploited within the recommendation process and what impact could this have?
 - Multiple recommendation schemes have been used in the proposed system to personalize the recommended list of programmes based on user's viewing history (content-based), viewing pattern (collaborative filtering), viewing environmental features (context-awareness). These recommendation approaches have been used simultaneously to generate the recommendation for users through their EPG.
 - Integrating the information that retrieved from social networks enhanced the recommendation process. The evaluation metrics showed that this information increased the precision of recommendation and decreased the predicting MAE. Moreover, the system provided a solution for cold-start problem through using the initial categories selecting and social networks information to recommend programmes for users before using the online EPG in watching any programme and before creating the users' viewing history. On the other hand, the system showed that viewing history needs to accumulate the information to a reasonable value in order to enhance the recommendation operation and being an effective parameter on the overall process. This meets with the other research approaches as shown in the literature review where the depending on only the user viewing history cannot

being a valuable method to obtain a satisfied recommendations and should be hybridized with other recommendation approaches in order to enhance the recommendation process.

5. What aspects of a user's context are important in recommending video content and how can these be integrated into the system design?
 - Filtering stage has been implemented on the final recommended list of programmes to match the user viewing device and available network bandwidth. The system was able to recommend programmes that are better tailored with the resolution of user viewing device and available network bandwidth. Moreover, the system integrated the time of viewing as an aspect in recommending the programmes for each user.
6. How can the time required to execute the functions of the EPG be optimised to deliver a quality of user experience?
 - All of the operations within the EPG system are implemented when the user logged off their EPG. These operations include searching, recommendation, retrieving information and clustering operations. The only operation that is implemented when the user using the EPG is filtering the recommended list of programmes based on time and environmental features. This strategy has been adopted in the proposed system to reduce the perceived time for recommending programmes for each user.
7. How can this generic EPG be designed to serve different displaying devices such as TV's, laptops, tablets and smartphones?
 - The EPG system is designed as a cloud-based system. All of the operations have implemented in the installed server and the resulted data has stored in the same server. The designed user interface is simply could be created as web page. The filtering stage needs information from the user device, therefore, in order to serve different devices, an application should be installed on that device to be able to provide the required information and receive the offered services.

The modular structure of this architecture provides flexibility in upgrading parts of the system without affecting the core functions and modules. Additionally, the unified format that has been adopted in the proposed system helped in using a small fraction of storage area that specified for user database server and content database server. That is because the

system stored only the important part of programmes' features that is necessary to describe any programme and employ it for the different system operations.

Finally, the findings from this work met with the contributions to knowledge that mentioned in Chapter one which are:

1. Design a generic online EPG system architecture. The EPG system architecture depends on a modular structure and the modules that are connected with the external sources are extensible to include any number of online sources without changing the main structure. That is implemented through adopting a new unified metadata representation for the different metadata formats which retrieved from the online sources where this unified format will be used within the entire system.
2. Exploiting features of social networks in a new way to support the recommendation system. The recommendation system has also been uniquely designed to employ four recommendation schemes to enhance the recommendation process and obtain a personalized list of programmes for each user.
3. The EPG system has been designed as a cloud-based service, rather than a locally hosted service, which can serve any device that, has the ability to connect to the Internet. Therefore, the EPG provides the ability to appropriately match content to the viewing device and its context.

7.2 Future Work

This research has thrown up many questions that need further investigation. For example, work could be done to establish whether this system can be applied to other viewing systems such as IPTV, and conventional terrestrial or satellite viewing. The system architecture should be capable of doing this for all that is required is to develop a search agent for each of these sources that interfaces to the necessary receiving hardware.

Further research could also be carried out to investigate different algorithms within the recommendation modules to provide more precise output.

Further work could also be carried out to improve the interaction with social media including, for example, adding the ability to invite friends to watch a recommended programme or the ability of the system to notify the users about updates relevant to their

favourite programmes or to provide information about programmes which match their interests but which are not compatible with their current viewing environmental.

REFERENCES

- Abbar, S., Bouzeghoub, M., & Lopez, S. (2009). *Context-aware recommender systems: A service-oriented approach*. Paper presented at the VLDB PersDB workshop.
- Adomavicius, G., & Tuzhilin, A. (2005). Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Knowledge and Data Engineering, IEEE Transactions on*, 17(6), 734-749.
- Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems *Recommender systems handbook* (pp. 217-253): Springer.
- Al-Shamri, M. Y. H., & Bharadwaj, K. K. (2008). Fuzzy-genetic approach to recommender systems based on a novel hybrid user model. *Expert systems with applications*, 35(3), 1386-1399.
- Alsabti, K., Ranka, S., & Singh, V. (1997). An efficient k-means clustering algorithm.
- Andrew Williams, "What is Youview? A guide to the catch-up TV system". (October, 2012). [Online]. Available:
<http://www.trustedreviews.com/opinion/what-is-youview-a-guide-to-the-catch-up-tv-system-2899429>.
- Andrew Laughlin, "What is Freeview Play?". (2107). [Online]. Available:
<http://www.which.co.uk/reviews/televisions/article/what-is-freeview-play>.
- Ardissono, L., Gena, C., Torasso, P., Bellifemine, F., Difino, A., & Negro, B. (2004). User modeling and recommendation techniques for personalized electronic program guides *Personalized Digital Television* (pp. 3-26): Springer.
- Atlas, "Reference". (2009). Metabroadcast. [Online]. Available:
<https://atlas.metabroadcast.com/#reference-feeds>.

- Barragáns-Martínez, A. B., Costa-Montenegro, E., Burguillo, J. C., Rey-López, M., Mikic-Fonte, F. A., & Peleteiro, A. (2010). A hybrid content-based and item-based collaborative filtering approach to recommend TV programs enhanced with singular value decomposition. *Information Sciences*, 180(22), 4290-4311.
- Barb Gonzalez, “How much bandwidth do you need for streaming video”. (July, 2012). Sound & Vision. [Online]. Available: <http://www.soundandvision.com/content/how-much-bandwidth-do-you-need-streaming-video#X4zWMIGVWxlwvIT8.97>
- Barton, J. M. (2000). *Television meta-data standardization overview and opportunities*. Paper presented at the Consumer Electronics, 2000. ICCE. 2000 Digest of Technical Papers. International Conference on.
- Berman, K. K., Iyer, R. S., Richardson, D. J., Rahurkar, A. S., & Seetharaman, A. (2012). Recommendation engine for social networks. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- Bing, B. (2015). Next-generation video coding and streaming.
- Bobadilla, J., Serradilla, F., & Bernal, J. (2010). A new collaborative filtering metric that improves the behavior of recommender systems. *Knowledge-Based Systems*, 23(6), 520-528.
- Bondad-Brown, B. A., Rice, R. E., & Pearce, K. E. (2012). Influences on TV viewing and online user-shared video use: Demographics, generations, contextual age, media use, motivations, and audience activity. *Journal of Broadcasting & Electronic Media*, 56(4), 471-493.
- Bonhard, P., & Sasse, M. A. (2006). 'Knowing me, knowing you'—Using profiles and social networking to improve recommender systems. *BT Technology Journal*, 24(3), 84-98.

- Boyer, F. E., & Demers, T. B. (2014). Television system with aided user program searching: Google Patents.
- Boyer, F. E., Demers, T. B., & Blackwell, B. A. (2014). Electronic mail reminder for an internet television program guide. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- Campos, P. G., Fernández-Tobías, I., Cantador, I., & Díez, F. (2013). *Context-aware movie recommendations: an empirical comparison of pre-filtering, post-filtering and contextual modeling approaches*. Paper presented at the International Conference on Electronic Commerce and Web Technologies.
- Carrer-Neto, W., Hernández-Alcaraz, M. L., Valencia-García, R., & García-Sánchez, F. (2012). Social knowledge-based recommender system. Application to the movies domain. *Expert systems with applications*, 39(12), 10990-11000.
- Cisco VNI, "The future of content marketing online video". (September 2015). Social Times. [Online]. Available:
http://www.adweek.com/socialtimes/the-future-of-content-marketing-is-online-video/626021#disqus_thread.
- Cha, M., Rodriguez, P., Crowcroft, J., Moon, S., & Amatriain, X. (2008). *Watching television over an IP network*. Paper presented at the Proceedings of the 8th ACM SIGCOMM conference on Internet measurement.
- Chang, J.-H., Lai, C.-F., Wang, M.-S., & Wu, T.-Y. (2013). A cloud-based intelligent TV program recommendation system. *Computers & Electrical Engineering*, 39(7), 2379-2399.
- Chang, N., Irvan, M., & Terano, T. (2013). A TV program recommender framework. *Procedia Computer Science*, 22, 561-570.

- Chen, Y.-C., Huang, H.-C., & Huang, Y.-M. (2009). Community-based program recommendation for the next generation electronic program guide. *Consumer Electronics, IEEE Transactions on*, 55(2), 707-712.
- Chen, Y.-L., Cheng, L.-C., & Chuang, C.-N. (2008). A group recommendation system with consideration of interactions among group members. *Expert systems with applications*, 34(3), 2082-2090.
- Chris Jackson, "Changes to Atlas access". (May, 2106). Metabroadcast. [Online]. Available:
[https://metabroadcast.com/blog/changes-to-atlas-access.](https://metabroadcast.com/blog/changes-to-atlas-access)
- Concolato, C. (2009). Generation, Streaming and Presentation of Electronic Program Guide. *Networked Television Adjunct proceedings of EuroITV 2009*, 46.
- Cui, Y.-b., Sun, J., Wang, B., & Li, X.-r. (2012). *Design and realization of EPG system and GUI in digital TV set-top box*. Paper presented at the Computer Science and Information Processing (CSIP), 2012 International Conference on.
- Dai, W., & Cohen, R. (2003). *Dynamic personalized TV recommendation system*. Paper presented at the Proceedings of the UM 2003 Workshop on Personalization in Future TV.
- Dajda, T., Ciślak, M., Hełdak, G., & Pacyna, P. (1996). DESIGN AND IMPLEMENTATION OF THE ELECTRONIC PROGRAMME GUIDES FOR THE MPEG2-BASED DVB SYSTEM.
- Dawes, C., "The past, present and future of entertainment discovery". (August 2015). Broadband Journal of the SCTE. [Online]. Available:
[http://www.thescte.eu/broadband/vol37_no3_QFRR/files/assets/basic-html/page42.html.](http://www.thescte.eu/broadband/vol37_no3_QFRR/files/assets/basic-html/page42.html)
- De Campos, L. M., Fernández-Luna, J. M., Huete, J. F., & Rueda-Morales, M. A. (2010). Combining content-based and collaborative recommendations: A hybrid approach

based on Bayesian networks. *International Journal of Approximate Reasoning*, 51(7), 785-799.

Debnath, S., Ganguly, N., & Mitra, P. (2008). *Feature weighting in content based recommendation system using social network analysis*. Paper presented at the Proceedings of the 17th international conference on World Wide Web.

Ding, C., & He, X. (2004). *K-means clustering via principal component analysis*. Paper presented at the Proceedings of the twenty-first international conference on Machine learning.

Drazin, J. P., & Kram, A. H. (2014). Electronic program guide for indicating availability of past programs: Google Patents.

Ehrmantraut, M., Härder, T., Wittig, H., & Steinmetz, R. (1996). *The personal electronic program guide—towards the pre-selection of individual TV programs*. Paper presented at the Proceedings of the fifth international conference on Information and knowledge management.

Ellis, M. D., Davis, B. L., & Knudson, E. B. (2014). Electronic television program guide schedule system and method with scan feature: Google Patents.

Ellison, N. B., Steinfield, C., & Lampe, C. (2007). The benefits of Facebook “friends:” Social capital and college students’ use of online social network sites. *Journal of Computer-Mediated Communication*, 12(4), 1143-1168.

Ellison, N. B., Steinfield, C., & Lampe, C. (2011). Connection strategies: Social capital implications of Facebook-enabled communication practices. *New media & society*, 873-892.

ETSI, T. (1983). 102 822-3-1 Broadcast and On-Line Services: Search, select and rightful use of content on personal storage systems (“TV-Anytime Phase 1”); part 3:

Metadata. Shibao Zheng (M'04) graduated from Xidian University PR China and received the BS and MS degrees in.

ETSI, T. (2007). 102 822-2 V1. 4.1. *Broadcast and On-line Services: Search, select, and rightful use of content on personal storage systems ("TV-Anytime"); Part 2: Phase 1-System description.*

Evain, J.-P. (2000). TV-Anytime metadata: a preliminary specification on schedule. *EBU Technical Review, 14.*

Fischer, D. (2012). Automatic generation of content recommendations weighted by social network context. Washington, DC: U.S. Patent and Trademark Office: Google Patents.

Freeview, "Freeview Play". (2017). [Online]. Available:
[https://www.freeview.co.uk/why-freeview/freeview-play.](https://www.freeview.co.uk/why-freeview/freeview-play)

Funk, J., & Noffsinger, B. (2014). System and method for searching multimedia: Google Patents.

Golbeck, J. (2006). *Generating predictive movie recommendations from trust in social networks.* Paper presented at the International Conference on Trust Management.

Han, X., Wang, L., Crespi, N., Park, S., & Cuevas, Á. (2015). Alike people, alike interests? Inferring interest similarity in online social networks. *Decision Support Systems, 69*, 92-106.

He, J., & Chu, W. W. (2010). *A social network-based recommender system (SNRS):* Springer.

Holanda, P., Guilherme, B., da Silva, A. P. C., & Goussevskaia, O. (2015). *TV Goes Social: Characterizing User Interaction in an Online Social Network for TV Fans.* Paper presented at the International Conference on Web Engineering.

- Hong, H.-K., & Lee, D.-H. (2015). A personalized refinement technique for semantic multimedia content search in smart TV. *IEEE Transactions on Consumer Electronics*, 61(4), 581-587.
- Hsu, S. H., Wen, M.-H., Lin, H.-C., Lee, C.-C., & Lee, C.-H. (2007). *AIMED-A personalized TV recommendation system*. Paper presented at the European Conference on Interactive Television.
- Huang, A. (2008). *Similarity measures for text document clustering*. Paper presented at the Proceedings of the sixth new zealand computer science research student conference (NZCSRSC2008), Christchurch, New Zealand.
- Huh, M., & Kang, S. (2010). *Standard and Metadata Development Considerations for IPTV Service Provider and Service Discovery*. Paper presented at the 2010 International Conference on Information and Communication Technology Convergence (ICTC).
- Jan Ozer, "What is MPEG DASH?". (November, 2011). Streaming Media. [Online]. Available:
<http://www.streamingmedia.com/Articles/Editorial/What-Is-.../What-is-MPEG-DASH-79041.aspx>.
- Jeong, J.-W., & Lee, D.-H. (2014). Inferring search intents from remote control movement patterns: A new content search method for smart TV. *IEEE Transactions on Consumer Electronics*, 60(1), 92-98.
- Jerding, D. F., Rodriguez, A. A., & Banker, R. O. (2013). Integrated searching system for interactive media guide: Google Patents.
- Jiang, M., Cui, P., Liu, R., Yang, Q., Wang, F., Zhu, W., & Yang, S. (2012). *Social contextual recommendation*. Paper presented at the Proceedings of the 21st ACM international conference on Information and knowledge management.

- Jonathan Tweed, "Getting full UK listing data from Atlas: It's easy". (November, 2012). Metabroadcast. [Online]. Available: <https://metabroadcast.com/blog/getting-full-uk-listings-data-from-atlas-it-s-easy>.
- Kamahara, J., Asakawa, T., Shimojo, S., & Miyahara, H. (2005). *A community-based recommendation system to reveal unexpected interests*. Paper presented at the 11th international multimedia modelling conference.
- Kim, E., Pyo, S., Park, E., & Kim, M. (2011). An automatic recommendation scheme of TV program contents for (IP) TV personalization. *IEEE Transactions on Broadcasting*, 57(3), 674-684.
- Kim, J. K., Kim, H. K., Oh, H. Y., & Ryu, Y. U. (2010). A group recommendation system for online communities. *International Journal of Information Management*, 30(3), 212-219.
- Kim, M.-E., Cho, J.-M., Yoo, J.-J., Hong, J.-W., & Kim, S.-H. (2013). *A Proposal of Semantic Analysis based Integrated Multi-Level Search System for Smart TV*. Paper presented at the Advanced Communication Technology (ICACT), 2013 15th International Conference on.
- Kim Staples, "Youview: everything you need to know", (January, 2017). [Online]. Available: <https://www.broadbandchoices.co.uk/guides/hardware/youview-box-review-all-you-need-to-know>.
- Kim, T., Kim, J., & Hahn, M. (2010). *Policy-based service overlay IPTV framework for open strategy*. Paper presented at the Network Infrastructure and Digital Content, 2010 2nd IEEE International Conference on.
- Knee, R. A., Reynolds, S. J., Ellis, M. D., & Hassell, J. G. (2009). Interactive television program guide system for determining user values for demographic categories: Google Patents.

- Konstas, I., Stathopoulos, V., & Jose, J. M. (2009). *On social networks and collaborative recommendation*. Paper presented at the Proceedings of the 32nd international ACM SIGIR conference on Research and development in information retrieval.
- Lai, C.-F., Chang, J.-H., Hu, C.-C., Huang, Y.-M., & Chao, H.-C. (2011). CPRS: A cloud-based program recommendation system for digital TV platforms. *Future Generation Computer Systems*, 27(6), 823-835.
- Lai, Y. X., Lai, C. F., Hu, C. C., Chao, H. C., & Huang, Y. M. (2011). A personalized mobile IPTV system with seamless video reconstruction algorithm in cloud networks. *International Journal of Communication Systems*, 24(10), 1375-1387.
- Lee, S.-Y., & Lee, S.-W. (2015). Online video services and other media: Substitutes or complement. *Computers in Human Behavior*, 51, 293-299.
- Legall, L. C., Masli, H., & Kulas, C. J. (1999). Integrated search of electronic program guide, internet and other information resources: Google Patents.
- Liddy, E. D., & Yu, E. S.-I. (2001). System for retrieving multimedia information from the internet using multiple evolving intelligent agents: Google Patents.
- Likas, A., Vlassis, N., & Verbeek, J. J. (2003). The global k-means clustering algorithm. *Pattern recognition*, 36(2), 451-461.
- Liu, F., & Lee, H. J. (2010). Use of social network information to enhance collaborative filtering performance. *Expert systems with applications*, 37(7), 4772-4778.
- Liu, X., & Aberer, K. (2013). *SoCo: a social network aided context-aware recommender system*. Paper presented at the Proceedings of the 22nd international conference on World Wide Web.

- Macedo, P., Cardoso, J., & Pinto, A. M. (2014). Enriching Electronic Programming Guides with Web Data.
- Mahimkar, A. A., Ge, Z., Shaikh, A., Wang, J., Yates, J., Zhang, Y., & Zhao, Q. (2009). *Towards automated performance diagnosis in a large IPTV network*. Paper presented at the ACM SIGCOMM Computer Communication Review.
- Mai, J., Fan, Y., & Shen, Y. (2009). *A neural networks-based clustering collaborative filtering algorithm in e-commerce recommendation system*. Paper presented at the Web Information Systems and Mining, 2009. WISM 2009. International Conference on.
- Mankovitz, R. J. (2014). System and method for searching a database of television schedule information: Google Patents.
- Martínez, A. B. B., Arias, J. J. P., Vilas, A. F., Duque, J. G., & Nores, M. L. (2009). What's on TV tonight? An efficient and effective personalized recommender system of TV programs. *IEEE Transactions on Consumer Electronics*, 55(1), 286-294.
- Martinsson, E. (2006). IPTV the future of television. *Report-Computer Communication*.
- Mathur, A. (2013). Social network based recommendation method and system. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- McKissick, P. L., & Forrer, M. A. (2013). Electronic program guide with advance notification. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- Messer, A., Kunjithapatham, A., Nguyen, P., Rathod, P., Sheshagiri, M., Cheng, D., & Gibbs, S. (2008). *Internet search on TV*. Paper presented at the 2008 5th IEEE Consumer Communications and Networking Conference.

Metabroadcast, "Release notes – week to 24th May 2017". (May 2017). [Online]. Available:

<https://metabroadcast.com/>.

Mitchell, K., Jones, A., Ishmael, J., & Race, N. J. (2011). Social TV: The impact of social awareness on content navigation within IPTV systems. *Computers in Entertainment (CIE)*, 9(3), 19.

Miyamori, H., Stejic, Z., Araki, T., Minakuchi, M., Ma, Q., & Tanaka, K. (2006). *Towards integration services for heterogeneous resources: An integrated search engine for web content and tv programs*. Paper presented at the Second International Conference on Semantics, Knowledge and Grid, 2006. SKG'06.

Mo, Y., Chen, J., Xie, X., Luo, C., & Yang, L. T. (2014). Cloud-based mobile multimedia recommendation system with user behavior information. *IEEE SYSTEMS JOURNAL*, 8(1), 184-193.

Mozilla, "Setting up adaptive streaming media sources". (2017). Mozilla. [Online]. Available:

https://developer.mozilla.org/en-US/Apps/Fundamentals/Audio_and_video_delivery/Setting_up_adaptive_streaming_media_sources.

O'Sullivan, D., Smyth, B., Wilson, D. C., McDonald, K., & Smeaton, A. (2004). Improving the quality of the personalized electronic program guide. *User Modeling and User-Adapted Interaction*, 14(1), 5-36.

Ofcom, "On-demand and online research: consumption and concerns". (January 2016). UK Communication Regulator. [Online]. Available:

https://www.ofcom.org.uk/data/assets/pdf_file/0013/53203/odo_consumption_and_concerns_jan16.pdf.

- Ofcom, "The UK is now a smartphone society". (August 2015). UK Communication Regulator. [Online]. Available: <https://www.ofcom.org.uk/about-ofcom/latest/media/media-releases/2015/cmr-uk-2015>.
- Ofcom, "The communication market 2016 - Television and audio visual". (August 2016). UK Communication Regulator. [Online]. Available: https://www.ofcom.org.uk/data/assets/pdf_file/0026/17495/uk_tv.pdf.
- Oh, H. J., Ozkaya, E., & LaRose, R. (2014). How does online social networking enhance life satisfaction? The relationships among online supportive interaction, affect, perceived social support, sense of community, and life satisfaction. *Computers in Human Behavior*, 30, 69-78.
- Oh, J., Kim, S., Kim, J., & Yu, H. (2014). When to recommend: A new issue on tv show recommendation. *Information Sciences*, 280, 261-274.
- Pham, M. C., Cao, Y., Klamma, R., & Jarke, M. (2011). A clustering approach for collaborative filtering recommendation using social network analysis. *J. UCS*, 17(4), 583-604.
- Punchihewa, A., De Silva, A. M., & Diao, Y. (2010). Internet Protocol Television (IPTV). *Multi-Media Research Group/ School of Engineering and Advanced Technology/ Massey University, New Zealand*.
- Quijano-Sanchez, L., Recio-Garcia, J. A., & Diaz-Agudo, B. (2011). *Happymovie: A facebook application for recommending movies to groups*. Paper presented at the 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence.
- Rao, P. S., Sloo, D. H., Danker, D., & Nyako, G. K. (2015). Searching electronic program guide data: Google Patents.

- Rik Henderson & Max Langridge, "What is Freeview Play, when is it coming to my TV and how can I get it?". (August, 2017). [Online]. Available: <http://www.pocket-lint.com/news/132739-what-is-freeview-play-when-is-it-coming-to-my-tv-and-how-can-i-get-it>.
- Ricci, F., Rokach, L., & Shapira, B. (2011). *Introduction to recommender systems handbook*: Springer.
- Rosengren, J. (1996). Electronic programme guides and service information. *Philips Journal of Research*, 50(1), 253-265.
- Rothschild, L. M. (2011). Systems, devices, and methods for integrated searching and retrieving internet or digital content across a communication network for a multimedia platform: Google Patents.
- Schaffer, J. D., Lee, K. P., & Gutta, S. (2011). Three-way media recommendation method and system: Google Patents.
- Scholl, H. R., Thelen, E., Kneissler, J., & Kellner, A. (2009). Program recommendation system. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- Shapira, B., Rokach, L., & Freilikhman, S. (2013). Facebook single and cross domain data for recommendation systems. *User Modeling and User-Adapted Interaction*, 23(2-3), 211-247.
- Shin, D., Lee, J.-w., Yeon, J., & Lee, S.-g. (2009). *Context-aware recommendation by aggregating user context*. Paper presented at the 2009 IEEE Conference on Commerce and Enterprise Computing.
- Shin, H., Lee, M., & Kim, E. Y. (2009). Personalized digital TV content recommendation with integration of user behavior profiling and multimodal content rating. *Consumer Electronics, IEEE Transactions on*, 55(3), 1417-1423.

- Smith, A., "Shared, collaborative and on demand: The new digital economy". (May 19, 2016). Pew Research Centre. [Online]. Available: <http://www.pewinternet.org/2016/05/19/OTHER-SHARED-AND-ON-DEMAND-SERVICES/>.
- Simsarian, J. E., & Duelk, M. (2007). *IPTV bandwidth demands in metropolitan area networks*. Paper presented at the Local & Metropolitan Area Networks, 2007. LANMAN 2007. 15th IEEE Workshop on.
- Smith, J. R., & Chang, S.-F. (1996). Searching for images and videos on the world-wide web. *IEEE Multimedia Magazine*.
- Smyth, B., & Cotter, P. (2001). Personalized electronic program guides for digital TV. *Ai Magazine*, 22(2), 89.
- So, E. C., Coden, A. R., & Mak, S.-W. (1999). Using multiple search engines to search multimedia data: Google Patents.
- Song, S., Moustafa, H., & Afifi, H. (2012). Advanced IPTV services personalization through context-aware content recommendation. *IEEE Transactions on Multimedia*, 14(6), 1528-1537.
- Song, Z., Wei, Z., & Jia, D. (2012). *The algorithms of intelligent TV programs recommendation based on data mining*. Paper presented at the Computer Science & Service System (CSSS), 2012 International Conference on.
- STATS, "YouTube Company Statistics". (2017). Statistics Brain. [Online]. Available: <http://www.statisticbrain.com/youtube-statistics/>.
- Stegg, F. v. (2015). Context-aware recommender systems.

- Stettner, A. P. (2012). System and method for dynamic electronic program guide (EPG) data downloads. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- Tan, C., Zhu, C., Wang, X., & Chang, G. (2009). *An EPG Recommender Framework with Information Exchange between Users*. Paper presented at the Hybrid Intelligent Systems, 2009. HIS'09. Ninth International Conference on.
- Taylor, C. R. (2015). On the new era of zipping and zapping: the need for research on how advertisers deal with an Era of time-shifted television program viewing on a variety of devices. *International Journal of Advertising*, 34(3), 403-405.
- Theodoridis, S., & Koutroumbas, K. (2003). Pattern recognition.
- Thomas, G. (2000). *Metadata for Enhanced Electronic Program Guides*. Paper presented at the Technical Conference and Exhibition, 142nd.
- Tu, H.-C., & Hsiang, J. (2000). An architecture and category knowledge for intelligent information retrieval agents. *Decision Support Systems*, 28(3), 255-268.
- Uchyigit, G., & Clark, K. (2003). *A Personalised Multi-Modal Electronic Program Guide*. Paper presented at the European Conference on Interactive Television: from Viewers to Actors.
- Walter, F. E., Battiston, S., & Schweitzer, F. (2008). A model of a trust-based recommendation system on a social network. *Autonomous Agents and Multi-Agent Systems*, 16(1), 57-74.
- Walgrove, A., "The explosive growth of online video". (July 2015). The Content Strategist. [Online]. Available: <https://contently.com/strategist/2015/07/06/the-explosive-growth-of-online-video-in-5-charts/>.

- Wu, J., Liu, Q., & Luo, S. (2008). *Clustering technology application in e-commerce recommendation system*. Paper presented at the 2008 International Conference on Management of e-Commerce and e-Government.
- Xu, J., Zhang, L.-J., Lu, H., & Li, Y. (2002). *The development and prospect of personalized TV program recommendation systems*. Paper presented at the Multimedia Software Engineering, 2002. Proceedings. Fourth International Symposium on.
- Xu, M., Berkovsky, S., Ardon, S., Triukose, S., Mahanti, A., & Koprinska, I. (2013). *Catch-up tv recommendations: show old favourites and find new ones*. Paper presented at the Proceedings of the 7th ACM conference on Recommender systems.
- Ye, M., Liu, X., & Lee, W.-C. (2012). *Exploring social influence for recommendation: a generative model approach*. Paper presented at the Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval.
- Yim, J., Lee, G., Lee, T., & Jeon, J. (2014). Review of IPTV System Architectures. *Advanced Science and Technology Letters*, 46, 83-86.
- YouTube, "Statistics". (2016). YouTube. [Online]. Available:
<https://www.youtube.com/yt/press/en-GB/statistics.html>.
- YouTube, "Delivering live YouTube content via DASH". (2014). Google. [Online]. Available:
<https://developers.google.com/youtube/v3/live/guides/encoding-with-dash>.
- Youview, "How to get Youview". (2107). [Online]. Available:
<https://www.youview.com/>.

- Yu, Z., Zhou, X., Hao, Y., & Gu, J. (2006). TV program recommendation for multiple viewers based on user profile merging. *User Modeling and User-Adapted Interaction*, 16(1), 63-82.
- Yuen, H. C., Mankovitz, R. J., Kwoh, D. S., & Leung, E. Y. (2000). System and method for grazing television channels from an electronic program guide. Washington, DC: U.S. Patent and Trademark Office: Google Patents.
- Zeadally, S., Moustafa, H., & Siddiqui, F. (2011). Internet protocol television (IPTV): architecture, trends, and challenges. *IEEE SYSTEMS JOURNAL*, 5(4), 518-527.
- Zhang, H., & Zheng, S. (2005). *Personalized TV program recommendation based on TV-anytime metadata*. Paper presented at the Proceedings of the Ninth International Symposium on Consumer Electronics, 2005.(ISCE 2005).

Appendices

Note: All the appendices are included in the attached CD.

Appendix (I) BBC Interface Code.

Appendix (II) Dailymotion Interface Code.

Appendix (III) YouTube Interface Code.

Appendix (IV) The Movie Database Interface Code.

Appendix (V) Facebook Application.

Appendix (VI) User Interface.

Appendix (VII) Test Sheet.

Appendix (VI-Implicit Function) Collaborative Filtering.

Appendix (VI-Implicit Function) Create BBC Content.

Appendix (VI-Implicit Function) Create Dailymotion Content.

Appendix (VI-Implicit Function) Create Facebook Content List.

Appendix (VI-Implicit Function) Create Facebook List.

Appendix (VI-Implicit Function) Create TMDb Content.

Appendix (VI-Implicit Function) Create VH Searching Content MAT.

Appendix (VI-Implicit Function) Create YouTube Content.

Appendix (VI-Implicit Function) Direct Search.

Appendix (VI-Implicit Function) Facebook Friends.

Appendix (VI-Implicit Function) Facebook Info.

Appendix (VI-Implicit Function) Facebook Matrix.

Appendix (VI-Implicit Function) Recommendation VH List.

Appendix (VI-Implicit Function) VH Searching.