Recommendation Systems in the Online Classified Industry
Studying the Online Classified Businesses and the Potential Role of Recommendation Solutions in their Development

Written by
Áron Erdős
Business Information Technology (BSc) – Specialization on E-Business
aron.erdos@stud.uni-corvinus.hu

Supervisor
Róbert Pintér
Institute of Informatics – Department of Information and Communication
robert.pinter@uni-corvinus.hu

April 2014
Abstract

This thesis aims to provide an overview on today’s online classified businesses and recommendation solutions based on a research conducted at the recommendation systems provider Gravity R&D and its various online classified partners.

Classified sites have become prominent players of the E-commerce and Internet Advertising sectors. Considering their business models, there are a number of challenges in making a classified site profitable, but recommendation systems can potentially provide solutions to many of these issues.

Recommendation systems are now widely applied in many of today’s data-rich industries as a result of a significant scientific development. These solutions, including the ones implemented at classified sites, are built up using several specific technologies. In fact, their utilization and goal can be very different site-by-site, largely depending on the maturity of the studied classified businesses and on the platforms’ actual interaction points with the different user segments.

It is advantageous to study recommendation systems in the classified industry not only due to the wide usage and vast variety of the applied solutions. The large amount of items and users that can be found on classified websites, as well as the available experience and good case practices make it easier to get a complete overview on the ways of the implementation and on the possible performance of recommendation solutions.

To gain this overview, this thesis will present both of the above subjects, and it will introduce and evaluate several examples and use cases of recommendation solutions implemented at different classified sites, illustrated with the necessary data, charts and figures.
Contents

1. Introduction ............................................................................................................................................. 6

2. An Overview on Classified Sites .......................................................................................................... 8
   2.1 Introducing Classified Sites ............................................................................................................. 8
       2.1.1 From E-commerce to Classifieds ............................................................................................ 8
       2.1.2 Development ........................................................................................................................ 9
       2.1.3 General Characteristics .......................................................................................................... 11
       2.1.4 General and Specialized Classifieds ..................................................................................... 13
   2.2 Revenues and Business Models ........................................................................................................ 14
       2.2.1 Monetization ........................................................................................................................ 14
       2.2.2 Possible Sources of Revenue ................................................................................................ 15
       2.2.3 Business Models .................................................................................................................. 16
   2.3 The Maturity Stages of Classified Sites ............................................................................................ 17
   2.4 The Users of Classified Sites .......................................................................................................... 22
       2.4.1 User Groups .......................................................................................................................... 22
       2.4.2 User Life Cycle ..................................................................................................................... 22

3. An Overview on Recommendation Solutions ......................................................................................... 25
   3.1 Recommendation Systems and Their Applications ........................................................................... 25
       3.1.1 Definition .............................................................................................................................. 25
       3.1.2 Development and Application ............................................................................................... 25
       3.1.3 The Role of Personalization .................................................................................................. 27
       3.1.4 The Main Functions of Recommendation Systems ................................................................ 28
   3.2 How Recommendation Algorithms Work ......................................................................................... 32
       3.2.1 Basics ..................................................................................................................................... 32
       3.2.2 Data and Methods ................................................................................................................. 32
       3.2.3 Architecture and Integration .................................................................................................. 34

4. Recommendation Solutions on Classified Sites ...................................................................................... 37
   4.1 Categorization and Measurement .................................................................................................... 37
   4.2 Attracting Visitors and Prolonging Browsing Sessions ....................................................................... 39
       4.2.1 Main Pages ............................................................................................................................. 39
       4.2.2 Dynamic Landing Pages ......................................................................................................... 40
       4.2.3 Zero Result and 404 Pages .................................................................................................... 42
       4.2.4 Ad Detail (Item) Pages ........................................................................................................... 43
       4.2.5 After Ad Reply Boxes and Pages ........................................................................................... 45
4.3 Easing Navigation and Decision Making

4.3.1 Result (Listing) Pages

4.3.2 Keyword Based Search Result Pages

4.3.3 Device Awareness

4.4 Monetizing Users

4.4.1 Optimized Gallery

4.4.2 Campaign Management and Reporting Tools for Operators

4.4.3 Insights for B2C Advertisers

4.5 Retaining Users

4.5.1 Email Personalization

4.5.2 Personalized Retargeting through Display Advertising

5. Further Challenges

6. Conclusions

References
Figures

Figure 1: The Maturity Stages of Classified Sites (Created by the Author) ........................................ 18
Figure 2: The Stages of the User Life Cycle (Created by the Author) ................................................. 23
Figure 3: The Architecture of Recommendation Methods (Gravity R&D, 2013) ......................... 35
Figure 4: The Most Relevant Functions of Recommendation Systems According to the Classifieds’ Maturity (Created by the Author) ................................................................. 37
Figure 5: Recommendation Placements on the Main Page (Created by the Author) ........... 39
Figure 6: CTR (%) on Recommendation Placements on the Main Page (Gravity R&D, 2014c) ... 40
Figure 7: Recommendation Placements on Zero Result Pages (Created by the Author) .......... 42
Figure 8: CTR (%) on Recommendation Placements on Zero Result Pages (Gravity R&D, 2014c) ................................................................................................................................. 43
Figure 9: Recommendation Placements on Ad Detail (Item) Pages (Created by the Author) .... 44
Figure 10: CTR (%) on Recommendation Placements on Ad Detail (Item) Pages (Gravity R&D, 2014c) ................................................................................................................................. 44
Figure 11: Recommendation Placements in After Ad Reply Boxes (Created by the Author) ..... 46
Figure 12: CTR (%) on Recommendation Placements in After Ad Reply Boxes and Pages (Gravity R&D, 2014c) ................................................................................................................................. 47
Figure 13: Recommendation Placements on Result (Listing) Pages (Created by the Author) .... 49
Figure 14: CTR (%) on Recommendation Placements on Result (Listing) Pages (Gravity R&D, 2014c) ................................................................................................................................. 49
Figure 15: Recommendation Placements on Keyword Based Search Result Pages (Created by the Author) ................................................................................................................................. 50
Figure 16: CTR (%) on Recommendation Placements on Keyword Based Search Result Pages (Gravity R&D, 2014c) ................................................................................................................................. 51
Figure 17: CTR (%) on Recommendation Placements in Newsletters (Gravity R&D, 2014c) ...... 58
Figure 18: CTR (%) on Advertisements Retargeted through Affiliate Sites (Gravity R&D, 2014c) ................................................................................................................................. 62
Figure 19: CTR (%) on Advertisements Retargeted through Ad Networks (Gravity R&D, 2014c) ................................................................................................................................. 63
1. Introduction

It may not be surprising that as the humanity becomes more and more familiar with the digital world and with the usage of the Internet and connected devices, furthermore as the World Wide Web becomes accessible to an increasing number of people, the demand towards the E-commerce services gets diversified. As a part of this process, the significance of decentralized forms of businesses grows steadily. With the increasing popularity of the Internet and with the wide possibilities that it provides, we can say that the appearance and the continuous development of the so called ‘C2C’ (customer-to-customer) marketplaces is part of a natural progress.

Although online classifieds are not necessarily marketplaces in the strict sense of the word (as they are representing a form of advertising), they are playing a significant role in the exchange of goods and services between individuals today. According to a study conducted in 2009 (Jones), 49% of American adults have ever used a Classified Site, which have probably increased since. Furthermore, although there are only estimates regarding their revenues, Classified Sites are important players of the Advertising and the E-commerce industries beyond doubt. However, while there are a large number of – both business and academic – materials available about both of these industries as such, the amount of public studies related to online Classifieds is quite limited.

On the other hand, systems that are capable of providing predictive analytics in real time just become known in the past 10-15 years with the spreading phenomenon of ‘Big Data’. It is also a result of the accelerating digitalization that we continuously produce such an amount of data that seems to be impossible to process or even to store. Among these, there is a mass of implicitly generated information related to users’ behaviors on various digital platforms.

A number of innovative companies – such as Gravity Research & Development, Ltd.¹, which has provided the necessary expertise for this study – were founded to develop technologies that are able to analyze and process this enormous amount of information in real time. With the help of different data mining techniques and adaptively learning algorithms, these technologies aim to transform all of the available inputs into a highly valuable output that can continuously contribute to the development of many Internet-based services and applications.

Currently – as the technology has proven, providing a growing number of use cases – Recommendation Systems have acquired a significant business potential in almost all fields of

¹ http://www.gravityrd.com/
life (including retail and financial sectors, logistics and healthcare), while they are already being applied extensively in the E-commerce and Internet Advertising industries. Moreover, due to some key characteristics, Classified Sites are particularly appropriate for these solutions to be implemented.

All the above, furthermore the available data, experience and good case practices make Classified Sites a suitable platform for studying Recommendation Systems and their contribution to the different user needs and business goals comprehensively, while we must not forget that both of these subjects are – even individually – representing an interesting and important research topic.
2. An Overview on Classified Sites

As a starting point, this section will provide a general overview on the classified industry by introducing the main elements of classified businesses, including the key characteristics, the different types of sites, the possible sources of revenue and the most commonly applied business models. As a combination of these, this section will also present a maturity model, furthermore it will discuss the typical users of Classified Sites briefly.

2.1 Introducing Classified Sites

There are some characteristics of Classified Sites that make them unique among other E-commerce platforms. These basic fundamentals are quite important to mention in order to give a complete overview about the topic, furthermore some parts of this section will also largely contribute to the comprehension of Recommendation Systems later on.

2.1.1 From E-commerce to Classifieds

The most common classification of E-commerce services is based on the orientation of transactions. These can be carried out among governments, businesses and customers in every possible combination, but the majority of them are business-to-business (B2B), business-to-customer (B2C) and customer-to-customer (C2C) deals. Classified Sites are listing both B2C and C2C offers; however, customer-to-customer interactions mean the main essence of these platforms.²

In case of C2C E-commerce services, both parties of a transaction are natural persons, and the online platform that facilitates their interactions is based either on an auction or on a classified model. In the auction model sellers are offering their products up for a bid and the highest bidders will be able to actually buy these, while sellers on Classified Sites are advertising their products on a fixed price determined by them. The advertised items can vary widely among relatively inexpensive second-hand products (such as electronics, clothes, collectibles, toys and the like), equipment (e.g. tools, vehicles), real estate, jobs and various services, among others.

The expansion of the Internet and the appearance of certain web services both “streamlined and globalized the traditional person-to-person trading”, making it possible for almost everyone to exchange goods and services economically, regardless of time and space. Today’s C2C websites are facilitating an easy exploration for buyers, while enabling sellers to immediately list any

² There is also a unique segment of classifieds focusing specifically on B2B offers, such as Alibaba.com.
items for sale (Bjornsson, 2001). Some of the most well-known and most successful C2C marketplaces are eBay\(^3\) and Craigslist.\(^4\)

Private advertisers on these sites are not operating a business since they are selling mostly their second-hand goods on an irregular basis in order to get rid of their redundant items or in order to generate temporary revenue. Although it is one of the biggest advantages of C2C platforms that individuals can sell their products themselves, business sellers on Classified and Auction Sites are playing the role of the middle man, thereby bringing in B2C transactions. In many cases, business sellers are not quite different from the operators of traditional E-Shops – as they are regularly purchasing (or in rare cases, they are producing) goods at a certain price and selling these with a margin that makes profit – only they are trading on a centralized platform that attracts potentially much more users than an independent online shop. Another group of business sellers are advertising and managing the sales process of properties belonging to multiple individuals, thereby acting as brokers or agents.

The grey area in between is formed by the regular sellers without any legal form of business. Although it is difficult to estimate, presumably it is not a small segment.\(^5\) On the other hand, one of the strongest motivations for a significant amount of sellers in founding a business is that they are successful as a private advertiser and this is a way of scaling up and legalizing their operation.

In conclusion, both Classified and Auction Sites are E-commerce applications that are focusing on C2C, while carrying out B2C transactions as well. The relative proportion of these transactions can differ, mostly depending on the type of the given site,\(^6\) on the applied business model and on the maturity of the platform, among others.

2.1.2 Development

**Origin**

According to the definition of the Random House Dictionary (2014), a Classified Ad is “an advertisement in a newspaper, magazine, or the like generally dealing with offers of or requests for jobs, houses, apartments, used cars, and the like.” The word ‘classified’ is derived probably from the fact that the section of the newspapers, where the classified advertisements where published, were divided or ‘classified’ by advertisement categories or ‘classes’.

\(^3\) [http://www.ebay.com/](http://www.ebay.com/)
\(^4\) [http://www.craigslist.org/](http://www.craigslist.org/)
\(^5\) The size may differ depending on each website’s policy and on local regulations.
\(^6\) For example it is unlikely to find many C2C advertisements on a Job Site.
The first classified ad was printed in 1704 (Advertising Age, 1999) and this became a major form of advertising over more than 300 years. Then, according to the Newspaper Associations of America (2013), U.S. newspapers’ revenue from classified advertising suddenly declined 71% (from $15,898 to $4,626 million) in 10 years, between 2002 and 2012. This was due to the presence and expansion of online Classifieds, which started in 1995 with two sites: AdOne (which was less successful) and Craigslist, which is still one of the market leaders (Wenger, 2011).

**The Size of the Industry Today**

Unfortunately, probably due to the lack of standard reporting, there are no particularly reliable estimates regarding classified revenues. However, there are a few reports and studies that might be able to provide a rough approximation about the size of the industry today.

According to survey sponsored by IAB (PwC, 2014), classified revenues totaled $2.6 billion in 2013 (up 7% from last year’s $2.4 billion), which represents 7% of the total internet advertising revenues in the United States.

A report prepared by ZenithOptimedia (2013, p. 7) is a bit more elaborate by stating “online classified has been subdued since the downturn in 2009, since it depends heavily on the weak property and employment markets in the developed world”, therefore they forecast an average annual growth of 6% until 2015. This means that (according to the estimate) the classified industry will account for 11% ($14.9 billion) of the global internet advertisement spending in 2015.

In comparison, AIM Group, a leading consultancy in classified advertising, describes the industry (including printed advertisements) as a $100 billion annual global market (2012a), and according to their Classified Intelligence Report (2012b, p. 6) “online classifieds have shown continuous double-digit revenue growth and remained robust during the 2008-2009 financial turmoil”.

Estimating the value of the traded goods and services turns to be an even more difficult task, since most of the Classified Sites are providing free listing options, moreover it is also hard to track what is actually sold, but the final number would probably exceed several thousand times the industry’s annual revenue.

Additionally – as whole E-commerce industry grows rapidly in these regions (eMarketer, 2013) – it seems to be certain that online classifieds have a substantial potential in developing markets like Asia-Pacific.
2.1.3 General Characteristics

This section will discuss the general characteristics of Classified Sites primarily from the Recommendation Systems’ point of view. If we consider this, there are three main elements: Items, Users and Events.

**Items**

An ‘Item’ is an advertisement on a Classified Site, uploaded by a ‘User’. An Item is described by multiple types of metadata like title, text, price, location, uploader, and by other Item specific information (e.g. number of rooms). Most Users are also uploading pictures about the Items to make their advertisements more attractive.⁷

Depending on its maturity and market reach, a Classified Site can list several hundred thousand or even millions of completely unique Items. It is also important that these are rotating relatively fast: daily up to 2-3% of the Items are replaced on a mature site⁸. Another challenge regarding Items is that they do not disappear when they have been bought. Unlike traditional E-commerce engines or Auction Sites, here Users themselves have to mark their advertisements when they sell something and – even though many operators use various incentives – this usually does not happen immediately afterwards.

In general, there is an ‘N:1’ (many to one) relationship between Items and Users, so a User can upload, buy or sell any number of Items, but an Item has just one owner. This unfortunately does not stop sites listing a significant amount of duplicates. It usually happens at real estate classifieds, when a property is advertised by multiple agencies.

An Item or advertisement can be about almost anything from electronics to real estate, and some Classified Sites even list (professional) services or actually Users.⁹ The Items are organized by categories and these are ordered according to a category tree. Most sites are building up their classes in a way that every parent category contains between 4-9 sub-categories in order to make the navigation easier.

**Users**

‘Users’ are the visitors of a Classified Site. A User can both browse existing Items and add new ones. One of the biggest challenges of E-commerce is that, although the Users and their goals can be completely different, unlike at traditional retail stores it is no quite easy to segment

---

⁷ Among these images there is one, called ‘Thumb’, that appears in lists by default.
⁸ That means if a site lists 1 million Items, then 30.000 are changing on a daily basis.
⁹ Dating Sites
them based on their appearance or ask them what they would like to buy. On the other hand, even Users that are experienced in online shopping can be lost among millions of Items.

From the Operators’ point of view there are two basic types of Users.

Unregistered Users are usually identified by a Cookie, therefore – with the help of appropriate tools – the Operators can follow them through multiple browsing sessions. However, to make this possible, Users should not block Cookies in their browser, moreover they should accept the site’s Cookie Policy. But applying this solution still does not give the opportunity to recognize the same User if he is browsing from multiple browsers or devices (e.g. home computer, work computer or smartphone). Cookies also expire after a time. Regardless of the fact that unregistered Users are mostly unidentified, many Classified Sites provide them with the ability to view Items, contact sellers or even to advertise their products.

Registered Users are willingly accepting the site’s Terms of Use and identify themselves with at least an email address. Other relevant data can be gender, date of birth, city or ZIP code. If they are logged in to the site, registered Users can be tracked regardless of the browser or device that they are actually using. Registering to a site is also often required to access some services such as email alerts or browsing history.

In general, the market value of a Classified Site mostly depends on the number of Users and Items.

Events
‘Events’ (or transactions) are the main connection and interaction points between Users, Items and the platform itself. Users can carry out various activities on a Classified Site, and each of these transactions can be tracked. Event types can be for example: view, click, search, rating and adding to favorites, among many others.

By tracking Events, Operators can get an overview about the general usage of their site, furthermore – if they have the proper applications – they can measure the performance of certain advertisements or study the behavior of specific Users or User segments.

In conclusion, logging Events is a simple way for Operators to observe the functions of their site and to receive feedback about their service, and thereby to perform the needed changes in order to achieve their business and marketing goals.

---

10 An Operator is somebody who owns or operates a Classified Site. They are also often referred as ‘service providers’. 
2.1.4 General and Specialized Classifieds

General Classifieds are sites that are not focusing on a specific advertisement category, but rather they try to be the market leader (in terms of the number of users and items) by listing any product or service.\(^{11}\)

Among General Classifieds, beyond the earlier referred Craigslist, it is important to mention the South African Naspers that is active in 107 countries with its classified brand OLX (Naspers, 2014b). Naspers also owns major E-commerce sites in the Central and Eastern European region through the Polish Allegro Group (Naspers, 2014a), which is present in 25 countries (Allegro Group, 2013). Additionally, the Norwegian Schibsted Media Group, operating General Classifieds in 29 countries (Schibsted ASA, 2010) is also notable. As many examples indicate, most of these major groups’ business strategies are quite simple: when they are entering a region, they are investing 3-5 years and enormous amounts of resources in their local subsidiaries, in the hope of once completely controlling the market.

Specialized Classifieds, in a sense, are the exact opposite of their general counterparts. They are focusing on a specific item category, and they often provide unique services and develop several additional functions in line with their range of products and target group. Trulia,\(^{12}\) for instance, makes it easier for its users to find their new home by building its service on an interactive map extended with additional information like crime rates, schools or amenities.

Specialized Classifieds can be divided along three traditional categories or classes: real estate, vehicles and jobs.\(^{13}\) It is beneficial to specialize on these categories since they are attracting the majority of the users therefore they are usually much easier to monetize (eNET-CUB, 2014).

In comparison, General Classifieds are focusing on ruling the market by attempting to reach every internet user through covering all product segments at once, but in most cases they are also forced to cross-finance their additional product categories with the help of the traditional classes. Additionally, it is not surprising that specialized sites can usually provide users with a better service, as they are much more focused on their visitors’ desired range of products.

Furthermore, although there is no particular research that would confirm this statement (therefore mainly based on the products listed), Specialized Classifieds tend to have a greater proportion of B2C ads, while General Classifieds are more significant in terms of C2C offers.

---

\(^{11}\) Naturally there are exceptions to this, like adult services.

\(^{12}\) http://www.trulia.com/

\(^{13}\) Some other common categories are: travel, electronics and dating.
One possible reason for this is that while private sellers often advertise on an ad-hoc basis, business advertisers usually concentrate on certain product classes (like real estate agents, auto parts dealers or employers), therefore they are looking for a more specific target group, which is easier to find on non-general sites. This may also contribute to the easier monetization of Specialized Classifieds.

2.2 Revenues and Business Models

2.2.1 Monetization

One of the biggest challenges in monetizing users on Classified Sites comes from the basic nature of these applications. In general, Classifieds only facilitate transactions in a sense that they are providing a platform to advertise and browse offers. The value of these transactions does not appear in the sites’ cash flow, as the deals are finished independently of the platform. The role of Classified Sites is to connect buyers and sellers, thus the final goal, conversion – which is known as ‘Ad Reply’ – does not generate any revenue that could be the basis of taking commissions.

Another challenge is fueled by a common Internet phenomenon: people generally do not like to pay for services on the Web, and it is not different in the case of online classified advertising. Unlike printed advertisements, users do not feel that they receive a tangible service by displaying their ads, since nowadays nothing is easier than publishing something online. This feeling is probably also supported by the Web 2.0 experience,14 moreover by the huge amount of sites that offer a wide range of advertising options for free.

Furthermore, the competition is not just about local challengers. As the Internet is a global platform, in many cases it is not possible15 to get around eBay, Amazon16, or the local sites of Schibsted Media, whose positions are secured with vast financial backgrounds. Additionally, if we consider that the form of classified advertisements are quite exempt from local characteristics, moreover shipping options are expanding (e.g. Amazon Prime Air17), and many sellers are operating with free or flat shipping rates,18 things are not getting easier for local start-up’s.

14 On a platform where users create most of the content, why exactly would they pay for the opportunity? Undoubtedly, without users and User Generated Content Classified Sites would not worth much.
15 Although this is mainly true for developed markets.
16 http://www.amazon.com/
17 http://www.amazon.com/b?node=8037720011
18 92% of the top 50 e-retailers studied by Forrester (Enright, 2013)
Consequently, Classified Sites should not be focused on monetizing their general, just browsing users or buyers in the first place. In fact, as these customers are forming the basis of classifieds, they and the quality of their user experience should be the first priority for the operators when considering the possible sources of revenue. Additionally, it is also important to note that sellers are the ones who are actually earning money as a result of advertising and trading their products, therefore also they are the ones who can be charged for using the site.

Many classifieds, such as the prominent Craigslist, are going even further by not charging C2C transactions at all, which means that individual advertisers can post any number of ads for free, while the site is maintained from the payments of business sellers (DealBook, 2006).

However, operating Classified Sites is still a profit oriented business, therefore the following sections will introduce some of the means of making online classifieds successful, also in terms of financial results.

2.2.2 Possible Sources of Revenue

**Fixed price per advertisement**
This is the most basic source of revenue. As some their offline equivalents, some online classifieds also charge a fixed amount of money per advertisement.

**Featured advertisements**
Most classifieds give their sellers the opportunity to highlight their advertisement by a different background color, by better rankings in listings or by displaying their featured ads in special galleries. It is also not rare that operators create special advertising packages (such as bronze, silver, gold, platinum etc.) in order to adapt to every need.

**Extended advertisements**
Some sites offer extended advertisement forms, for example longer descriptions are allowed, links can be embedded or more and higher resolution pictures can be uploaded.

**Value-added services**
Value-added services can be alerts (e.g. SMS service), accessible statistics about the ad performance and comparison with similar ads, among others.

**B2C advertising packages**
Mostly Thematic Classifieds offer business packages (for company sellers and agents) including extended services like ad management tools or customizable sub-pages.

---

19 However, it is more likely that newspapers charge per line.
Subscription
Dating sites for example tend to have a periodical subscription or a one-time registration fee.

Display Advertising
Displaying banners from third-party advertisers is a possible way of indirectly monetizing users, while maintaining smooth user experience. Therefore this is an important source revenue for many sites, especially in an early stage of maturity when a platform might already attract a lot of visitors but operators are still struggling with monetization. Advertising on Classified Sites can be particularly effective if these ads are customized according to the content of the actual site, since apart from higher response rate, customers will also feel that they are provided with relevant offers. For example, promoting au pair services next to baby furniture listings might be a profitable option.

2.2.3 Business Models
Most online classified business models are based on the various combinations of the above mentioned sources of revenue. Furthermore, many operators provide private and business sellers with different opportunities, therefore this serves as the main organizing principle for the following categorization.

C2C Business Models
Free
In this model, the full range of the services is provided free of charge. The primary goal is improving the quality of services, building loyalty among users, furthermore raising the number of users and items, thus increasing market share.

Freemium
Freemium means that the basic services are provided free of charge, but the access to the full functionality costs a certain amount of money per ad.

The sites operating with a Free or Freemium model also tend to spend a significant amount of money on building their brand, since their goal is not (yet) to achieve high profitability.

In accordance with the above, the Free and Freemium models are usually applied only temporarily (until reaching a certain market share), or only for some categories cross-financed by others.

20 It is not uncommon that they are investing in offline ATL (Above The Line or mass media) advertisements and TV commercials.
**Premium**

Premium, value-added services (e.g. highlighting, alerts, statistics) are offered for a certain amount of money per ad or for a subscription fee. Free advertising options are firmly limited in this model.

**Subscription**

This means that multiple one-time or a periodic subscription options are offered in order to provide different levels of services.

**B2C Business Models**

*Fixed price per advertisement*

This business model is mostly used by job sites targeting minor employers who advertise their vacancies irregularly.

*Scalable business packages*

This model offers a range of business packages in order to adapt to the needs of small, medium and large advertisers as well. Some packages include the option to create a company branded sub-page structure that operates as an own catalog.

2.3 **The Maturity Stages of Classified Sites**

It may be important to differentiate Classified Sites (beyond specialization and business models) according to their maturity, since – as in case of many other industries as well – this indicator helps a lot in understanding these businesses better.

Although there are a handful of approaches for determining the maturity of a business, on the one hand the scope of this paper might not be wide enough to introduce a model in detail, and on the other hand the accessible models might not be specific enough to study the classified businesses precisely.

These reasons gave the inspiration for creating the following categorization, based on the experience and common knowledge available at Gravity R&D, a Recommendation System Provider company, which has served a great variety of altogether more than 20 Classifieds over the past few years. The classification is also illustrated by Figure 1 (Created by the Author), which shows how the revenues of a typical Classified Site change over the time according to its maturity.

---

21 And neither would that particularly contribute to the goal of this thesis.

22 The initial inspiration for the categorization comes from Boston Consulting Group’s growth-share matrix (Doyle, 2011).
**Immature**

Classified Sites at the Immature stage are considered mostly as question marks. Although most of their main functions are operating, these are usually quite limited and they are still under development or testing.

Apart from developing functions, one of the biggest challenges here – and it is not so much different from other businesses – is finding the first clients. The operators are usually obtaining their first advertisers by approaching and partnering potential B2C advertisers that have a significant operation (e.g. real estate agents) based on unique agreements.

At this stage, the sites are only used by a number of early adopters who become aware of the new service mostly thanks to the ‘word of mouth’ phenomenon, as operators are not yet spending on major advertising. Additionally, apart from promoting the site, first users have another – and equally important – role of providing the operators with relevant customer feedback.

---

**Worth Spending On**

If the first results are promising and the site will potentially succeed on the market – based on the previous operational experience, on the customer feedback and on the number of initial sellers – it is time to go mainstream and reach out for as many users as it is possible.
Some say that the competition at this stage is clearly about building brand awareness while increasing the number of users and items at any cost, therefore operators with a stable financial background tend to spend huge amounts of money on advertising. Hungary’s two leading General Classifieds Sites, ‘Apród’\(^{23}\) (Allegro Group, Naspers) and ‘Jófogás’\(^{24}\) (Schibsted Media) are notable examples of this phenomenon, as they are harshly competing since 2011 to dominate the Hungarian market. Although they are not publishing many details, thus there are only estimates for the ‘size’ of the sites, it seemed that until August 2012 their traffic improved similarly, then during 2013 Apród/OLX took the lead (eNET-CUB, 2013) and the difference has just increased since then (SimilarGroup, 2014). Both of the sites have spent immense amounts of money on advertising – both of them ran TV and outdoor (billboard) campaigns as well – without actually generating any income or having an operating a business model (Szalay, 2012), thereby making huge losses (eNET-CUB, 2014). Eventually, it seems that until now the operators of Apród/OLX have been spending their money somehow better, although Jófogás has just acquired some of their smaller competitors (Schibsted ASA, 2014).

However, it is not quantity that only matters. Business Unit Manager of Apród/OLX claims that laying a special emphasis on the quality of the ads, checking every newly added, and automatically deleting all outdated items are the keys to stand out (Szalay, 2012).

Another essential point here is to provide users with a clear and easy-to-use interface, furthermore with plain item categories and with an efficient search function, thus creating a seamless browsing experience.

Although increasing brand awareness and developing user experience are both vital, another crucial step during this phase is to determine the means of revenue generation by adapting the business model that will lead to financial sustainability and eventually to success. Until this point, most sites’ initial revenues are mainly coming from third-party display advertising and partly from unique arrangements with significant B2C sellers, therefore the first paid services (still aiming business sellers in the first place) are usually introduced at this stage of maturity. Recommendation Systems are also deployed around this period in the hope of ensuring certain competitive advantages.

The phase lasts until reaching a certain market share (this means being among the top 3-5 players, depending on the size of the market), while optimistically incomes will also start balancing expenses till the end of the stage.

\(^{23}\) Apród has been recently renamed to OLX (http://olx.hu/) as Naspers is currently working on unifying the brand of its sites.

\(^{24}\) http://www.jofogas.hu/
Sound Market Reach

Sound Market Reach means that the Classified Site achieves stable growth and a more or less leading position on the market. Incomes are continuously growing during this period, but in order to make the business profitable, the spending on advertisements should be also decreased, which can have an adverse effect on the achieved market share. Certainly, one of the biggest challenges in this phase is to find the right balance between profitability and growth.

The main functions and the business model of the sites are usually finalized at this stage, while the range of services is typically widening with more premium solutions. Therefore, in addition to the new customers acquired as a result of advertising and referrals, cross-sells and upsells become the most important drivers of revenue growth. The period lasts mainly until a significant part of the customers (including some private advertisers) will become willing to pay for these services thus securing the profitability of the site. However, as the Hungarian situation also indicates, massive monetization (which means charging for C2C ads too) in the classified industry is only possible after becoming the market leader in a given region (or at least after capturing a significant share of the market), since users will not likely to pay for certain services until they can turn to other strong alternatives that offer lower prices or does not charge at all.

Consequently, one of the most important tasks here is first finding the optimal balance, and then gradually shifting it from free services (which aim to build customer loyalty) towards paid and premium features (which represent the main sources of revenue), without losing the leading position. As the related section will also indicate, Recommendation Systems can be a major help in easing this process and speeding up this phase.

Monetizing and Maintaining

Reaching these stages means that the site came out as the winner of the intense competition, and it is more or less able to permanently control the market. But still, after a time, revenues’ growth rate will reach its maximum and after slowing down, incomes will eventually slightly decrease. However, there are a number of areas that operators can focus on in order to keep their sites around the top of its profitability as long as it is possible.

The main keys of keeping up against strengthening competitors and disappointed customers are: the constant improvement of the general ad quality, the continuous development of the key functions and the introduction of additional comfort features and premium services. Moreover, Business Intelligence solutions tend to have a significant role at this point, since...
monetization can be greatly supported by analyzing user segments and providing unique offers according to the customers’ (current or potential) profitability. Consequently, long lasting success is mostly achievable by maintaining user loyalty and decreasing churn rate.

**Expanding**

In general, as the site is leaving the top of its profitability, operators have two choices. Either they are expanding their operation geographically to new target countries and regions, or they are gaining a foothold in new markets in terms of additional product categories.

In case of the former, operators are usually introducing an earlier version of the mature site and slightly adapting it to the new country’s reality, and then the whole process starts over. Another option is launching a white label product and licensing the platform (or the source code) and providing support and good case practices to a number of (foreign) partners.

In case of targeting new markets by broadening the scope of products, the first step might be to introduce new categories in order to cover all the potentially profitable user segments, but many examples show that there are a couple of further, more advanced options.

For example, Allegro Group (Naspers) in Hungary operates currently two Auction Sites, a Coupon Site, an Online Plaza and a Price Comparison Site apart from its classified Aprod/OLX. Therefore the group has a strong, leading market position in the E-commerce industry in Hungary, offering a vast variety of products and reaching millions of customers (Allegro Group HU, 2014).

However, there are even more sophisticated solutions other than dominating the industry with multiple E-commerce sites. Ingatlan.com, a major real estate classified in the CEE region, chose another way to scale up revenues. The main element of this strategy is to widen the portfolio with services that can be easily linked to the initial business unit. Providing online content for free (e.g. launching a news site) is more than appropriate for this purpose. If the quality is sufficient, a significant amount of visitors can be expected and they can be easily driven to the initial site by displaying relevant ads next to the content. Ingatlan.com developed this strategy even further when they launched Koponyeg.hu, a weather forecast website. Visitors here are naturally choosing their current location, of which’s weather they are interested in, and this information is also used to display real estate ads from the related regions and cities, where they are probably looking for new homes.

---

26 [http://koponyeg.hu/](http://koponyeg.hu/)
2.4 The Users of Classified Sites

Without having an overview of the typical users of Classified Sites and without determining the main stages of their life cycle, it would be difficult to understand the visitors’ needs towards online classifieds and their recommendation solutions.

2.4.1 User Groups

A research conducted by the Pew Internet & American Life Project (Jones, 2009, pp. 3-11), provides detailed insights about the demographic characteristics of the ‘average’ online classified user.

The research states that “the number of online adults who have used online classified ads has more than doubled” between 2005 and 2009, and the proportion of those who have ever used a Classified Site grew to 49%, while around 9% of all internet users visit Classified Sites on a typical day.

The report also reveals that age is a significant factor here: Internet users ages 25-34 are the most likely (62%) to visit online classifieds, followed by 35-44 year olds (57%), who are still ahead of 18-24 year olds (45%) and 45-54 year olds (49%), and finally – but not surprisingly – the least likely are the 55-64 year olds (35%) and older (26%).

According to the study, it is also important that “college graduates and higher income earners as well as urban and suburban internet users are more likely than their counterparts to use online classified ads”.

It is also noted that the findings about demographics seem to be stable over the time, “though younger age groups grew even more likely than older groups to use classified ad websites”.

2.4.2 User Life Cycle

User Life Cycle is a well-known phenomenon in many industries and it is also applicable to Classified Sites. The stages of this cycle help segmenting users according to their potential value and expected behavioral patterns. Also, users in distinct stages will probably interact with the site differently through dissimilar interaction points, which may justify the use of different strategies in monetization.

The following categorization – similarly to the previous one – is created based on the experience and common knowledge available at Gravity R&D.
Figure 2 (Created by the Author) shows the stages of the User Life Cycle along with the expected revenues over the time.

![Image of User Life Cycle diagram]

**Acquire**
First, operators have to drive users to their sites by acquiring visitors, which is possible by increasing brand awareness through investing in various advertisement platforms and promotions. Additionally the site has to offer a number of desired items, furthermore it has to possess certain qualities (e.g. attractive design), not to make visitors bounce\(^{27}\) and to make them proceed to the next phase.

**Activate**
Visitors have to be activated in order to create a positive return on the investment spent on the acquisition. This phase is mainly about building loyalty, therefore making the future customers believe in the service provided by the site. This can be achieved by a smooth user experience, by quality ads, by positive experience with the sellers and by a helpful customer service, among others.

Active users are visiting the site regularly, browsing categories, clicking on third-party banners, sending Ad Replies and hopefully soon they will upload their first advertisements and sell their first products.

\(^{27}\) Bounce means that the visitor leaves the site immediately after landing, without performing any events.
**Monetize**

The real value lies in those customers who advertise their products regularly and who – as an effect of the previously developed trust towards the site – are willing to pay for premium services.

Regular buyers are also valuable in a sense that they create the demand for the products advertised by sellers. Also, regardless from the fact that without them it would be difficult to satisfy sellers, they can largely support the site’s expansion by recommending the service to the members of their personal network.

**Retain**

Most C2C sellers have a limited amount of products to advertise (this might be also true for some B2C advertisers as well, e.g. for a small employer), and the activity of buyers is also decreasing after obtaining the wished items. Therefore the role of the Retain phase is to make sure that next time if they want to sell or buy something, they will use the services provided by the site again.

Going beyond promotions, some of the most effective tools for keeping users interested are Event Driven Marketing (e.g. email alerts) and Retargeting Campaigns.
3. An Overview on Recommendation Solutions

3.1 Recommendation Systems and Their Applications

3.1.1 Definition

According to Ricci, Rokach and Shapira (2011, p. 1) Recommendation Systems or Recommender Systems “are software tools and techniques providing suggestions for items to be of use to a user”. These suggestions help users carrying out certain decisions (such as what to buy, rent and watch, or which services to use) by providing and ranking a number of relevant alternatives. From another point of view, Recommendation Systems try to predict users’ responses (or the results of their decisions) in various situations by modelling their personal taste, knowing what they want before they actually do (Grossman, 2010).

3.1.2 Development and Application

The development of Recommendation Systems can be easily associated with the evolution of the Internet and E-commerce. Online sales made it possible to list a whole new variety of products and services thus online retailers now are able to offer several times more items than their offline equivalents. In line with this, some websites today provide probably more content than the whole Web 10-15 years ago – while clearly the number of sources is also growing dramatically – which eventually leads to a content glut, and finally to the ‘paradox of choice’ (Schwartz, 2004): the users become increasingly overwhelmed with information.

The main challenge is, if the most users remain interested in only a thin slice of this information, then how is it possible to gain exactly the needed particles? Unfortunately these needs often cannot be easily expressed or specified in the form of a search query, or simply the result is either beyond manageability or not informative enough to make a satisfying decision. That is when Recommendation Systems come into the picture, since – at least ideally – they are able to identify and answer the users’ actual need.

Recommendation Systems have already become an independent research area in the mid-1990s, and there have been also many conferences and workshops organized during the last decade dedicated to the topic, such as ACM Recommender Systems (RecSys), which was established in 2007. Additionally, several journals are covering the results of the research and development in this field – like the AI Communications (2008), the IEEE Intelligent Systems (2007) and the International Journal of Electronic Commerce (2006) – furthermore nowadays many institutions of higher education around the world offer popular courses related to this area (Ricci, et al., 2011).
Beyond the scientific interest, nothing indicates the significance of Recommendation Systems better than the websites and other Internet based services that are deeply integrated with, or actually built on this technology.

Based on a taxonomy and classification by Montaner, López, de la Rosa (2003, p. 287) and Ricci, et al. (2011, p. 14), Table 1 (Created by the Author) shows the classes and domains where Recommendation Systems are most commonly applied.

Table 1: The Most Common Applications of Recommendation Systems (Created by the Author)

<table>
<thead>
<tr>
<th>Classes</th>
<th>Domains</th>
<th>Examples and References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Entertainment</td>
<td>Music Streaming</td>
<td>Deezer (DEEZER, 2014)</td>
</tr>
<tr>
<td></td>
<td>Video Sharing</td>
<td>YouTube (YouTube LLC, 2014)</td>
</tr>
<tr>
<td></td>
<td>OTT Video on Demand</td>
<td>ivi.ru (Gravity R&amp;D, 2014a)</td>
</tr>
<tr>
<td></td>
<td>Interactive TV</td>
<td>SaskTel (SaskTel, 2012)</td>
</tr>
<tr>
<td>Content</td>
<td>Search, Discovery</td>
<td>Google Search (Kamvar, 2005)</td>
</tr>
<tr>
<td></td>
<td>News, Mass Media</td>
<td>Taboola (Taboola, 2014)</td>
</tr>
<tr>
<td></td>
<td>Social Media</td>
<td>Facebook (Falahi, et al., 2012)</td>
</tr>
<tr>
<td></td>
<td>Messaging</td>
<td>Gmail (Leichtberg, 2009)</td>
</tr>
<tr>
<td>E-commerce</td>
<td>Online Shopping</td>
<td>Amazon (Amazon.com, Inc., 2014a)</td>
</tr>
<tr>
<td></td>
<td>Classified Media</td>
<td>Schibsted (Gravity R&amp;D, 2014a)</td>
</tr>
<tr>
<td></td>
<td>Online Auctions</td>
<td>Vatera.hu (Gravity R&amp;D, 2014a)</td>
</tr>
<tr>
<td></td>
<td>Daily Deals</td>
<td>Bónusz Brigád (Gravity R&amp;D, 2014a)</td>
</tr>
<tr>
<td>Services</td>
<td>Holiday, Travel</td>
<td>RCI.com (Gravity R&amp;D, 2014a)</td>
</tr>
<tr>
<td></td>
<td>Leisure, Places</td>
<td>Foursquare (D’Orazio, 2013)</td>
</tr>
<tr>
<td></td>
<td>Dating</td>
<td>Randivonal (Gravity R&amp;D, 2014a)</td>
</tr>
<tr>
<td></td>
<td>Advertising</td>
<td>Google Ads (Google Inc., 2014a)</td>
</tr>
<tr>
<td></td>
<td>Expert Systems (e.g.</td>
<td>VITA²⁸ (Felfernig, et al., 2007)</td>
</tr>
<tr>
<td></td>
<td>Financial Services,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Logistics, Wholesale</td>
<td></td>
</tr>
</tbody>
</table>

As the above indicate, Recommendation Systems have currently an important role in many areas of life, especially in information and data-rich industries where the process of decision making is rather unstructured and the amount of alternatives is difficult to comprehend.

In addition to this, there are multiple reasons for which it would not be a surprise if the above classes would further expand.

²⁸ Virtualis Tanacsado: financial services recommendation environment
First, – actually the development of Recommendation Systems can be also initiated from this rather simple observation – it seems to be a very basic human characteristic that individuals often rely on recommendations provided by others even in making routine, daily decisions (Ricci, et al., 2011). It is really common that the personal experience, suggestions and opinions of people we may not even know or admire play a much more influential role in our decisions than proven facts, while it does not really matter if it is about looking for a coffee place, buying a computer or coming to a conclusion on long term investments. We might call this ‘the power of recommendations’.

On the other hand, industries that have been lately starting to capitalize on ‘Big Data’ and Business Intelligence solutions (such as healthcare and the energy industry) may get to the point where real time data analysis and recommendations will be essential for certain decisions, in order to maintain competitive advantages and an efficient operation.

3.1.3 The Role of Personalization

It is important to note that non-personalized recommendations have already existed before the development of Recommendation Systems – only they were based on simple statistics or on the expertise of professionals – in the form of bestsellers, top lists and other selections (Ricci, et al., 2011).

There are two main problems with these: firstly, they are only representing the general taste of the majority, therefore they are not providing suitable alternatives for many individuals. Furthermore, they are not taking into consideration the large number of items that are not able to get into the mainstream. This may finally result in the ‘long tail’ problem (Anderson, 2004), which means that a large amount of products are sold in small quantities, in contrast with a small number of best-selling products (Stevenson & Waite, 2011).

The long tail phenomenon is also a major challenge for Recommendation Systems, since – as mentioned earlier – a mature Classified Site for instance can expand with thousands of new items on a daily basis, thus it is not surprising that there are many studies addressing this issue in detail (Park & Tuzhilin, 2008).

Nevertheless, non-personalized suggestions still play an important role in modern Recommendation Systems as well. The so called ‘cold start’ problem – when user data or item related ‘view’ and ‘rating’ events are not yet available for generating suggestions – can be addressed with non-personalized recommendations, among others.
In contrast, personalized recommendations are based on unique user profiles generated by tracking visitor behavior and by identifying similarities and differences, making it possible to tailor customized recommendations according to every individual’s taste separately. Naturally, the solutions that are providing recommendations on the level of the individual (not on the level of certain user segments for example) have to be built up from more sophisticated algorithms that may need much more time to process. Therefore, another important challenge here is to serve recommendations in real time (below a few hundred milliseconds, in order to provide users with a seamless experience) considering continuously changing datasets that are containing millions of items and hundreds of thousands of users.

Later on we will mostly deal with the personalized form of recommendations.

3.1.4 The Main Functions of Recommendation Systems

After all, why are Recommendation Systems so essential in the E-commerce industry? Basically, they are addressing really important issues related to two main groups: users and operators.

User related functions

It would be difficult to evaluate Recommendation Systems without understanding the goals for which they are being used by the most important end-users, customers. Herlocker et al. in a paper (2004, p. 9) – that has become a classical reference in this field – define two main user functions that a Recommendation System can serve.

Annotation in Context

Highlighting or emphasizing some items in an existing list or other context. The context itself does not change (both the list of items and their order remain the same) but the most desired items are annotated. Additionally, in some cases the most undesired ones are downgraded (or filtered out). This is useful in situations where it is important to keep a list in its original form (for example categories A-Z) but due to the obscurity of the context (high number of items in the list) it is vital to help users distinguish between relevant and less relevant content.

Find Some or a List of Good Items

Suggesting a number of specific items to individual users, or recommending a list of items ordered by relevance. During the process, each and every item in the recommendation scope is provided with a certain value by the system (also known as ‘predictor value’) that indicates the current relevance. In the former case the top items (or optionally a number of randomly
selected ones from a bigger pool of top items)\textsuperscript{29} are displayed for the user. In the latter case the whole list is displayed. This is the core function of Recommendation Systems and it recurs in a vast variety in various applications.

Thinking both the above further and focusing mainly on the functions that are most commonly related to Classified Sites, Recommendation Systems can effectively help users in the following situations.

\textit{Lack of Knowledge}

It is not rare that users do not exactly know what they want. They sense a problem or need (the recognition of these can be also quite uncertain) but they cannot define the exact method to get to the solution. Furthermore, even if the satisfying solution is known, they cannot express or specify it in the form of a search query. This might be due to the complexity of the problem, to the inexperience of the user or due to a poor user interface.

\textit{Just Browsing}

This case might be the effect of the previous one, or separately – as it was also defined by Herlocker et al. (2004) – the users are simply happy to browse, they do not recognize any need or problem, and most importantly they have no initial intention to purchase anything.

\textit{Too Many Alternatives}

It may occur that the result of a search query is beyond manageability, thus the users eventually have no time or capacity to process the information, so they cannot turn it into knowledge in order to carry out a satisfying decision. Too many alternatives can easily lead to the ‘paradox of choice’ (Schwartz, 2004).

\textit{Making Poor Decisions}

Even if the users are conscious about their choices and they think they are making the right decisions, it is possible that they are continuously doing the opposite, thus on the long run they become unsatisfied, finally blaming the service in the first place.

\textit{Lack of Diversity}

There might be problems that certain users were previously unable to solve, moreover some needs might be constantly present. In both cases it can be a struggle for many users to find new and different, but still relevant solutions without being exposed to browsing through many irrelevant ones.

\textsuperscript{29} This aims to increase diversity, and nowadays it used in most of the recommendation scenarios. This way it is unlikely to display the same set of suggested items more than once.
Operator related functions

Users are the main target group of recommendation solutions, but obviously operators are the ones who are deciding on spending remarkable resources on implementing any of them, therefore (apart from satisfying user needs) there must be strong benefits and clear performance indicators associated with these systems. The following list sums up the findings of Ricci et al (2011, p. 5) briefly.

Increase the number of items sold

Unsurprisingly, the most important goal of a Recommendation System (from the operators’ point of view) is to increase sales related performance indicators and ultimately the number of items sold. This can be achieved through multiple sub-goals, such as – if we stick to classified example – lowering bounce rates while also preventing users from leaving the site, increasing CTR’s\(^{30}\) and Ad Views, and eventually scaling up conversion rates. The methodology behind this (apart from prolonging browsing sessions) is that the recommended items are much more likely to suit the users’ needs and wants, therefore customers will buy more and more frequently.

Sell more diverse items

Another major function is addressing the long-tail phenomenon. Even though the operators are interested in selling all the items listed (or providing the sellers with the opportunity to sell any of their products, not just the popular ones) some niche items may be difficult to find and therefore to sell. Through proper recommendations it is also possible to promote unpopular products to the right audience without being exposed to the risk of suggesting something undesired to the majority of the users.

Increase user satisfaction

As mentioned earlier, Recommendation Systems can fulfill a number of user related functions. If the solution is designed well, users will find interesting and relevant recommendations that will positively affect their subjective evaluation of the system.

Increase user loyalty

The increasing user loyalty is mostly an effect of satisfaction, moreover the longer the users interact with the site, the more refined user profile can be built up, and the more the recommendations (and thereby the whole user experience) can be customized according to their individual taste. Additionally, valuable visitors can be also targeted with special offers.

\(^{30}\) Click-Through Rates
Understanding user needs

Operators can utilize the enormous amount of data and the detailed user profiles (either collected explicitly or predicted) in many ways apart from serving recommendations. For example, the website or the application can be redesigned (and dynamized) in order to match to the unique needs of certain user segments, furthermore it is possible to support off-site (or even offline) business and marketing goals, and the capacity and stock management can be also simplified, among others.

Common Recommendation Scenarios

In order to present more tangible use cases, this section will introduce the four possible combinations of recommendation scenarios between users and items. This list and the examples are based on common examples at the company Gravity R&D (2013).

For a given item, provide the top-N recommended items

This is the standard item-to-item recommendation use-case. Items can be similar (high similarity), related (some similarity) or best next items (items that the user will most likely to click on after the initial item).

For a given user, provide the top-N recommended items

This is the main scenario for personalized item recommendations. It can be optimized according to various business goals such as conversion, discovery, serendipity, etc.

For a given item, provide top-N recommended users

This user-to-item scenario is mainly used for direct marketing purposes, such as personally targeted promotional emails.

For a given user, provide Top-N recommended users

For privacy reasons this scenario is usually not provided to end users but it can be efficiently used in back office applications to fight against churn or to optimize marketing campaigns.

Some of the solutions – as Gravity R&D’s – also support forming virtual items. This means that all the scenarios (both in case of the referenced and recommended subjects) can be also applied to groups of users and items. These groups can be defined based on any metadata or contextual information. Therefore it is also possible to recommend items to a selected item group (‘basket recommendations’), or – although it is not very realistic – two or three bedroom apartments and houses in a given price range from a specific region to female users who are browsing a given category between 4-6 p.m. from a mobile device.
3.2 How Recommendation Algorithms Work

This section will briefly introduce the very basic concept and technical background of Recommendation Systems. It is important to note that the aim of this thesis is far from presenting the technology in details, the goal is rather to provide exactly the needed (or at least a satisfactory amount of) information, which is essential for evaluating the utility of these solutions in a general E-commerce environment.

3.2.1 Basics

Ricci et al. (2011, p. 2) sum up the main basics precisely: “In their simplest form, personalized recommendations are offered as ranked lists of items. In performing this ranking, RSs\(^{31}\) try to predict what the most suitable products or services are, based on the user’s preferences and constraints. In order to complete such a computational task, RSs collect from users their preferences, which are either explicitly expressed, e.g., as ratings for products, or are inferred by interpreting user actions. For instance, a RS may consider the navigation to a particular product page as an implicit sign of preference for the items shown on that page.”

Suggestions are generated mostly in real-time upon (an indirect) request of the user (or its browser), based on the applied recommendation scenario, and on the given user’s current context and needs, utilizing various item and user related metadata, and taking into consideration previous events and external data sources, among others. After displaying the recommended items, the user is able to browse them and decide whether they are satisfying or not, providing an explicit or implicit feedback. Furthermore, all these user actions and feedback events are stored in a database and queried later on, in order to generate accurate recommendations in further user-system interactions, presumably also for different users and items (Ricci, et al., 2011).

3.2.2 Data and Methods

As mentioned earlier, users, items and events, furthermore the data that describes these as detailed as possible, are the most important elements for Recommendation Systems to generate proper suggestions. In this section, this statement will be slightly refined based on the main components of the technology applied for serving recommendations at Gravity R&D.

The types of collected data and the related recommendation algorithms can be classified into the following major categories.

\(^{31}\) Recommendation or Recommender Systems
**User Behavior**

User behavior is mainly analyzed by tracking codes embedded to the operators’ website. These are logging all the on-site activities, such as clicks, searches, page and product views, viewing times, ratings, favorites, etc. Additionally, there are certain methods for tracking user behavior off-site. An example is tracing clicks in emails, in mobile applications and in their push notifications (if available) moreover affiliate sites and ad networks can provide further options with the help of third-party cookies.

User behavior related data is typically processed with Collaborative Filtering methods. The main concept here is studying the correlation between users: if a user liked some items in the past, other users with similar taste will probably like the same items. The similarity between user profiles is calculated based on the similarities of the users’ behavior. Collaborative Filtering is probably the most widely implemented technique among Recommendation Systems (Ricci, et al., 2011).

Some of the most common Collaborative Filtering methods are the Nearest-Neighbor Algorithms\(^{32}\) and the Latent Factor Models, like Matrix Factorization\(^{33}\) (Sarwar, et al., 2000).

**Item Details**

Item details can be any metadata that describes an item, such as title, category, price, description, location, owner, inventory, ratings and other product related information (e.g. number of bedrooms).

Item related metadata is mostly utilized by Content Based Filtering methods. Although these techniques are less sophisticated, they fulfill a significant role in case of new items (that have not really occurred in user histories yet), but most importantly they can serve accurate item-to-item recommendations. These can be used to suggest items that are similar to the ones that a certain user liked in the past\(^{34}\) (Ricci, et al., 2011) and also to generate ‘you may also like’ type recommendations on product pages.

---

\(^{32}\) For example Pearson correlation
\(^{33}\) For example Singular Value Decomposition (SVD)
\(^{34}\) For example if a car from a certain brand has been added to the favorites, other cars from the same or similar manufacturers might be also interesting for the user. Furthermore, by comparing the selected car’s price to the average prices in the category, the system might be able to determine the price sensitivity of the given user, which can be used for recommendations in other categories. Therefore, if the car was relatively expensive (e.g. falls into the seventh decile), the suggested apartments should be relatively expensive as well.
Contextual Information

Contextual information is particularly important to refine the main recommendation algorithms. Some patterns in a user’s behavior may differ periodically according to the time of the day, the weather and the general mood, and additionally some items may be affected largely by seasonality. However, the used device, the current location (and its change) and the referral URL – which indicates the source of the traffic and in some cases the original search query as well – play an even more substantial factor when studying user activities. A well-designed solution has to take these influences also into consideration\(^\text{35}\) and serve recommendations accordingly.

Social Events

As social networks have become the major players of today’s internet (both in terms of the number of users and the time spent on site) Recommendation Systems outside social media platforms also have to react to the related challenges and turn them into opportunities. The main point is that users carry out several activities,\(^\text{36}\) furthermore they are having their friends\(^\text{37}\) on social networking sites, and both of these can strongly contribute to more accurate suggestions. The only problem is that operators of social sites – not unexpectedly – do not give away this information for free since their revenue is largely based on it. However, there are ways to access the social data of a site’s visitors by linking their social profile to their user accounts with the help of social media applications. Depending on the platform, users might have to confirm this access and – making the situation more complicated – passing the data to third parties (such as Recommendation System Providers) can also raise privacy issues.

Although Collaborative Filtering and Content Based Filtering are the most fundamental methods, – as the above categories indicate – modern Recommendation Systems operate as hybrids of these approaches by combining different techniques enhanced with various knowledge sources to provide operators with a highly valuable output.

3.2.3 Architecture and Integration

This section aims to present an overview on the very high level architecture of a Recommendation System and it also intends to briefly summarize the main tasks to be carried out for implementing such solutions.

\(^{35}\) This can be done either by post-filtering the suggestions from the output that are not matching to the given context or by using contextual data in the prediction model explicitly (Ricci, et al., 2011).

\(^{36}\) Such as likes, shares, retweets, pins, etc.

\(^{37}\) On whose recommendations, past purchases and other interactions they might rely much more.
Architecture

In order to help the simpler comprehension and interpretation of the above discussed topics, Figure 3 (Gravity R&D, 2013) is showing a sample, high level structure of some of the main recommendation methods used at Gravity R&D.

![Diagram of Recommendation Methods](Gravity R&D, 2013)

The illustration indicates that (in this example) the contextual data is taken into consideration in a phase called Intelligent Weighting after processing all the other available information with the company’s core logic, which consists of a number of different methods. After that, the output is transmitted to different post-processing modules such as the Business Rule Engine and the search function. The results of the recommendation process are then presented with the help of various interfaces on websites, in emails, in applications running on certain devices, in Business Intelligence systems and on other platforms.

Integration

Today, although many websites and online services are starting their initial operation by already applying a deeply integrated, internally developed Recommendation System, there still

---

38 CF stands for Collaborative Filtering, while CBF means Content Based Filtering
39 The Business Rule Engine is responsible for representing and realizing different business and marketing goals in recommendations.
40 API means Application Programming Interface and SNMP stands for Simple Network Management Protocol
41 STB stands for Set-top box (TV)
seems to be a great need for outsourced solutions developed independently by specialized companies. These are implemented usually after the initial service reaches certain maturity, and they can be highly customized according to the given site’s needs.

Generally, an implementation process consists of the following steps (Gravity R&D, 2014b).

**Definition**

As a starting point, the integration should be based on clearly defined business requirements, furthermore explicit KPI objectives and transparent measurement methods should be set in order to properly evaluate the solution’s contribution to the organizational goals.

**Technical integration**

The process of the technical integration can be divided into three main parts.

1. Some initial data should be transmitted towards the Recommendation System Provider. The most important elements are the item catalog (the database of the items and related metadata) and the user catalog (the database of registered users along with demographic information, if available). Furthermore, if they have been logged on the site previously, user events can be also very helpful.

2. Setting up the continuous synchronization of the item and user catalog, moreover inserting tracking codes to the website in order to start collecting behavioral data.

3. After finishing with the first two steps of the technical integration, recommendations are ready to be displayed on the site. A number of boxes should be placed to the different page types and recommendation scenarios should be set up according to the business requirements. The methods for requesting and displaying suggestions can be divided into two main categories. Client side integration is when the user’s browser communicates directly with the recommendation engine (for example through a JavaScript API). In case of the server side integration; however, the operator’s E-commerce server is the one that communicates with the recommendation engine (for example through a PHP API), thus the operator’s server (that is accessed by the browser) requests and displays the suggestions.

**Fine-tuning and evaluation**

Not surprisingly, after the solution is implemented, there is a need for regularly evaluating and continuously fine-tuning the Recommendation System’s performance.

---

42 Privacy is an important issue when considering Recommendation Systems. Most of the solutions manage users anonymously without misusing any of their sensitive data (such as name, exact address or links to social profiles), identifying them with a single string. Demographic data (such as age, gender and region) are needed to make recommendations more accurate.
4. Recommendation Solutions on Classified Sites

After getting an overview on both the classified industry and Recommendation Systems, this section will introduce some of the most common applications of recommendation solutions on Classified Sites. Where it is available, the use cases will be illustrated with data charts (provided by Gravity R&D) and wireframes (created based on actual website designs), which aim to help comprehending the general implementation of these techniques better. However, due to confidentiality reasons, the exact sources of the examples will not be specified.43

Some findings of this section will also appear on Gravity R&D’s corporate blog in a number of forthcoming posts (Gravity R&D, 2014d).

4.1 Categorization and Measurement

In general, the final goal of implementing Recommendation Systems on Classified Sites is increasing conversion, but by breaking this goal down to more detailed objectives, it might be also important to categorize the main applications of these solutions around the actual issues that are the most relevant at the different maturity stages of classifieds.

---

43 This is also the reason for using wireframes and hiding the actual dates from data charts. Some of the data presented in this thesis comes from websites that are treating Recommendation Systems related information as sensitive, therefore publishing these is only possible by not mentioning the actual websites, moreover by hiding the exact dates and by presenting simplified illustrations instead of actual screenshots, thus making the sources unidentifiable.
Therefore, the following classification is based on the earlier introduced maturity model, which has been extended with some of the most general Recommendation Systems related functions, as it is visible on Figure 4 (Created by the Author).\textsuperscript{44} Similarly to the initial model, this version is also created based on the experience and common knowledge available at Gravity R&D. The description of the solutions will mosty focus on the placements of the suggestions, furthermore on the primary recommendation scenarios behind these.

To indicate the efficiency of the solutions in each category, the following examples are also supplemented with some actual data about the Click-Through Rates (CTR) measured on the displayed suggestions. CTR is a percentage value and – in this case – it is defined as the number of unique clicks on a set recommendations divided by the number of times a set of recommendations is requested. CTR is usually measured separately by placements or page-types. It is also important to note that according to the applied measurement, if a visitor clicks on multiple recommended items from a given set of recommended items, it still counts as one unique click, furthermore the number of recommendation requests will also increase if the set of recommendations have not been visible for the visitor (e.g. the suggestions are displayed at the bottom of the page and the visitor does not scroll down).\textsuperscript{45}

CTR is used as the main performance metric in this thesis for two key reasons. First, it is always available, while it is difficult to estimate – at least in case of the classified industry – for example the sales amounts or the Recommendation System’s exact contribution to a website’s revenue. On the other hand CTR’s are relatively simple to measure (compared to conversion rates for instance), while they are indicating the performance of Recommendation Systems quite accurately.

It is also important to note that – although the goal was to show as recent figures as it is possible – some of the following data are derived from slightly different time-periods in order to present a clear and relevant\textsuperscript{46} overview on Recommendation Systems’ capabilities and possible performance.

\textsuperscript{44} It is important to note that this is not a strict categorization but rather a guide that shows a possible grouping of the issues and some of the related solutions (provided by Recommendation Systems), organized along the maturity stages of Classified Sites. Furthermore, most of these groups of issues are valid throughout multiple maturity stages, only they are not equally relevant everywhere.

\textsuperscript{45} For example if a set of recommendations is requested 100 times on a given page type (the page was loaded 100 times) and it receives 7 clicks from 6 page views (one visitor clicked on 2 items at once by opening them in new browser windows), regardless of how many visitors have actually seen any recommendations, the Click-Through Rate on the recommendation placement will be 6 / 100 = 6% in the given time period.

\textsuperscript{46} Clear and relevant means avoiding underperforming set-up phases as well as major fluctuations and declines caused by significant changes or disadvantageous business decisions.
4.2 Attracting Visitors and Prolonging Browsing Sessions

As described earlier, the operators’ most important task at the first two stages of maturity (apart from deciding on the future sources of revenue) is to attract the largest amount of buyers and advertisers possible, and to provide them with a seamless user experience. For this, it is also vital to keep visitors on the site until they become familiar with the usage of the interface, discover some of their desired items, and send a number of Ad Replies.

This is not quite difficult to achieve if one basic fundament is taken into consideration: generally it does not really matter where the visitors are coming from (e.g. search, referrals, banners, etc.) or which page they are landing on until it is not a ‘dead end’. The site should highlight such further items to see that are personally relevant for the given user and that are related to the actual listing or item, or that are at least generally popular.

The following recommendation placements and scenarios are addressing this issue in more detail.

4.2.1 Main Pages

As Figure 5 (Created by the Author) shows, the Main Page of a typical Classified Site usually consist of the list (or map) of regions where items are available, and displays the main categories. The categories are usually represented by schematic figures or icons. Bouncing often depends on the site’s design, therefore on the users’ subjective evaluation of these illustrations. Moreover, the clarity of the different item classes also play a significant role in not leaving the site.
However, item groups can also be clarified and efficiently promoted by selecting the best performing items from each category, and utilizing their images (thumbs) as illustrations of the corresponding categories. The link behind a selected image can either lead to the represented category’s Result (Listing) Page or to the Ad Detail Page of the matching item.

The items and their images presented on the Main Page (and also the categories they are representing) are automatically selected by the recommendation solution after assessing their popularity and quality.47

Although the recommendations in the given example are not particularly emphasized, they can still achieve a CTR around 3%, as Figure 6 (Gravity R&D, 2014c) indicates. Clearly, this contributes to lowering bounce rate and engaging new visitors; however, it would be also possible to achieve higher results by highlighting the recommendations to a greater extent.

![CTR (%) on Recommendation Placements on the Main Page](Gravity R&D, 2014c)

### 4.2.2 Dynamic Landing Pages

Thinking the above further, it may be also worth enhancing not only the Main Page of a website, but improving every Landing or Entry Pages with customized offers to make sure that the users will continue browsing. Dynamic Landing Pages are frames, filled dynamically – at the very moment when the visitor enters the site – with the currently most relevant products, navigation and design elements, in order to convert visitors at the first impression with the highest likelihood (Gravity R&D, 2014e). For this, Recommendation Systems are taking into consideration both user related (if available) and contextual information in the associated scenarios. Among contextual data the referral URL is the most important, since this indicates the source of the traffic. Based on where the visitors are coming from, the following four main groups can be distinguished (Gravity R&D, 2014f).

---

47 The considered characteristics can be potentially the number of Ad Views and Ad Replies in a given time period, furthermore the resolution of the uploaded images, the completeness of the ad, the existence of certain keywords and the length of the description, among others.
From Organic Search Results

As of March 2014, according to Net Applications (2014) Google is strongly dominating the search engine market with its 69.55% global share while its biggest challenger, the Chinese local player Baidu has a thin share of 16.77%. Therefore this section is focusing only on Google’s search services.

Until recently, Google provided website operators with their visitors’ initial search keywords when visitors clicked on organic search results. Since the search terms are hidden (Advertising Age, 2013), it is hardly possible to customize Landing Pages for users coming from organic search results.

However, Google’s search engine is also often used as a ‘shortcut’. Many are relying more on this service than on the target websites’ navigation options, so they are including the target site’s name among the Google Search keywords when they want to look for something on a specific website. Eventually, users are typing ‘{product} {name of the Classified Site}’ into the search field instead of navigating to the URL of the Classified Site and then starting to look for the wanted product.

In most cases like this, users are landing on a Result Page that lists products from a specific category. If these Category Pages can provide visitors with recommendations based on popularity (if the visitor is unknown) or based on behavioral information (if the user is a returning visitor), they will less likely to leave the site unsatisfied.

From Paid Search Results

Google however still provides operators with the visitors’ initial search terms if they are clicking on paid advertisements displayed next to organic search results through Google AdWords. The keywords can be easily obtained from the referral URL.

When the visitor clicks on the advertisement, the Recommendation System can generate a unique landing page based on the search term. Naturally, operators could do the same thing by creating static pages for every keyword they advertise, but this becomes quite difficult when advertising for thousands of keywords in search marketing. Recommendation Systems are also able to automatically optimize these pages based on the implicit and explicit feedback of the visitors.

48 Many browser applications are also strengthening this phenomenon by providing users with the opportunity to start searches from the address bar.
From Referrals

Operators are usually creating unique landing pages for display advertising campaigns. Affiliate sites and other referrers however are generally leading to the Main Page or to Category Pages. It is also possible to customize these landings based on the given referrer site’s target group (based on its users’ general interests) and based on its content (which may be identified with the help of the URL).

From Personalized (Retargeted) Advertisements

Personalized Retargeting will be discussed in detail later in the ‘Retaining Users’ section. In general, these advertisements are directing visitors to Ad Detail (Item) Pages, which are also explained in depth as an independent topic.

4.2.3 Zero Result and 404 Pages

Unsurprisingly, as usually they have nothing to offer, Zero Result and 404 Pages are among the most common exit pages. Therefore it is highly suggested to display item recommendations on these page types as Figure 7 (Created by the Author) indicates. These suggestions are able to keep users on the site mainly by helping them in ‘Finding Some Good Items’.

![Search interface](image)

The recommendation scenarios can vary on these sites. By default, the personally most relevant items are recommended, taking the given user’s interests and the details of the current browsing session also into account. In case the empty page is generated as a result of a search query, it is also possible to eliminate some less relevant search filters (e.g. the neighbor regions may also provide appropriate items), furthermore Recommendation Systems can also attempt to correct typing errors or identify matching parts of the keywords using text mining techniques.
In case of neither a sufficient user history nor the keywords are available (e.g. a new user navigates to a 404 Page) usually the most popular items are recommended. However, if a visitor looks for a hidden or outdated item (the latter is likely to happen if an advertisement have been shared and the item has been sold since then), it is also advisable to include some items in the recommended set that are similar to the outdated (or hidden) ones.

As it is also visible on Figure 8 (Gravity R&D, 2014c) CTR on Zero Result Pages can reach even 40%, which clearly demonstrates the undeniable efficiency of recommendation solutions on these typical exit pages.

4.2.4 Ad Detail (Item) Pages

Recommendations on Ad Detail or Item Pages are among the most basic placements. Most probably these were the earliest implementation of recommendations, and today it is hard to find any major product listing website that does not operate a solution like this. Figure 9 (Created by the Author) shows a typical realization of suggestions on Ad Detail Pages.

‘You May Also Like’ type recommendations are able to attract visitors by giving users the impression that the site lists a great variety of their desired items and thereby that it can provide several alternatives to their needs, even if these are very specific.

More importantly, by displaying further appropriate items (that are mainly similar but slightly different from the current one) bounce rates can be significantly lowered on this page type, since there will be no need to start new searches, the users who are navigating to Ad Detail pages will continuously see some items that they like. Therefore user satisfaction and the average time spent on site are also likely to increase.
Regarding user related functions, primarily the ‘Just Browsing’ and the ‘Lack of Knowledge’ issues are addressed by recommendation scenarios displayed on these page types. For example, it is very likely to happen that a user finds an almost perfect item, but it lacks some of the desired features (e.g. perfect car, wrong color). In this case, users’ frustration can be considerably decreased if they are both ‘helped out’ by similar items and getting confirmed by the message “We may not have the perfect car right now but do not worry, consider these ones too!”
Although the recommendations on these pages are generally based on simple item-to-item scenarios, they are still highly efficient, providing 5% CTR from day one and being able to achieve above 20% after proper fine-tuning, as Figure 10 (Gravity R&D, 2014c) indicates. Apart from fine-tuning, the reason for the significant deviation of the above results is that they are also strongly depending on certain, site related business decisions such as the applied recommendation scenario and the position of the suggestions.

An additional demonstration of the versatility of recommendations on Ad Detail Pages is that they are also strongly belonging to the next group of issues (Easing Navigation and Decision Making), since they are able to change user paths fundamentally. These placements tend to cannibalize the number of page views on Result Pages, as they drive users into a loop of navigating from Ad Detail Page to Ad Detail Page without going back to the Result Pages. On one hand, not having to navigate back the result list after every Ad Detail Page visit makes user experience much smoother, but it is also important to note that this can be a drawback for sites that are earning most of their revenue from display advertising. The reason behind this is that less Result Page views mean fewer impressions, resulting in lower CTR’s on third-party ads, and smaller revenues from displaying these.

4.2.5 After Ad Reply Boxes and Pages

The original purpose of After Ad Reply Boxes and Pages is confirming that the message was successfully sent. The only difference between the two types is that in the former case the message is displayed in place of the Ad Reply Form or in a pop-up window (so the initial item remains visible) while in the latter case the user is redirected to a new (empty) page. In general – if there are no recommendations displayed – there seems to be no evidence that could prove which solution is better, it is usually the matter of the actual site.

It is not surprising that the operators are eager to supplement these messages with further suggestions since users are very likely to leave the site after having their ‘job’ (sending an Ad Reply) done, therefore – similarly to Zero Result and 404 Pages – After Ad Reply Pages are also typical exit points. Figure 11 (Created by the Author) shows an example for a pop-up After Ad Reply Box.

---

49 Some sites also implement ‘Those Who Viewed This Item Also Viewed’ type and other personalized scenarios, but in general, these tend to be slightly less efficient.

50 The variability in the individual data series belonging to a particular site is mostly the result of changes, while measurement errors can also occur occasionally.

51 Third-party ads are usually displayed next to result lists.

52 Therefore, the necessity of ‘sacrificing’ advertising spaces might be an important challenge when implementing Recommendation Systems.
The recommendation scenarios of these placements are usually not much different from the ones described at Ad Detail Pages. However, to maintain the diversity of the suggested items, the recommendation sets are usually generated as a number of randomly selected items from a pool of top items.\(^{53}\) If this is done similarly at Ad Detail Pages, users will probably not see any of the same items in the two sets.\(^{54}\)

Although it is not quite certain, there seems to be some correlation between the placement type and the CTR results. Figure 12 (Gravity R&D, 2014c) shows data from four different sites, where ‘Site A’ indicates a Classified Site that uses After Ad Reply Pages, ‘B’ and ‘C’ operate with pop-up boxes, and ‘Site D’ displays the After Ad Reply Box in place of the Reply Form. If it is presumable that the actual scenarios are set-up equally well, it seems that the CTR’s mainly depend on the size of the placements, or in other words on how much of the initial page is left visible. Therefore, probably the key indicator is the ‘emptiness’ of the new page, thus the number of the options users have, other than clicking on a recommendation.

\(^{53}\) This logic has a great importance at many other recommendation scenarios as well.

\(^{54}\) Naturally it is also possible to filter out those items from the recommended set (displayed in the After Ad Reply Box) that have already been suggested on the Ad Detail Page.
Consequently, when considering the performance of a scenario, CTR’s do not necessarily provide enough information after a certain depth. Therefore – from this point, without further website related statistics or an A/B test\(^5\) – it is hard to tell which one of these placements is the best to use. On one hand, if the users are redirected to an ‘empty’ page after replying to an ad, they will more likely leave the site, but those who will stay will be forced to click on such items that ideally they are interested in. Therefore they will remain in the previously mentioned loop and will continue sending Ad Replies. On the other hand, the exit rate will be probably lower if the content of the initial site (that is already proven to be interesting) remains visible, but in this case the suggestions are less emphasized therefore users will more likely to navigate back to a Result Page and start new browsing sessions instead of viewing further items. This might mean that the whole conversion path\(^6\) starts over, which will probably result in less Ad Replies.

Obviously there are many additional factors that can influence the above, therefore every placement options should be tested in their actual environment eventually ensuring the highest performance possible. But it is clear beyond doubt that recommendations in After Ad Reply Boxes and Pages can largely contribute to the success of Classified Sites.

\(^5\) A/B testing is a simple tool for evaluating the efficiency of two (or more) different solutions. It is usually done by creating a control group (A) and randomly selecting another group of users (B) who are provided with different scenarios or placements throughout a certain period of time. When this is over, the response rates of the two groups are compared in order to help the decision of the operators.

\(^6\) Conversion path is the path that leads from the Landing Page to the Ad Reply button, or in other words the process of turning a potential customer (prospect) into an actual lead.
4.3 Easing Navigation and Decision Making

The most important user related functions of Recommendation Systems are usually implemented at the ‘Sound Market Reach’ stage. In general, these functions are really difficult to apply simultaneously therefore operators have to choose different recommendation scenarios depending on their actual goals. These can be for example, decreasing the number of average clicks before conversion, helping discovery, creating a feeling of serendipity, or increasing the diversity of viewed items thus addressing the ‘long tail’ problem, among others.

However, nowadays it is indeed an increasingly valid expectation towards Recommendation Systems that they should be able to satisfy the needs of completely different user groups (such as ‘just browsing’, ‘determined’ and ‘shopaholic’ users) at the same time, by successfully identifying these segments and by being able to customize the recommendation scenarios according to their needs.

In conclusion, ideally all the users should see exactly those items that are somehow relevant for them during the actual browsing session, but none of those that are completely irrelevant. Therefore the Recommendation Systems implemented on Classified Sites should break away from serving the most popular offers, instead they have to be able to create highly accurate user profiles relatively quickly (just after a few events), knowing and utilizing not just the primary but the secondary, tertiary, etc. interests of the users, while identifying and taking into consideration the goal of the current browsing session and still maintaining a general sense of diversity.

The following recommendation placements and scenarios mainly aim to fulfill these requirements.

4.3.1 Result (Listing) Pages

Recommendation placements on Result Pages are universal in the sense that they are able to provide a solution for most of the user related functions of Recommendation Systems.

A possible implementation of ‘Annotation in Context’ is highlighting the most important search filters and the list of the related categories, while preserving the clear layout or the alphabetical order. It is also feasible to provide users with the opportunity to reorder the result lists based on personal relevance\(^{57}\) to help them in ‘Finding a List of Good Items’, while recommendation placements next to these lists can also serve multiple goals. These

\(^{57}\) By default, most Classified Sites are ordering result lists by the date of upload.
suggestions are most usually displayed above, below or next the results, but many Classified Sites are also displaying recommendations among the list of items, as Figure 13 (Created by the Author) indicates.

Based on the applied recommendation scenario, these can increase the diversity of the items, guide uncertain and ‘just browsing’ users, or highlight the best offers to make user decisions easier and more satisfying.

Not surprisingly, the performance of the recommendations on Result Pages also depends on the applied scenarios, on the placements and on the actual business requirements. In general, as Figure 14 (Gravity R&D, 2014c) shows, the CTR’s can range from 5% to 12%.
4.3.2 Keyword Based Search Result Pages

Some sites are applying unique scenarios for those listing pages that are generated as a result of a keyword based search query. In this case, the recommended set of items is typically placed to a more recognizable part of the page, such as above the result list, as Figure 15 (Created by the Author) indicates.

![Recommendation Placements on Keyword Based Search Result Pages](created_by_author)

These scenarios are related primarily to the ‘Lack of Knowledge’ issue, as they help users in finding the most relevant items even if their search terms were not too specific or misspelled. A typical example is a user who is interested in car parts (this is known from his history) but he only types a car manufacturer’s name (maybe even misspelled) in the search field. Although actual cars are generally much more common and popular on Classified Sites than car parts, in this case the Recommendation System will suggest the latter on the top of the page, while by default, the result list will remain displaying every item that matches the keyword (mostly actual cars) sorted by date.

Figure 16 (Gravity R&D, 2014c) shows the CTR results of a scenario similar to the above. What is not visible though, that the slight but continuous growth that the chart outlines was an effect of more than 40 changes in the scenario during the given period. Although 8% is not a very high result, but undoubtedly, this solution helped thousands of users in finding their desired products (the data below is derived from more than 12.5 million clicks) and the further growth only seems to be the question of continuous fine-tuning.
4.3.3 Device Awareness

Although different market players are publishing various results – and these may highly depend on the actual site and on the mobile device penetration of the given region – undoubtedly, the amount of users visiting Classified Sites from mobile devices is becoming more and more significant. For example, La GESTE claims that 30% of French mobile users were using phones to visit Classified Sites in Q2 2012 (Shipside, 2012), moreover according to GeekWire, Zillow.com’s mobile visits reached 60-70% in 2013 with a yearly growth of 150% (Soper, 2013), and finally a study conducted by Placed and Cars.com reveals that 81% of car buyers used smartphones to do research when purchasing a vehicle (Placed Inc., 2014).

The main point is though, that modern Recommendation Systems have to be prepared to the challenges related to visitors using multiple devices.

One of these challenges is that users’ browsing habits turn out to be slightly different on mobile devices. Recommendation Systems have to be aware of these differences and the related scenarios should be formed accordingly. In general, the main reasons behind the distinct behavior are the smaller screen size, the different navigation, and the on-the-go usage.\footnote{On-the-go usage does not necessarily mean outside home, but short-term, temporary usage.} Therefore suggestions should be more direct and they need to be displayed differently (e.g. in scrollable boxes instead of lists), moreover Recommendation Systems have to provide additional functions (such as predictive search and instant results) in order to help users in finding their desired products with even fewer interactions. Consequently, the similar item recommendations on Ad Detail Pages can also represent an increased value on mobile screens, since these can always provide users with a next step without having them return to a Result Page.
The other big challenge is linking user profiles from multiple devices together. It can be an absolutely valid user need to provide the ability to continue the browsing on a smartphone at the same point where it has ended on a PC. On the other hand, knowing every detail about a given user’s interests will cease to be a valuable asset if he will prefer visiting the Classified Site from another device, since the different devices will not share the same cookies. But instead of making Recommendation Systems create duplicated profiles, operators should encourage users to give their email addresses at least once on every device they use. For example, it is usual that even when a user is not logged in to the site, leaving a reply email address is prerequisite for sending an Ad Reply, moreover most of the native mobile applications also request a one-time authentication. If the email address is obtained, it becomes possible to merge the corresponding user profiles that were identified with different cookies on various devices.\(^5\) This method also enables the accurate analysis of the device specific differences in the behavior of certain user segments.

Furthermore, in the near future it may also become possible to ultimately identify the users across devices without knowing their email addresses. An advertising technology company called Drawbridge claims that they can connect mobile and desktop visits throughout 15-20 observations over 2-3 weeks with 60-70% accuracy (Geron, 2012).

Other than challenges, mobile devices also provide operators with further opportunities. Native mobile applications can for example capitalize on event driven push notifications (e.g. a recently viewed item’s price has just decreased), on location and other context based services (e.g. a favored item might be available for viewing nearby, upon calling the seller) and on the devices’ camera (e.g. the app can help shooting a well converting photo by showing sample images of well performing ads from the given category), among others.

\(^5\) Another way of identifying users on different devices is making them click (or tap) on a link in a personalized email at least once on every device they use. In this case, the URL’s in the personalized email can be supplemented with a unique user ID that can be read by the Recommendation System’s tracking code after the user lands on the website, and then this ID can be merged with the device’s cookie.
4.4 Monetizing Users

As mentioned earlier, although it is also possible to monetize non-seller users (mostly indirectly by third-party ads), Classified Sites’ main revenues come from a number of services provided specifically for advertisers. Even though this group is relatively small compared to the total user base of a mature site, sellers are the ones who are actually earning money as a result of advertising and trading their products, therefore they are the ones who can be charged for using the site in the first place. Furthermore, as Classified Sites do not facilitate actual transactions, there is no option for taking extra commissions thus only the advertising opportunity (and certain extra services) can be sold by the operators.

Apart from tailoring each and every item to the appropriate audience thus making it possible to sell practically anything, Recommendation Systems’ main role in monetizing users, as the following examples indicate, is to turn buyers into advertisers and to convince them that it is worth paying additional fees for premium services.

4.4.1 Optimized Gallery

Many Classified Sites provide sellers with a service called Gallery that operates as an additional display advertising opportunity. In general, this function works really simply: sellers pay a certain amount per ad, and operators guarantee a minimum number of impressions.

By default, operators can chose between two main strategies for managing these advertisements. Either they can display everything for everyone (rather randomly), or they can focus on the most popular ads. In the former case, probably fewer users will click, moreover they will become unsatisfied with the irrelevant offers and sellers will not be satisfied either. In the latter case, although users will feel more contented and they will probably click more, it will become difficult to guarantee the minimum number of impressions for the unpopular items.

As described earlier, Recommendation Systems are able fulfill the operator related ‘Sell more diverse items’ function by targeting the right advertisements to the right audience. Therefore, by personalizing the Gallery, operators can guarantee the minimum number of impressions (while the more relevant ads will also result in higher CTR’s and more Ad Replies) and at the same time the user experience will also turn to be positive as an effect of the useful offers.

---

60 Beyond the users’ individual interests and history, similarly to the other scenarios, the targeting can also take into consideration the current search terms, the applied filters and contextual information, among others.
Additionally, there is also a notable potential in automatically assessing the advertisements (just as described at Main Pages) and rewarding high-quality ads by charging lower Gallery fees, thus operating a system similar to Google AdWords.\(^{61}\)

### 4.4.2 Campaign Management and Reporting Tools for Operators

Some Recommendation System Providers are also developing tools that allow operators to individually adjust scenarios in order to meet changes in business strategies and to reach specific monetization goals. The main requirements towards these Campaign Management tools are to provide a clear, user friendly and easy-to-use interface, while (by combining adaptively learning algorithms and manual rules) making it possible to create, test, maintain and follow up sophisticated scenarios and campaigns that are able to manage various user groups throughout a number of interaction points, furthermore that can also comply with different business requirements simultaneously.

As described at the ‘Understanding user needs’ operator related function, the huge amount of data collected, processed and analyzed by Recommendation Systems can be also further utilized in various ways (such as in email marketing, in social, search and display advertising and in offline promotions, among others). This would not be possible without Reporting tools that support multidimensional queries in order to help operators learn more about their customers’ behavior, interests, demographics and other insights, and without further Business Intelligence solutions such as market trend analytics and forecasting.

Ultimately, utilizing these tools can result in more qualified leads, in improved engagement and higher conversion rates, furthermore in greater user lifetime value (Gravity R&D, 2014e).

### 4.4.3 Insights for B2C Advertisers

Some of the above mentioned insights can also be provided for B2C advertisers. Real estate agents for example, might be particularly interested in the performance of their advertisements (such as the daily number of impressions, Ad Views and Replies) and they might be also concerned about the actual return on their investments in premium services. These data can be displayed in the form of simple reports, supplemented with further information, such as the prediction of the advertisements’ future performance and the potential effects of subscribing on additional premium services. Furthermore it is also possible to provide advertisers with a collection of the most competing ads.

---

\(^{61}\) Google AdWords measures a so called Quality Score when determining the position, where ads are shown on a result page. If their advertisements are excellent and also highly relevant, advertisers can win a higher position at a lower price than their competition (Google Inc., 2014b).
In conclusion, it is probably much easier to make B2C sellers interested in improving their advertisements (thereby improving the overall ad quality of the site), if they know more about their competition, and on the other hand, they will be also much more motivated to spend on premium services, if they are provided with a proof of the past and a forecast of their ads’ future performance through these insights (Gravity R&D, 2014e).

4.5 Retaining Users

Letting previously acquired users go\textsuperscript{62} tends to be a really expensive practice in case of Classified Sites. Loosing users means not only weakening the given site’s market position and probably strengthening the competitors’ (as every user represents a certain value), but it is also particularly dangerous due to the fact that well-connected users are possibly dragging some of their relations too. Additionally, the transition barriers between Classified Sites are not really substantial\textsuperscript{63} due to the lack of long term contracts, and because of the number of alternative services and the similarities in their usage.\textsuperscript{64}

Consequently, it is vital to perform clear actions in order to avoid churn. This is especially true for mature sites, whose possibilities to attract completely new users are usually very limited.

Apart from determining users’ value and providing special offers with the help of CRM\textsuperscript{65} and other Business Intelligence solutions, the following placements and scenarios are the most effective tools for retaining users.

4.5.1 Email Personalization

Email Marketing is a standard and cost-effective tool in the E-commerce Industry for acquiring and retaining customers, and most of the Classified Sites are also extensively leveraging this direct marketing technique. Similarly to websites, it is also possible to extend emails with dynamic content. Recommendation solutions can significantly increase the efficiency of Email Marketing, since this way (from the undesired by-products of having a user account registered), emails can become personally relevant messages that are providing customers always with timely and useful content.

\textsuperscript{62} Also known as Churn  
\textsuperscript{63} Compared to Telecommunication Service Providers for example  
\textsuperscript{64} Most Classified Sites are practically built on the same logic therefore it is not too difficult for users to get familiar with the usage of a new interface.  
\textsuperscript{65} Customer Relationship Management
Content
Based on the classified businesses and the capabilities of Recommendation Systems, this section will focus on the following four types of emails.

Newsletters
Newsletters are electronic publications that are sent to the subscribed customers on a regular basis. Edited by the operator, they are serving general marketing purposes, such as customer engagement.

In accordance with their content, newsletters can be enhanced with personally relevant item recommendations, as well as with a report from top trending product categories and a number of best case practices regarding how to sell them. The former scenario can be tailored to regular buyers based on the activities that they performed since the last newsletter, while the latter can target advertisers to give them inspiration what to sell. Ideally, a well conducted newsletter can create – or at least increase – both the supply and the demand for certain (seasonal) product categories.

Transactional Emails
Transactional Emails are automatically generated, necessary, formal elements of certain user-website interactions, such as registration, account changes, Ad Reply confirmations and purchases, among others. The recommendation scenarios in these messages can provide buyers either with a list of personally relevant or further similar items to consider, or they can contain call to actions in order to turn buyers into sellers. Furthermore they can also lead to user satisfaction surveys that can eventually supplement the existing user profiles. Additionally, these emails can provide a good opportunity to cross-sell and upsell premium services for sellers (e.g. an advertiser can receive a notification about an Ad Reply, pointing out the potential positive effects of spending more on premium advertising options), and they can also contain incentive offers.

Triggered Email Alerts
Triggered Email Alerts are automatic notifications that are activated in response to specific events. These can inform potential buyers about new, presumably favorable items or price changes, while advertisers can be also notified if their advertisements are underperforming, additionally providing them with useful hints and premium offers in order to improve.

It can be also useful to trigger additional messages after an Ad Reply confirmation, if the advertiser of the given item does not respond in 24 hours.
**Remarketing Emails**

Remarketing Emails aim to drive back inactive users to the Classified Site. Again, one of the best ways to address buyers is suggesting a set of the most relevant items to consider, while potential advertisers can be efficiently targeted with a report from the most trending, easiest to sell product categories, highlighting those item groups that are matching well with their interests.

**Timing**

Apart from content, timing is also a key aspect in increasing the efficiency of promotional emails. As personalization is able to improve response rates, good timing can guarantee higher open rates. A research conducted by GetResponse (Andrzejewska, 2012) points out that 23.63% of all email opens occur within the first hour after delivery, while the average open rate after two hours is 9.52%, which decreases steadily to 6.33% and 4.8%. After one day, this eventually drops to 0.63%.

Not surprisingly, some recommendation solutions are also able to predict the time when a given user is most likely to read his emails, based on previous opens. Additionally, the weather, weekends, national holidays and general mood can be also taken into consideration when deciding on when to send the messages.

Nevertheless, even if the above techniques are applied well, there will be still a huge number of emails that will not be read in the first few hours or days after sending. In these cases, the question may arise how Recommendation Systems are capable of providing timely suggestions in emails that have been already sent, when the content of Classified Sites are changing so rapidly.

Clearly, there is not much use of recommending items that are already sold or deleted. Moreover, it is also possible that while the email stays unopened in his inbox, the given user starts to be interested in new item categories, therefore the initial suggestions may become irrelevant over the time. Modern recommendation solutions are managing this issue by generating the dynamic contents at the very same moment when the given message is opened. For this, operators only have to insert a HTML code into their email design that identifies the user as well as the proper scenario, and then this code will call the recommendation provider’s API to serve the user with suggestions in real-time (Gravity R&D, 2014b). This way, the recommendations will always be based on the currently available item catalogue and on the most recent user history.
Figure 17 (Gravity R&D, 2014c) shows CTR results measured in the personalized newsletters of a Classified Site that sends emails often on a daily basis. The values are calculated as the number of unique clicks divided by the number of newsletters sent. The CTR’s around 3.5%-4.5% are slightly better than non-personalized but well designed, timely newsletters (this site does not use personalized timing), therefore these results can be considered average compared to the full capabilities of dynamic emails.

As the above examples also indicate, if applied well, Email Personalization solutions are capable of much more than ‘just’ retaining users efficiently, they can scale up revenues by contributing to the monetization of users to a great extent.

4.5.2 Personalized Retargeting through Display Advertising

Display Advertising was among the first internet advertising techniques, and the image, flash and rich media banner campaigns are still widespread across the Web. Although during the early days of the Internet banners have been also utilized to drive traffic, nowadays they are serving mainly long term objectives, such as increasing brand awareness and contributing to offline goals. Therefore, display advertising spaces can be bought relatively inexpensively (at least in case of impression based pricing), which is not surprising if we take into consideration that according to Google (DoubleClick), in 2009 the average CTR’s on display advertisements were 0.1% in the United States (2010a), and 0.07-0.14% in the Europe, the Middle East and Africa Region (2010b).

In comparison, Personalized Retargeting is able to efficiently utilize the cost-effective advertising spaces and it has a large potential to revolutionize Display Advertising, since this
Technique is proven to contribute to sales and conversion goals to a great extent – therefore it can radically improve ROI – while still remaining capable of increasing brand awareness. The huge potential and the indisputable advantages of these solutions are based on two main elements. One is the failure of Display Advertising in its current form, and the other is the outstanding technological advance behind this technique.

By default, retargeting is a form of advertising that focuses on users that have already visited the target website. Retargeting solutions are addressing these users by targeting them with banners on various domains across the Internet. This means that normally only these users are provided with the given website’s advertisements, or at least they are provided more frequently compared to users that have not been on the target website. Users are identified and followed across the Web by third-party cookies that can notify the retargeting provider when to serve an advertisement, which is then usually displayed through an ad network. Sometimes the retargeting provider and the ad network is the same organization, as it happens in the case of Google Ads (2014a).

Although retargeting is a great way to re-engage and convert high-intent prospects (Gravity R&D, 2014e), it does not necessarily provide an answer to what to promote. This issue is particularly valid if the advertiser’s site lists thousands of products, since even though all of the targeted users have visited the site at least once, obviously not all of them will be interested in the same items. Google (2014a) addresses this issue by making it possible for operators to tag their product pages with different labels (such as ‘TV’) and allowing them to create specific banners for each of these categories. Those users who have been browsing the ‘TV’ category will then be served with the related advertisement.

Compared to this basic form of retargeting, Personalized Retargeting solutions served by Recommendation Systems can ensure the following additional advantages.

1. Thanks to the continuous item catalog synchronization, there is no need for tagging product pages. From the retargeting provider’s point of view, every item is unique, characterized by detailed metadata.

2. Users are not only targeted based on their last page views, but based on their elaborate user profiles and interests. The information collected from on-site activities is also linked to the cookies that identify the users. Furthermore, when generating the ad, users’ current context (e.g. location, device) is also taken into consideration.

Many studies can confirm this statement with surprising statistics, such as only 8% of all internet users account for 85% of the clicks on display ads (comScore Inc., 2009).
3. The displayed banners are dynamic. This means, on the one hand, that there is no need to create unique banners for every item category. On the other hand, dynamic banners can attract users much more efficiently by displaying actual items (instead of general illustrations) based on personal relevance. Rich media advertisements are also allowing interactions inside the banner (e.g. the users can request a new set of recommended items by a refresh button). Moreover, while regular banners tend to burn-out after the first few weeks of a campaign as users loose interest, there is no tendency of declining performance in case of Personalized Retargeting. On the contrary – as the examples will also indicate –, CTR’s are usually growing steadily during the campaign, as the system ‘learns’. Additionally, while most of the regular retargeting banners are focusing mainly on the ad performance by targeting only those users that are already aware of the brand, the dynamic technology is also providing the opportunity to display optimized banners for unknown users. These are able to create brand awareness by an additional starting slide for example, and then they can drive the visitor to the target website by displaying popular items, taking contextual information also into account.

4. Personalized Retargeting banners can also track implicit and explicit feedback, and supplement user profiles with these, off-site collected information.

Consequently, Personalized Retargeting banners are not much different from on-site recommendation placements. Both techniques operate with content generated by various algorithms based on unique user profiles, and both of them are served in real time upon a request. However, the process of displaying Personalized Retargeting banners is a bit more complicated, especially if it is done through ad networks. This process can be summed up by the following steps.

1. A user visits the target site, but does not convert and leaves. This site has both the ad network’s and the Recommendation System’s code implemented, therefore two cookies will be placed in the user’s browser: one by the ad network and one by the Recommendation System.

2. The user visits another website that is a part of the ad network (it also has the ad network’s code implemented), and that is also among those domains that the operator of the target site wants to advertise on.

3. The ad network’s ad server is responsible for specifying which advertiser’s banner should be displayed at a given ad space. This is optimized based on certain parameters in order to
maximize the ad network’s revenues. The ad network identifies the user by reading his browser’s corresponding cookie, and in this case, it finds out that he has already been on the target website.

4. Then, if the ad network’s algorithm selects the target site’s advertisement, the recommendation provider’s iframe\textsuperscript{68} will be loaded into the given ad space.

5. At this point, the recommendation provider’s code is called. First, it identifies the user by the matching cookie, and analyzes the current context based on the information that is provided by the user’s browser. As the user is identified, the Recommendation System will generate him a personalized banner, which will contain suggestions based on his user profile. In those cases, when the user does not have a cookie placed by the Recommendation System,\textsuperscript{69} he will be served with a non-personalized advertisement that will show him popular items based on the available contextual information, and also with a cookie that will register his interactions with the banner.

Personalization can be integrated both to HTML and Flash banners. This can be done by implementing three code parts: one for requesting the recommendations, one for displaying the results of the recommendation calls and one for tracking user interactions (Gravity R&D, 2014b).

Including ad networks, personalized advertisements are usually displayed through the following platforms.

**Through Affiliate Sites**

Affiliate or referrer sites are webpages that are displaying certain advertisements of the target site based on a unique agreement with its operator. Sometimes – as the earlier mentioned Ingatlan.com’s example also indicates –, operators themselves are establishing these websites in order to attract a massive number of users with certain contents (e.g. news, weather), and thereby to drive traffic to their E-commerce site.

Figure 18 (Gravity R&D, 2014c) shows the CTR results of a Classified Site that capitalizes on the Personalized Retargeting technology through affiliates. The data below is an average, calculated from the results of various banners placed on multiple sites. It is visible that non-personalized, but dynamic, contextual information and popularity based product

\textsuperscript{68}An inline frame is an HTML element that is able to host an individual HTML document.

\textsuperscript{69}This can happen if unknown users are also targeted by the ad network (at the operator’s request), or in those rare cases when the Recommendation System does not provide some users with a cookie due to an error.
recommendations (these are served for unknown users) can also perform almost five times better than the earlier cited industry average, while the CTR’s on personalized advertisements can exceed even 2% on average.

**Figure 18:** CTR (%) on Advertisements Retargeted through Affiliate Sites (Gravity R&D, 2014c)

### Across multiple E-commerce Sites

If an operator have a portfolio of Classified Sites (e.g. separate sites specialized on cars, real estate and jobs) or any other E-commerce businesses, it is also possible to track user behavior across multiple domains, while building an integrated user profile. This means that even those users, who have never visited the real estate classified for example, can be accurately targeted with apartments while browsing cars.

In this case, Personalized Retargeting is actually equivalent with cross-site recommendation placements.

### Through Ad Networks

Ad networks, such as Google Display Network, can display banners literally anywhere on the web. Therefore, they can contribute to a huge reach while providing a relatively inexpensive opportunity to target certain user groups based on specific domains and ad placements (e.g. it might be possible to buy only side banners in the business columns of certain news sites).

**Figure 19** (Gravity R&D, 2014c) outlines the results of a Personalized Retargeting campaign run through an ad network. Similarly to the previous chart, this also indicates CTR’s both on non-personalized and personalized advertisements. As it is visible, the former are mostly above 0.5%, while the latter perform outstandingly between 2.5% and 7%.
As all the above examples and data indicate, Personalized Retargeting has a great potential in contributing to the success of Classified Sites, not just by efficiently retaining previously acquired visitors, but also by successfully attracting new target groups.
5. Further Challenges

As Recommendation Systems are evolving and their utilization becomes more and more diversified, there are a number of emerging topics that might be important to consider, especially if we are not indifferent regarding the future opportunities of these solutions. Among these, there are both general issues and matters specifically related to Classified Sites. Based on the earlier findings of this thesis and on the conclusions of Ricci et al. (2011), this section aims to provide a brief outlook on the emerging topics by shortly introducing them, rather as a selection of the most important challenges and as a supplement to the upcoming conclusions, than as a complete overview.

Scalability

If the current trend further evolves, the development of recommendation algorithms has to keep pace with the continuously increasing amount of data generated by various platforms. Sticking to the online classified example, it is not just the growing number of users and items that can make serving suggestions in real time more and more difficult, but both the number of user-platform interactions (such as ratings, preferences, reviews, etc.) and the mass of the available external and contextual data sources (such as social networks, locations, devices, news, weather, etc.) are also growing rapidly.

On the other hand, Recommendation Systems will still have to be able to provide accurate suggestions from relatively small data sets (e.g. for the users of an immature site) while continuously adapting to the given platform’s development.

Privacy and Control

Most users are probably not aware of how recommendation solutions work, therefore it is not likely that they are conscious of the extensive data collection on the related platforms. Thus the main question is: to what extent should users be warned? Furthermore, should they be able to get to know the profile built about them and modify or delete it?

Although there were a number of attempts to introduce comprehensive a regulation, the answer is not so obvious. Apart from the fact that most sites notify their visitors about the use of cookies, operators usually grant different levels of user involvement, as the following examples indicate. Some providers like Amazon (2014b) give detailed explanations for the recommendations and also let users to edit their browsing history. Apart from managing

70 For example the ‘EU cookie law’ (e-Privacy Directive) addresses the issue (ICO, 2014).
71 Such as ‘Customers Who Bought Items in Your Recent History Also Bought’
search history and providing the option to turn tracking off (Google Inc., 2014c), Google also allows its users to see and modify what the company thinks about them, and users are also entitled to opt out services that are operating based on this information (Thier, 2012). Facebook, in comparison, makes it possible eliminate undesired posts from the ‘News Feed’ and control content recommendations by taking feedback surveys (Wagner, 2014).

Naturally, operators should prevent users from becoming frustrated from the feeling that the system knows too much about their true preferences. On the other hand, Recommendation Systems should ensure that the knowledge about users cannot be freely accessed by malicious applications (Ricci, et al., 2011). However, as mentioned earlier, most of the recommendation solutions manage users anonymously without misusing any of their sensitive data (such as name, exact address or links to social profiles), identifying them with a single string.

Another aspect of this issue is related to the operators’ ability to easily and flexibly modify, test and apply any elements of the recommendation algorithms, scenarios and placements. This might be possible with highly modular and customizable systems and with advanced Campaign Management Tools.

**Proactivity**

In most cases Recommendation Systems and their applications are following a ‘pull’ model, only serving suggestions when explicitly requested (e.g. when navigating to a page). However, there is an emerging need of so called ‘push’ recommendations that are served based on implicit requests, predicting not only what to suggest, but also when and how (Ricci, et al., 2011). ‘Google now’ is a good example for the latter approach, since the service is able to provide its users with the right information and with the relevant suggestions at just the right time, operating as a personal assistant (Google Inc., 2014d).

**Diversity**

Users will probably find their desired items easier if there is a certain degree of diversity among the items included in a recommended set, furthermore “there are many situations, especially in the early stage of a recommendation process, in which the users want to explore new and diverse directions” (Ricci, et al., 2011, p. 26). In comparison, Recommendation Systems are often accused with narrowing the number of choices too severely when trying to find the perfect items. On the other hand, irrelevant suggestions often cause dissatisfaction, therefore perhaps the solution lies in somehow assessing the given user’s actual need for diversity and serving recommendations accordingly.
Advertising
As mentioned earlier, it might be necessary to discard third-party advertising spaces on websites in order to create room for recommendations. Additionally, suggestions on Ad Detail (or any product) Pages tend to cannibalize Result Page views. Both can lead to fewer impressions, lower CTR’s and eventually smaller revenues from third-party ads.

However, on the positive side, Recommendation Systems might also be able to increase the efficiency of the remaining advertisements by providing the ability to personalize them. Amazon, for example, operates with a system of displaying ‘Interest-Based’ third-party ads that are targeted upon information collected from on-site activities (Amazon.com, Inc., 2014c).

Supply and Demand
Classified (and Auction) Site operators’ interest is naturally to ensure that every buyer has the opportunity to find his perfect product, and also that every advertiser can sell practically anything. If we assume that Classified Sites are not perfect markets, Recommendation Systems can have a significant role in automatically controlling the internal supply and demand. On the one hand they can efficiently tailor niche products (that the sellers may have an excess supply of) to their appropriate audiences, and on the other hand they can highlight the most demanded item categories (that the site may have a shortage of) to activate potential advertisers.

Short-Term and Long-Term User Preferences
Many of today’s algorithms are focusing either on the short-term or on the long-term user preferences when generating recommendations. Obviously, even if we suggest that the former are departing from the latter, there might be a large difference between the related needs (Ricci, et al., 2011). A solution that can implement both of these approaches is possibly taking external data sources and contextual information extensively into consideration, in order to decide how much it is needed to rely on the previously collected user information compared to the given user’s current behavior.
6. Conclusions

This thesis discussed the significance and the main characteristics of online classifieds, moreover it introduced the basics of Recommendation Systems and considered their current and potential role in today’s and tomorrow’s data-rich industries. Finally, it presented and evaluated a great variety of recommendation solutions implemented on Classified Sites covering various user-platform interaction points, and provided a brief outlook on further, emerging topics.

C2C E-commerce, and the appearance of online classifieds (that have almost completely taken over their paper-based predecessors’ place in the past two decades), both streamlined and globalized the traditional person-to-person trading, making it possible for almost everyone to exchange goods and services economically, regardless of time and space.

Although these platforms possess a number of unique characteristics that make difficult for their operators to turn them into profitable businesses, this thesis introduced several possible options for revenue generation, and also a number of applicable business models. These were then combined in a model that described classified businesses along certain maturity stages, indicating the most important issues and opportunities during each phases. After the overview on the correlations between the maturity and monetization, potential revenues were also discussed on the level of the users, studying their demographic characteristics and presenting their typical life cycle.

Consequently, although Classified Sites may never become the number one trading platform, they are already important players of both the Advertising and the E-commerce industries, and their significance will probably further increase as more and more users are becoming aware of the simple, direct, free and independent buying and selling methods that they provide.

Moreover, considering today’s Internet trends, and the fact that online classifieds already provide a highly decentralized form of commerce, it would be not surprising if they would further develop in the direction of community based models.
On the other hand, we can think about the appearance of Recommendation Systems as the data scientists’ sophisticated answer for today’s overwhelming content glut. Apart from Big Data, the development of these solutions was strongly motivated by the observation that individuals in unclear situations often rely on recommendations provided by others, regardless of the recommender and of the given decision’s actual weight.

As this thesis indicated, Recommendation Systems are already commonly applied throughout various domains, and they still have extensive opportunities in many additional areas of life. The real strength of these systems lies in the fact that, unlike previous solutions, they are able to provide fully personalized suggestions by building individual user profiles.

Throughout their outstanding abilities, Recommendation Systems can effectively address a number of issues that are particularly important for both the users and the operators of E-commerce platforms. However, it would not be possible to fulfil these functions without the extensive collection of certain data (such as user actions, item details, contextual information and social events), and without the application of complex algorithms that are capable of both processing all inputs and serving masses of end-users with accurate suggestions in real time. Additionally, the thesis also discussed the architecture and the integration of the presented solutions briefly.

After the general overview of the above topics, the earlier introduced maturity model was completed with some of the most important online classified related issues that can be addressed by implementing recommendation solutions. Thereby the model could cover many areas, such as bounce rates, the average time spent on site, navigation, decision making, monetization and retention.

Then, the thesis presented and assessed a number of placements and scenarios for each of the groups of issues, and illustrated these solutions with the necessary data, charts and figures. Therefore there have been many successful implementations identified, such as different on-site recommendation use cases on each classified page types, various tools for monetizing users through suggestions, and a number of opportunities that lay in personalizing emails and advertisement platforms, among others.

The last section shortly discussed the most important emerging topics related to the above solutions, such as scalability, privacy, control, diversity, proactivity, and managing supply and demand, among others.
In conclusion, recommendation solutions and other intelligent systems are very useful applications that can guide people through many areas of life, since they are not only able to help us in finding our desired products, but with their assistance, any of our complex decisions can become much easier to make. Although the growing number of positive use cases are often offset by different drawbacks and challenges, there might be still an enormous yet undiscovered potential in these systems, therefore many innovative companies are working on their further development in order to make them serve their users better, by providing an affordable option to proactively and efficiently convert today’s and tomorrow’s increasingly extensive and truly incomprehensible data overflow into practical knowledge.

Consequently, although in the most cases the final goal of implementing Recommendation Systems is to scale up revenues, this would not be possible without actually achieving a higher goal: user satisfaction.
References

Available at: http://aimgroup.com/purchase/classified-intelligence-report/
[Accessed 14 April 2014].

Available at: http://aimgroup.com/wp-content/uploads/2012/05/cir-1301.pdf
[Accessed 14 April 2014].

Available at: http://adage.com/article/special-report-the-advertising-century/ad-age-advertising-century-timeline/143661/
[Accessed 2 March 2014].

Available at: http://adage.com/article/dataworks/google-hides-search-terms-publishers-marketers/244949/
[Accessed 5 April 2014].

Available at: http://allegrogroup.hu/
[Accessed 6 March 2014].

Available at: http://allegrogroup.com/about-us
[Accessed 10 March 2014].

Available at: http://www.amazon.com/gp/help/customer/display.html?ie=UTF8&nodeId=16465251
[Accessed 14 March 2014].

Available at: http://www.amazon.com/gp/help/customer/display.html/ref=pd_ybh_help?ie=UTF8&nodeId=201145440
[Accessed 19 April 2014].

Available at: http://www.amazon.com/b/?&node=5160028011
[Accessed 19 April 2014].

Available at: http://www.wired.com/wired/archive/12.10/tail.html
[Accessed 15 March 2014].

[Accessed 8 April 2014].

Available at: [http://www.cs.brandeis.edu/~magnus/ief248a/eBay/history.html](http://www.cs.brandeis.edu/~magnus/ief248a/eBay/history.html)  
[Accessed 13 April 2014].

comScore Inc., 2009. *comScore and Starcom USA Release Updated “Natural Born Clickers” Study Showing 50 Percent Drop in Number of U.S. Internet Users Who Click on Display Ads - comScore, Inc.* [Online]  
[Accessed 9 April 2014].

[Accessed 15 April 2014].

Available at: [http://support.deezer.com/customer/portal/articles/1349609-how-does-deezer-make-recommendations-](http://support.deezer.com/customer/portal/articles/1349609-how-does-deezer-make-recommendations-)  
[Accessed 14 March 2014].

D'Orazio, D., 2013. *Foursquare will soon push personal recommendations to your phone | The Verge*. [Online]  
[Accessed 14 March 2014].


[Accessed 2 March 2014].

eNET-CUB, 2013. *eNET » Blog Archive » eNET-CUB – The online small ads market is not even so small*. [Online]  
[Accessed 10 March 2014].

eNET-CUB, 2014. *eNET » Blog Archive » There can be only one – battle on the online classified ad market*. [Online]  
Available at: [http://www.enet.hu/news/there-can-be-only-one-battle-on-the-online-classified-ad-market/?lang=en](http://www.enet.hu/news/there-can-be-only-one-battle-on-the-online-classified-ad-market/?lang=en)  
[Accessed 24 April 2014].

Enright, A., 2013. *Fulfillment/Delivery - Free shipping, more than fast shipping, draws consumers back - Internet Retailer*. [Online]
Available at: http://www.internetretailer.com/2013/10/09/free-shipping-more-fast-shipping-draws-consumers-back
[Accessed 3 March 2014].


[Accessed 5 April 2014].

[Accessed 8 April 2014].

[Accessed 8 April 2014].

[Accessed 14 March 2014].

[Accessed 5 April 2014].

[Accessed 19 April 2014].

[Accessed 19 April 2014].

Gravity R&D, 2013. RECO Architecture (Internal material, with the permission of the owner). Budapest: Gravity Research & Development Ltd.

[Accessed 14 March 2014].

Gravity R&D, 2014b. RECO Integration (Internal material, with the permission of the owner). Budapest: Gravity Research & Development Ltd.
Gravity R&D, 2014c. *RECO Statistics (Internal data, with the permission of the owner).* Budapest: Gravity Research & Development Ltd.


Gravity R&D, 2014e. *RECO Product Descriptions (Internal material, with the permission of the owner).* Budapest: Gravity Research & Development Ltd.


Available at: http://marketshare.hitslink.com/search-engine-market-share.aspx?qprid=4&gpcustomd=0&gptimeframe=M
[Accessed 5 April 2014].

Newspaper Associations of America, 2013. *Annual Newspaper Ad Revenue.* [Online]
Available at: http://www.naa.org/Trends-and-Numbers/Newspaper-Revenue.aspx
[Accessed 3 March 2014].

Lausanne, ACM (RecSys ’08).

[Accessed 7 April 2014].

Available at:
http://www.iab.net/media/file/IAB_Internet_Advertising_Revenue_Report_FY_2013.pdf
[Accessed 12 April 2014].

Available at: http://dictionary.reference.com/browse/classified%20ad
[Accessed 2 March 2014].


SaskTel, 2012. *SaskTel to introduce new TV recommendation engine.* [Online]
[Accessed 14 March 2014].

Available at: http://www.schibsted.com/en/About-Schibsted/
[Accessed 27 February 2014].

[Accessed 24 April 2014].


Shipside, S., 2012. *30 per cent of French mobile users using phones to check classifieds | AIM Group.* [Online]
[Accessed 7 April 2014].

Available at: http://www.similarweb.com/website/jofogas.hu/#/aprod.hu
[Accessed 6 March 2014].

Available at: http://www.geekwire.com/2013/zillow-mobile-stats/
[Accessed 7 April 2014].


Available at: http://www.mediapiac.com/digitalis-lap/2012-1-2-szam/Kis-penz-nagy-igeret/924/
[Accessed 6 March 2014].

Taboola, 2014. Under the Hood | Taboola - Content you may like. [Online]
Available at: http://www.taboola.com/under-hood
[Accessed 14 March 2014].

Available at: http://www.forbes.com/sites/davidthier/2012/01/27/does-google-know-more-about-you-than-you-think-take-this-test/
[Accessed 19 April 2014].

Wagner, K., 2014. How to Curate Your Facebook News Feed. [Online]
Available at: http://mashable.com/2014/01/19/facebook-news-feed-curation/
[Accessed 19 April 2014].

Available at: http://continuations.com/post/8214667310/a-brief-history-of-disruption-classified-ads
[Accessed 28 February 2014].

Available at: https://www.youtube.com/feed/recommended
[Accessed 14 March 2014].

[Accessed 12 April 2014].