

Visualization of music suggestions

A visual explanation system for collaborative filtering

Joris SCHELFAUT

Supervisor: Prof. E. Duval
Affiliation (optional)

Co-supervisor: (J. Klerkx)
Affiliation (optional)

Mentor: (J. Klerkx)
Affiliation (optional)

Thesis presented in
fulfillment of the requirements
for the degree of Master of Science
in Applied Computer Science

Academic year 2012-2013

Contents

1	Introduction	1
1.1	Music suggestions and explanation systems	1
1.2	Thesis objective	2
1.3	A multi-focal perspective	2
1.3.1	Item recommendation	2
1.3.2	Insight gaining and sensemaking	2
1.3.3	Information visualization and visual data mining	3
1.3.4	Graph drawing	3
1.3.5	Human-computer interaction and usability	3
1.4	The visual explanation system	4
1.5	Next chapters	4
2	Literature study	5
2.1	The user	5
2.1.1	Insight gaining	5
2.1.2	Interactive visualization	8
2.2	The system	11
2.2.1	Properties of recommendation algorithms	12
2.2.2	A classification of recommendation algorithms	13
2.2.3	Challenges for recommender systems	13
2.3	The interface	14
2.3.1	Types of data	15
2.3.2	Visual encoding and visual channels	15
2.3.3	Visualization techniques	16
2.4	Related work	16
2.4.1	Smallworlds	17
2.4.2	Pharos	17
2.4.3	TasteWeights	19
2.4.4	SFVis	19
2.4.5	PeerChooser	20
2.4.6	Comparative study of visual explanation systems for item recom- mendation	20
3	Designing a white box model for collaborative filtering	22

4	Iterative development	23
4.1	Methodology	23
4.2	Iteration 1: paper prototype	23
4.3	Iteration 2: digital prototype	23
4.4	Iteration 3: application implementation	23
5	Implementation: the SoundSuggest application	24
5.1	Software design and application architecture	24
5.2	Technologies	24
5.2.1	The Last.fm API	24
5.2.2	D3.js JavaScript Library	24
5.2.3	Chrome extensions	24
5.3	Implementation	24
6	Conclusion and future work	25
	References	26
	List of Figures	29
	List of Tables	30
	Index	31

Chapter 1

Introduction

1.1 Music suggestions and explanation systems

Imagine you have a music library with a number of tracks in it. No doubt you will like certain tracks more than others. At a certain point you will want to expand your library. It is only natural that you will want to add music that is similar to the music you already like, but where should you begin to look for this kind of music? Over the last decade, systems have been developed to compute what tracks, or in the more general case, items or information, would be of interest to you based on your listening history and/or track ratings. This kind of systems are called recommender systems.

In their essence, recommender systems can be seen as filters applied on a large data sets. Ever since computer engineers started to develop this kind of systems, a wide range of algorithms have been designed and implemented to compute item recommendations[5, 19, 23, 24]; each of them with their own advantages and disadvantages.

Let us assume you have plugged some recommender system into your music library and you have received a list of music suggestions. Which of these recommendations should you choose? Of course you could go through them all one by one, but that might take up quite some time. What it comes down to is that you don't know how the recommender system computed these recommendations, and as a result, you have a hard time making an educated decision where to start.

To solve this problem, you will need some kind of explanation system that provides a reasoning to arrive at the results. An ambitious approach would be to explain each step of the recommendation algorithm, but this is not always possible or desired. Indicating which of your tracks are closely related to the given recommendations, the system's confidence in the accuracy of the suggestions, et cetera, also help giving some additional context in explaining why a particular recommendation would be interesting[11]. Over the course of the last decade a wide range of explanation systems have been implemented. Many of these also use visualizations to explore user and/or item relationships[3, 6, 8, 9, 10, 22].

Let's say that you have installed the recommender system with an integrated explanation system. The explanation system visualizes how the items in your library are related to the recommendations, and provides additional statistics. Now, finding new, interesting music will hopefully become easier than ever before.

1.2 Thesis objective

The goal of this thesis is to design, implement and evaluate visualization and interaction techniques that allow the user to gain insight into the recommendation process as well as actively steer the process. The elaboration of this thesis consists out of a literature study on the topic of visualization of music suggestions, and secondly a similar application is designed and implemented[17].

In the design of the application findings of the literature study and comparative study of visual explanation systems are taken into account. From this initial design a hypothesis and expected result are derived. This hypothesis is tested through user tests. If possible, the application is improved. This process can be repeated, incrementally improving the application.

The scenario in section 1.1 sketches the context in which this application can be used.

1.3 A multi-focal perspective

Although the focus of this thesis lies on the creation of a visual explanation system for music recommendation, this thesis covers aspects of several research domains. Therefore, this text can be viewed from a multi-focal perspective. The various facets that can be distinguished are as follows: item recommendation, which is in turn a subfield of data mining, insight gaining and sensemaking, and information visualization and visual data mining. Also the graph drawing problem is an important topic in this paper. And last but not least, human-computer interaction can be seen as the overarching concept.

To make the forementioned topics more concrete in the context of this thesis, the following subsections try to explain how each of these topics is relevant, as well as explain how they are interrelated.

1.3.1 Item recommendation

Recommender systems form one of the major topics of this thesis. In particular this thesis tries to address the black box problem. In a paper by Herlocker et al. [11], recommender systems are compared to a black box. This black box generates item recommendations seemingly at random, leaving the user guessing as to how recommendations were computed. Herlocker et al. propose to create an explanation system, i.e., the white box, to overcome this black box problem.

Although there are many variations of recommendation algorithms, the focus of this thesis is on collaborative filtering-based (CF) item recommendation. This algorithm is studied in the context of music recommendation. The CF algorithm tries to establish similarities between user profiles, in this case music libraries with additional listening history and/or item ratings. Items in the difference between items sets of similar profiles are then candidate recommendations[24].

1.3.2 Insight gaining and sensemaking

Another aspect of this thesis is related to insight gaining. In order to develop an explanation system, it is interesting to explore how a user arrives at insight. We will try to define the insight gaining process and the related concept of sensemaking. Sensemaking is not

uniquely defined as its meaning may slightly vary depending on the context[34]. In short, sensemaking can be explained as the effort by an individual to understand the underlying information structures with the objective to answer task-specific questions[25, 34].

1.3.3 Information visualization and visual data mining

As the explanation system will be using visual elements, concepts from the field of information visualization will be used. We will look at general characteristics of information visualization and interactive visualization. Also, various visualization techniques, and clutter reduction techniques are discussed.

1.3.4 Graph drawing

To determine whether or not graphs are applicable to visualize a concept, Herman et al. [13] pose the following question: "is there an inherent relation among the data elements to be visualized?" As the structure behind collaborative filtering can be interpreted as a network of users and items, the answer to this question is yes. As a result, it should come as no surprise that graphs are a popular way to visualize collaborative filtering-based recommendation[3, 10], and the application developed for this thesis will also contain a graph-based visualization.

Herman et al. [13] define the graph drawing problem as follows: "given a set of nodes with a set of edges (relations), calculate the position of the nodes and the curve to be drawn for each edge". This thesis can then be seen as an effort to solve the graph drawing problem in a particular context, i.e., specific constraints in terms of screen size, data dimensionality, data quantity, et cetera.

1.3.5 Human-computer interaction and usability

In its most basic form, this is a thesis about humans and computers. The human uses a computer to find new item recommendations. The computer provides these recommendations. The human wants to understand the recommendation process. The computer explains how it computed the item suggestions. The user interacts with the system to further his/her understanding of the subject.

In this thesis we will try to understand the user, the system or computer, and the interface that allows interactions between them. Human interaction (CHI) is defined as "is the study of how people design, implement, and use interactive computer systems and how computers affect individuals, organizations and society." [31].

To improve the interaction between humans and computers, we will aim to develop an interface with high usability. The evaluation methods in this thesis draw from the evaluation techniques from the field of human-computer interaction. Paper prototyping, think aloud protocol, subjective evaluation methods among others, are examples of these techniques[20]. In this sense, this thesis can be seen as a case study of designing, implementing and incrementally improving an application using CHI techniques.

1.4 The visual explanation system

The application created for this thesis is a page action Chrome Extension that injects *HTML* and *JavaScript* into the recommendations page of *Last.fm* at <http://last.fm/home/recs>. The application makes use of several *JavaScript* libraries, such as D3¹ and jQuery², as well as a specific JavaScript library by Felix Bruns³ to facilitate using the Last.fm API⁴.

The application can be found in the Google Chrome web store⁵.

1.5 Next chapters

The rest of this thesis text is organized as follows. First we will present a literature study on the topics discussed in section 1.3 of this introduction. The next chapter is a comparative study of recommender systems with visual explanation systems. After this we will try to design a white box for CF-based recommendation by applying some of the conclusions and techniques discussed in the literature study. This concludes the theoretical part of the thesis.

In the next part of the thesis we will look at the evaluation of the application's design through user tests, as well as the software design and implementation of the application. In these chapters we will try to identify liabilities in the design, highlight implementation details, and discuss evaluation methods.

The thesis text ends with an analysis of the application's evaluation results, further conclusions, and a reflection on future work and opportunities.

¹A library using SVG, HTML and JavaScript[4]; available at: <http://d3js.org/>

²Available at: <http://jquery.com/>

³Available at: <https://github.com/fxb/javascript-last.fm-api>

⁴Available at: <http://www.last.fm/api>

⁵The SoundSuggest application can be found at: <https://chrome.google.com/webstore/detail/soundsuggest/jimblcjmmjjfaklclmohcnabndlidmb>

Chapter 2

Literature study

The term 'human-computer interaction' describes a phenomenon where two actors, the human or *user* and the computer or *system*, share a communication channel. The communication channel is a representation of data of the system towards the user. It is clear that both actors will impose certain restrictions on this visual communication channel[28, 33]. In this chapter we will try to establish who these actors are, what this communication channel looks like, and what kind of restrictions are imposed upon the visual communication channel.

2.1 The user

Who is the first actor? The answer to this question is of course very broad, so first we will try to list what we actually want to know. The goal is to allow the user to gain insight into the system through an interactive, visual explanation system. In conclusion, two specific questions arise:

- how does a human gain insight?
- what kind of limitations are imposed by the user on the design of an interactive visualization?

The following subsections try to establish an answer to these questions. Most of the ideas in these subsections are drawn from a papers by Yi. et al. [34], North et al. [21], Klein et al. [15, 16], and a book by Colin Ware [33].

2.1.1 Insight gaining

What is insight? In [21] it is argued that insight is not a well-defined term. A formal definition might be too restrictive to capture its essence, and yet too broad to be useful. To quantify insight, [21] and [34] list characteristics that allow a finer evaluation:

- **Complex:** insight is complex in the sense that involves large amounts of data that form cognitive constructs, rather than individual units;
- **Deep:** insight is self-generating in a way, as insight provides a starting point for insight on the next level;

- **Qualitative:** insight is subjective, uncertain and can have multiple levels of resolution;
- **Unexpected:** insight is usually unpredictable, serendipitous and creative;
- **Relevant:** insight is deeply embedded in the data domain: it gives data meaning as it connects data to the existing domain knowledge;

The quality of insight can then be determined by quantifying each of these characteristics[21]. The previously described collection of properties defines insight. Now we will look at the closely related concept of sensemaking. The next paragraphs describe how a user can arrive at insight.

Sensemaking

Sensemaking plays an important part in insight gaining [34]. The definitions for sensemaking may vary. We adapt the definition presented by [15] and [34].

In [15] sensemaking is looked at from a psychological perspective, a perspective of human-centered computing, and the perspective of naturalistic decision making. Sensemaking is then defined as follows: "sensemaking is a motivated, continuous effort to understand connections in order to anticipate their trajectories and act effectively"[15].

Based on the discussion in [16] and [34], Soo Yi et al. describe the process of sensemaking. Sensemaking is a:

- **Cyclic and iterative procedure:** consisting out of a generation loop searching for representations, a data coverage loop instantiating the representations and finally shift representations;
- **Creation procedure:** being more about reasoning than discovery;
- **Retrospective procedure:** as people construct a framework and assign relevant information to a place within this framework. If the data fits the framework well, the framework is confirmed, otherwise it may be updated or discarded;

An important remark made in [15] is that data fusion algorithms can reduce information overload, but they also pose challenges to sensemaking if the human can't form an accurate mental model of the machine, to understand why and how the algorithms are doing what they are doing.

Processes of insight gaining

Although sensemaking can play an important part in gaining insight, it is not the only path to arrive at insight [34]. Soo Yi et al. [34] identify four processes through which insight is established. Note that these processes are intertwined and often used together to generate insights. The processes are as follows[34]:

- **Provide overview:** in this process the individual gains understanding of the big picture of a dataset of interest. It allows the user to make a distinction between what is known to him/her and what is not;

- **Adjust:** in this process a person will explore a dataset by adjusting the level of abstraction and/or the range of selection. Typical actions involve filtering and grouping of data;
- **Detect pattern:** in this process the user will try to identify specific distributions, trends, frequencies, outliers or structure in the dataset;
- **Match mental model:** in this process the gap between data and cognitive model is bridged, reducing cognitive load and linking the present visual information with real-world knowledge.

The link with sensemaking is found in the cyclic and iterative nature of sensemaking - provide overview, adjust and detected pattern can be applied iteratively, as well as its creative and retrospective aspects - adjust and detect pattern create hypotheses and test them through various interaction techniques[34].

Improving insight

Yi et al. [34] identify several ways in which the insight gaining process can be made more efficient. They list the system's interactivity, the quality of visual encodings and usability among others, as possible enablers for increased insight gaining. Naturally, improvident designs will act as barriers rather than enablers in the insight gaining process.

Interactivity of the system promotes the user's engagement into the dataset. Spending more time with the data will allow users to form more detailed and accurate hypotheses, and as a result greater insight[34]. At the same time, while using the visualization, the user will become more skilled at a task over time. Nonetheless, bare in mind that when performing long and tedious search tasks, vigilance will become an important aspect as well in the efficiency of data exploration[33].

Similarly visual encodings that are counter-intuitive will also increase the cognitive load. Other barriers on insight gaining are clutter, occlusion and data overload[34].

Usability is another aspect that may have an impact on the insight gaining process, as controls that are hard to use will inevitably occupy some of the cognitive capacity of the user[34]. In the ISO standard ISO 9241-11, usability is defined as "the extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency and satisfaction in a specified context of use"[32].

Note that usability should not be considered a one-dimensional property of a user interface. Nielsen identifies several characteristics of usability in applications[20]:

- **Learnability:** if the system is easy to learn, the user can get started quickly;
- **Efficiency:** if the system is efficient to use, it will be possible to complete more work in less time;
- **Error rate and severity:** if the system should be robust and minimize faults;
- **Memorability:** once the system is learned, acquired skills should not be forgotten easily;
- **Satisfaction:** the system should be pleasant to use.

2.1.2 Interactive visualization

As we now have a better understanding of what insight is, we will try to establish how the insight gaining process works through visual data mining and interactive visualization. The relation between insight gaining and data visualization has been pointed out in other research. Colin Ware [33] describes interactive visualization as an "internal interface between the user and the computer in a problem solving system". Keim [14] notes that "idea behind visual data exploration, is to present data in a visual form, allowing the user to gain insight into the data, draw conclusions, and directly interact with the data".

In a chapter on visualization in [28], Tamara Munzner describes visualization as follows: "visualization allows the user to offload cognition to the perceptual system, using graphical data representations as a form of external memory. Therefore, by augmenting human capabilities, the data analyst is aided to understand, explore and form hypotheses of the data" [28]. In conclusion, the visual data exploration process can then be understood as a hypothesis generation process [14].

In what follows, we try to describe how a human interacts with interactive visualization on a cognitive level. It will be clear that some parallels can be drawn with the insight gaining process. This should come as no surprise, since these processes are intertwined [14, 33, 34].

In [33] interactive visualization is characterized by three classes of feedback loops:

- **Data selection and manipulation loop:** the user selects and moves objects that are selected through simple interactions based on eye-hand coordination;
- **Exploration and manipulation loop:** the user tries to find his/her way through a large visual data space;
- **Problem solving loop:** the user forms hypotheses about the data and refines them through an augmented visualization process.

The next subsections describe each feedback loop in greater detail.

Data selection and manipulation loop

The quality of performance of selecting and manipulating data on a screen, depends on certain factors. Colin Ware discusses the following attributes:

- **Reaction time:** This is the amount of time for a user to identify and select certain objects [33].
- **Types of interaction:** different types of interaction will have a different influence on user performance.
- **Learning:** The speed at which a user performs a task may decrease over time, as the user becomes more skilled at executing the task.

Each of these factors has been evaluated and since, they are captured in various laws. We will discuss some of them, as listed by Ware et al. in [33].

The reaction time is given by the *Hick-Hyman law*: $Time_{reaction} = a + b \log_2(C)$ with C the number of choices and a and b empirically determined constants. This formula

has been derived from experiments in which subjects had to press one of two buttons depending on the color of a light that was turned on or turned off[33].

The reaction-time may be influenced by many other factors such as the amount of visual noise, the distinctness of the signal and so on. If a person is allowed to make mistakes, the subject will respond faster, but at a cost of loss of accuracy[33]. When performing long and tedious search tasks, vigilance will become an important aspect as well[33].

In [33] different kinds of interactions are discussed such as selection, hover queries and path tracing. The time required for selecting an object in a two-dimensional space is determined by *Fitt's law*: $Time_{selection} = a + b \log_2(D/W + 1.0)$ with D the distance to the target, W the width of the target and a and b empirically determined constants[33].

A common way of selecting objects is through hover queries: the user drags a cursor over an object[33].

Another type of interaction is path tracing. The speed at which a user can trace or follow a given path is given by $v = W/\tau$ with velocity v , W the path width and τ a constant depending on the motor control system of the user[33].

When a task is repeated over time, the user will become more efficient at performing this task. Depending on the difficulty of the task, this learning effect will be more or less prominent. The *power law of practice* describes this speed up [33]: $\log(T_n) = C - \alpha \log(n)$ in which T_n is the time required to perform the task at the n -th trial, with $C = \log(T_1)$ and α the steepness of the learning curve.

Some actions are easier to learn than others. This is related to the *stimulus-response (S-R) compatibility*. The stimulus-response compatibility refers to the way in which skills that have been learned before through everyday experience, are applied to computer control design. However it is not necessary to produce a whole virtual reality to create excellent computer interfaces. Research shows that mismatches between the interface and real world experience are not necessarily detrimental to the efficiency of human-computer interaction. Ware concludes that "it would be naive to conclude that computer interfaces should evolve toward VR simulations of real-world tasks (...). The magic of computers is that a single button click can often accomplish as much as a prolonged series of actions in the real world"[33].

Exploration and manipulation loop

In the second loop the user navigates through the data space. The basic navigation control loop is described as an iterative process that involves two distinct aspects[33]:

- **Human:** the user gains understanding of the data space through a logical, spatial model. Parts of this model may be encoded in the longterm memory, on the condition that the data space is maintained for a long enough period of time;
- **Computer:** the visualization may be updated and refined from data mapped onto the spatial model.

When exploring spatial maps, a user is confronted with the *focus-context problem*, i.e., "the problem of finding detail in a larger context"[33]. The objective is to see the relation between the larger context and the details, rather than finding details. Ware goes

on to discern between spatial, structural and temporal scales in which the focus-context problem manifests itself[33].

Ware and Mitchell remark that the human visual system is already well-adapted to the spatial focus-context problem[33]. Therefore, they argue that when designing a display, it should already try to take a maximal advantage of these perceptual skills.

The time it takes to navigate between two points in the data space to gather information is called the *cost of knowledge*[33]. To keep the cost of knowledge low, various interaction techniques have been developed, such as distortion, rapid zooming, elision, and multiple windows[33]. In section 2.3 we will take a closer look at some of these techniques.

Problem solving loop

The problem solving loop can be described through means of a *visual thinking algorithm*[33]. Such an algorithm combines perceptual and cognitive actions into a process, as the user interacts with the visualization and explores the data space. As we want to keep the *cost of knowledge* low, it is obvious that the cost and time complexity of each of these actions should be kept at a minimum. The cognitive system that runs these algorithms, is made up out of several different components. Ware lists the following components[33]:

- **Early visual processing:** Early visual processing occurs at the lowest level, where elementary features such as color, texture information, local edges, and motion are extracted from the visual image.
- **Pattern perception:** Pattern perception occurs by combining the extracted features, patterns can be detected within the visual image.
- **Eye movements:** A map schedules tasks to explore proto-patterns through eye-movements. These proto-patterns correspond to the patterns that have a high probability to be relevant to the current task.
- **The intrasaccadic scanning loop:** Information is processed serially when focusing on an area of interest.
- **Working memory:** The working memory forms an intermediary between the incoming patterns and long-term memory. The currency of information in the working memory are *object files*. Object files are a combination of visual attributes and semantic meaning.
- **Mental imagery:** The ability to visualize internal images, which can be combined with external images, such as data visualization, to construct and test hypotheses about the visualized data.
- **Epistemic actions:** Epistemic actions are actions that support information search, such as eye movement, and mouse selections. Out of all epistemic actions, eye movements have the lowest cost.
- **Visual queries:** A visual query translates a hypothesis into a cognitive task. The result of this query can be a pattern or lack of a pattern. To reduce errors and increase efficiency, patterns are typically small, as the capacity of the visual working memory is rather limited.

- **Computational data mappings:** The computer maps data onto the screens. Sometimes the data is transformed before doing so. An important consideration is that the interval between the human action and the result on the screen should be low, promoting the so-called *principle of transparency*. The objective of transparency is that the user will have the illusion of directly manipulating the data on the screen.

In [33], Ware and Mitchell list ten different visual thinking algorithms. We will describe our own visual thinking algorithm in chapter 3, based on these algorithms.

2.2 The system

Before describing the visual communication channel between human and computer, it is necessary to know what exactly has to be explained through visualization. Therefore we will take a closer look at the system, i.e., the recommender system.

A recommender system is a system that computes item suggestions for users based on a ratings of related items in the user's profile and/or history. Additional information can be incorporated into the recommendation algorithm to refine suggestions. Several categorizations of these techniques are proposed in literature [3, 5, 11, 19, 26].

One of the incentives behind creating recommender system is the 'long-tail phenomenon'. This phenomenon can be explained as follows. Physical retail and warehouses can only keep a subset of all the available items in stock. These items are usually the most popular items on the market. Online vendors however, such as *Amazon*¹, can offer a vastly larger subset of these items to clients, including also less popular and/or less known items[24]. Typically the long-tail phenomenon is visualized in a graph in which items are ordered by their popularity on the horizontal axis against the popularity rating on the vertical axis, as can be seen on figure 2.1. Physical stores will offer only items in the first part of the graph, whereas the online vendors will also sell items from the remaining 'long tail' of the graph[24, 26]. Recommender systems then provide a means to find relevant items within this much larger range of items[24]. They enable to "connect supply and demand, introducing consumers to these new and newly available goods and driving demand down the tail "[1, 26].

The success of recommender systems to achieve this objective has been the subject of some research, e.g. [18] and [26]. Different classes of recommender algorithms favour different properties[5, 27]. This is elaborated further in subsection 2.2.3. For now its enough to see that the quality of recommendations is not uniquely defined. For example, in order to increase coverage of the item space, achieving high serendipity and novelty may be desired more than high recommendation accuracy, depending on the application context[27, 31, 26].

Typical applications of recommender systems are product recommenders for online retailers, movie and music recommenders such as *Netflix*² and *Last.fm*³, and news article recommenders in online news services[18, 24, 26].

Recommender system have opened up new possibilities in the landscape of online retail, and as a result, spurred the interest of businesses in this field. A remarkable initiative was

¹<http://www.amazon.com/>

²<http://www.netflix.com/>

³<http://www.last.fm/>

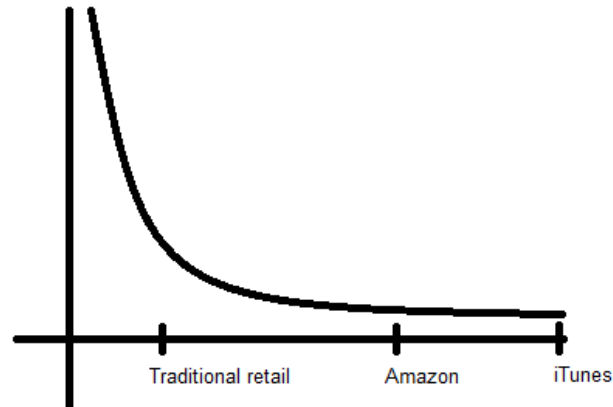


Figure 2.1: The long-tail: Items ordered by popularity are layed out against their popularity rating. Most of the items reside in the long tail of the graph. Companies such as Amazon can offer a vastly greater subset of the total item space.

the Netflix challenge. In 2006, Netflix Inc. offered a prize to beat the performance of their recommendation algorithm by 10 percent. It gave a significant boost to the research on recommendation algorithms, and yielded a winning algorithm in September 2009[2, 24].

2.2.1 Properties of recommendation algorithms

in [12] and [27], Herlocker et al. and Shani et al. respectively, compare the performance of recommender systems. They list a number of metrics for recommendation algorithms. The resulting set of properties consist out of the following characteristics:

- **Accuracy:** The accuracy of item recommendations. There are three broad classes of prediction accuracy measures:
 - the prediction of the rating given by a user;
 - the prediction whether or not a user will actually use the item (for example adding to a queue) opposed to predicting the rating itself;
 - the prediction of a ranking among items rather than an explicit rating of each item independently.
- **User preference:** The opinion of certain users may be more valuable than the opinion of others.
- **Coverage:** The proportion of items that the recommender system can recommend is referred to as catalog coverage. Another measure in this respect is the percentage of all items that are recommended to users. Finally we can also look at the diversity of the recommended items. Coverage can also mean the proportion of users or user interactions for which the system can recommend items. The cold start problem relates to coverage as it measures the coverage for a specific type of users, namely new users.

- **Confidence:** Confidence in the recommendation can be defined as the systems trust in its recommendations or predictions. The most common measurement of confidence is the probability that the predicted value is indeed true, or the interval around the predicted value where predefined portion of the true values lie. Confidence bounds can be used to filter recommended items where the confidence in the predicted value is below some threshold.
- **Trust:** Trust refers to the user's trust in the system, as opposed to confidence.
- **Novelty:** Novel recommendations are recommendations for items that the user did not know about.
- **Serendipity:** Serendipity is a measure of how surprising the successful recommendations are. One can think of serendipity as the amount of relevant information that is new to the user in a recommendation, or alternatively as deviation from the 'natural' prediction.
- **Diversity:** Diversity is generally defined as the opposite of similarity. Note that an increase in diversity may correlate to a decrease in accuracy.

2.2.2 A classification of recommendation algorithms

Based on classifications presented in [5] and [26], a categorization of different types of recommendation strategies can be identified. We will only discuss the two most prominent ones, namely collaborative filtering (CF) and content-based filtering (CB)[11, 24], and list some hybrid strategies. In the literature on recommender systems other general approaches that are commonly identified, are utility-based filtering, knowledge-based filtering, demographic filtering, and expert-based filtering[5, 3].

Collaborative recommendation

Collaborative recommendation aggregates item ratings by users. By establishing overlaps between ratings in the corresponding user profiles, the system generates new item recommendations[5, 11]. A typical user profile in a collaborative system consists of a vector of items and their ratings, that continuously augmented as the user interacts with the system over time[5].

For CF-based recommendation, there are two classes of entities: users U and items I . The data itself can then be represented by a utility matrix A . The entries $a_{i,j}$ of the utility matrix represent what is known about the degree of preference of user u_i and item i_j [24].

Content-based recommendation

Content-based recommendation learns a profile of the users interests based on the features present in objects the user has rated. New recommendations can then be generated based on a similarity function on these features[5, 23].

2.2.3 Challenges for recommender systems

Each recommendation technique has benefits as well as drawbacks. Some of these apply to all or most types of recommendation strategies, while others are only relevant to certain cases.

Both CF and CB-based recommendation algorithms suffer from the ramp-up problem in one way or the other. The 'ramp-up' or 'cold start' problem (although they may refer to slightly different problems depending on the literature) is dual problem that encompasses two distinct, yet related problems as defined in [5]:

- **New User:** when a recommender system uses ratings by its users to compute item recommendations, it is hard to find neighbours for a user, who has a limited profile. As user profiles tend to build up over, new users usually fall in this category.
- **New Item:** a new item will most likely not have that many ratings associated with it, and as a result will not be easily recommended. This 'new item problem' typically emerges when new items are constantly added to the system; for example when browsing a constant stream of news articles. When new articles are introduced, not many users have had the chance yet to rate these items. In the case of a news feed, an additional problem is that these items are short-lived, meaning that at some point these item profiles will most likely stop receiving any ratings at all.

A problem that is typical of collaborative filtering is the 'gray sheep problem'[5, 11]. The gray sheep problem occurs when a user falls between different clusters of users that may have contradicting item ratings. As a result, it is hard to determine how to classify the user[5].

Another issue with recommendation systems is that these system often appear as 'black boxes' towards the end user. The complexity of the algorithms used prevents the user from understanding the recommendation rationale[35]. This problem decreases the acceptance by the user of item suggestions. One of the solutions for this problem, proposed by Herlocker et al. in [11], is to provide an explanation system, i.e., the white box, on top of the recommender system that explains the recommendation process. This can be done through providing a transcript of the system's reasoning or through visualizations[11].

2.3 The interface

In this section we will take a closer look at the visual communication channel.

Shirley et al.[28] lists three distinctive limitations:

- **Computational capacity:** time complexity and memory usage of algorithms must allow a responsive user interface, especially in the case of interactive visualization.
- **Display capacity:** there is a trade-off between the benefits of maximizing the information density, i.e., the measure of the amount of encoded information against the amount of unused space, and causing visual overload.
- **Human perceptual and cognitive capacity:** optimizing the cognitive cost is one of the key aspects that make up a successful visualization, as visual and non-visual memory capacity are limited[33].

There has been done extensive research in this domain. Over the last two decades various new information visualization techniques have been developed. These techniques can be classified based on three criteria [14]:

- **Data:** a classification based on the structure and type of data;
- **Technique:** a classification based on characteristics of visualization techniques;
- **Interaction and distortion:** a classification based on the way in which interaction between user and visualization is enabled.

2.3.1 Types of data

Information visualization has been focusing on on data sets that lack inherent spatial semantics, thus posing a challenge to map the abstract data onto a two-dimensional screen space[14].

There are different types of data and their characteristics will have an influence on the type of visualization. Tables of data consist out of rows, representing items, and columns, representing the data dimensions, or 'attributes'. The number of dimensions is referred to as the dimensionality of the data set[14]. There are three different kinds of dimensions, namely[28]:

- **Quantitative:** numerical data on which arithmetic can be applied;
- **Ordered:** an enumeration that has a definite order;
- **Categorical:** data that has no specific ordering, and is distinguished by name only.

Relational data on the other hand consists out of nodes and links or 'edges'[14, 28]. Both nodes and edges can have associated attributes.

In [14] also text and hypertext, and algorithms and software are discussed as examples of other types of data. In the case of text and hypertext, standard visualizations are hard to use, as they cannot be described easily in terms of numbers. As a result, the data is first transformed into description vectors. Next, these vectors can be used in a visualization. Examples of software and algorithm visualization are flow diagrams, presentation using a graph-based structure of source code, and so on[14].

2.3.2 Visual encoding and visual channels

Visual encoding is defined as the mapping of data set attributes to a visual representation. The choice of visual encoding is one of the central problems in the visualization design[28].

Visual encoding takes place through visual channels. A visual encoding corresponds to a graphical element, or mark. Examples of visual channels are spatial position, color, size, et cetera. The dimension of the mark may vary: a point is a zero-dimensional mark, a line a one-dimensional one, an area a two-dimensional one and so on.

A visual encoding has the following characteristics, as described in[28]:

- **Distinguishability:** the ability of a user to distinguish between visual encodings;

- **Seperability:** Separable visual channels are opposed to integral visual channels, which are focused together on a pre-conscious level. Separable visual channels are safe to use for encoding multiple dimensions;
- **Pop-out:** selecting a channel and make it visually stands out from all the others.

There is a variety of possible visual channels that a designer can turn to in order to create a visual encoding, such as color, spatial position, size, shape, orientation, and so on. The performance of the visual encoding (through a visual channel) depends on the type of data, i.e. quantitative, ordered or categorical [28]. Figure ... gives an overview of the performance for each category, adapted from [28]. Note that spatial position is the most accurate for each data type.

In [28] colour is considered in terms of three separate channels: hue, saturation and brightness. This allows for different encodings. Just like for most visual channels, the choice of the channel (hue, saturation or brightness) depends heavily on the type of data:

- **Quantitative data:** uses a color map, a range of color values that can be continuous or discrete. It is recommended to use lightness instead of hue, as lightness has an implicit perceptual ordering. Moreover the human eye responds most to strong luminance. Hue on the other hand has a small range (around twelve values that can be reliably distinguished, including background and neutral colors);
- **Ordered data:** lightness and saturation are advised. As mentioned before, these have an implicit perceptual ordering;
- **Categorical data:** hue can be successfully applied for categorical data, keeping in mind its small range. An important remark is that roughly 10% of men is red-green color deficient. If a coding uses red and green, it may be wise to apply redundant coding using lightness or saturation in addition to hue [28].

Spatial layouts form other visual channels. Although these tend to be the most accurate, spatial layouts in two and three dimensions have several weaknesses [28]:

- **Occlusion:** parts of the data set become hidden by others. In the case of the mapping of abstract dimensions onto spatial positions, understanding the details of a three-dimensional visualization may be challenging, even if the user is allowed to change viewpoints;
- **Perspective distortion:** again, in the case of the mapping of abstract dimensions onto spatial positions, distances may convey meaning that may be distorted through perspective.
- **Text in arbitrary orientations:** special care has to be taken with text, as it may become very hard to read depending on the orientation.

2.3.3 Visualization techniques

2.4 Related work

In this section we will take a look at visual explanation systems that already exist and that have been evaluated as well. Four different applications are discussed: SmallWorlds, Pharos, TasteWeights, and SFVis.

Next a comparison is made between these systems, based on a number of objectives for explanation systems listed by Tintarev and Masthoff in [30].

2.4.1 Smallworlds

In [10], Gretarsson et al. used the Facebook API to create an application to generate social recommendations for Facebook users. Unfortunately the Facebook API does not support unauthorized reading of item preference information beyond the immediate friend group. This would not have been a problem unless traditional collaborative filtering strategies tend to produce item suggestions of inferior quality for small items. In this case however the research team relies on the social filtering through the active user's peer group[10].

SmallWorlds is "a visual interactive graph-based interface that allows users to specify, refine and build item-preference profiles" [10]. The system promotes transparency in the recommendation process, and gives the user a sense of control over the recommendation process through interactions. This way, Gretarsson et al. try to further overcome the limitations of their recommender system[10].

SmallWorlds uses a five-layered design to create suggestions:

1. the active user's node;
2. the active user's profile items;
3. friends who have items in common with the active user;
4. items that are not in the active users profile, but are liked by friends in layer 3;
5. friends who have no items in common with the active user and items in their profiles, but not items in the profiles of friends in layer 3.

The user can move nodes in each layer further or closer towards the active user's node to adjust the weights of each node. This is used in combination with similarity functions to calculate the suggestions. Figure 2.2 shows a screenshot of the application.

2.4.2 Pharos

Content-centric social websites, such as blogs and discussion forums, contain vast amounts of fast growing information. Recommendation systems have been developed to help users find the information they are looking for. The *Pharos* application tries to address two distinctive problems that present themselves in this context: the cold start problem and the black box problem[35], as described earlier in section 2.2.3.

As they hope to overcome previously defined problems, Zhao et al. [35] collect and visualize content-related social behaviour. The resulting data set is transformed into a

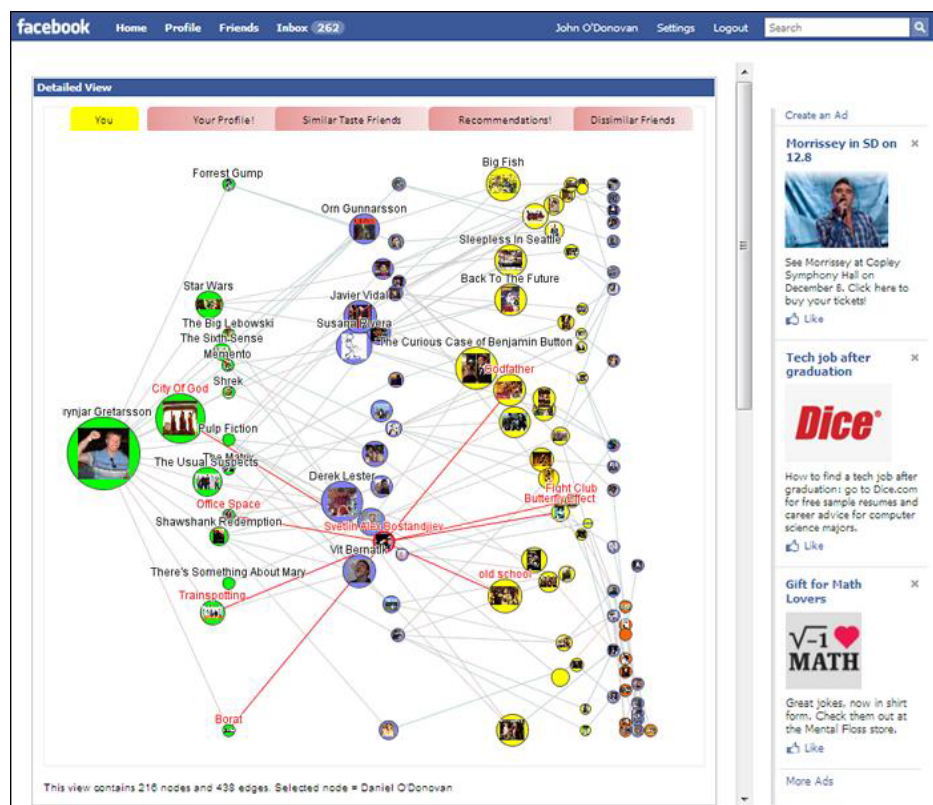


Figure 2.2: The SmallWords interface.



Figure 2.3: The Pharos social map. Colours indicate activity within a certain group.

social map. The social map provides a context for new users, addressing the cold start problem. Secondly, the user can explore the social map to increase understanding and user interaction, in an effort to overcome the black box problem.

The generation of the social map takes place through the following three step process:

- **Community extraction:** a map depicting 'which users are talking about what'. Starting from either relationships, people or content communities can be derived;
- **Community/item/people ranking:** the next step is to rank these communities. The 'hotness' can be measured on content, people authorities and so on;
- **Community labeling:** describing what each community is about.

An example of what the resulting visualization looks like, is depicted in figure 2.3.

2.4.3 TasteWeights

Social web APIs and other data sources are constantly evolving, and traditional recommender system techniques such as automated collaborative filtering need to adapt to the changing environment of the social web (cf. Smallworlds).

Bostandjiev et al. introduces two enhancements for the traditional techniques. Multiple web sources, namely from Facebook, Twitter and Wikipedia are combined when computing the recommendation. This combination provides a hybrid of different recommendation strategies, namely: collaborative filtering, expert-based and content-based respectively. The other feature is a new user interface that provides transparency into the recommendation process. The TasteWeights application uses this approach.

There are three levels to be distinguished that are represented visually as well:

A profile layer: liked items on Facebook, A context layer: items coming from different sources, A recommendation layer containing the actual recommendations. Nodes of a

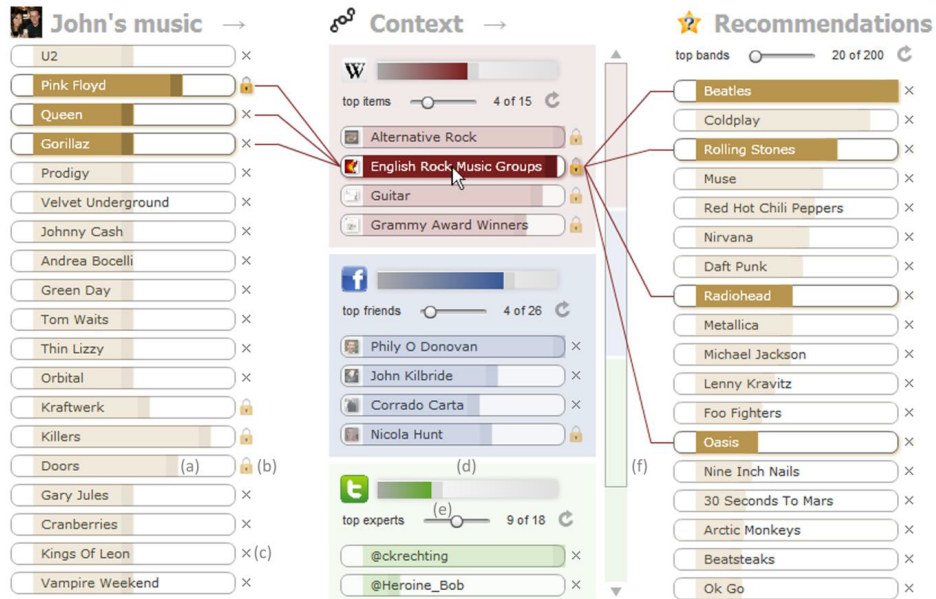


Figure 2.4: The TasteWeights interface.

graph reside in one of these layers. The user can influence the outcome of the recommendation by attributing weights to the nodes in the profile layer and context layer.

2.4.4 SFVis

SFVis (Social Friends Visualization) is an application that helps users explore and find friends interactively under a context of interest. The system is a hybrid approach of social tags and social networks. The SFVis framework transforms a data model (social tags and social networks) into a visual form. Users can both manipulate the input and visuals on demand.

Social tags can form a network. Within this structure clusters may arise. From this cluster tag network a hierarchy is derived. A compound graph is generated from the tag hierarchy and social networks.

A mapping function assigns actors in a social network to a tag tree. The actor similarity algorithm in SFVis considers both structure similarity in a social network and semantic similarity in a tag network. These scores will allow the recommendation system to compute friend suggestions.

SFVis uses circular visualizations for the different trees and graphs for both views as well as interaction with the user.

2.4.5 PeerChooser

Due to black-box nature of collaborative filtering O'Donovan introduces PeerChooser. This system provides visual explanations for collaborative filtering recommendations and enables users to manipulate their neighborhood at varying levels of granularity.

This way the researchers hope to overcome the gray-sheep problem, defined earlier in section 2.2.3. This occurs when a users history does not help to identify a set of similar

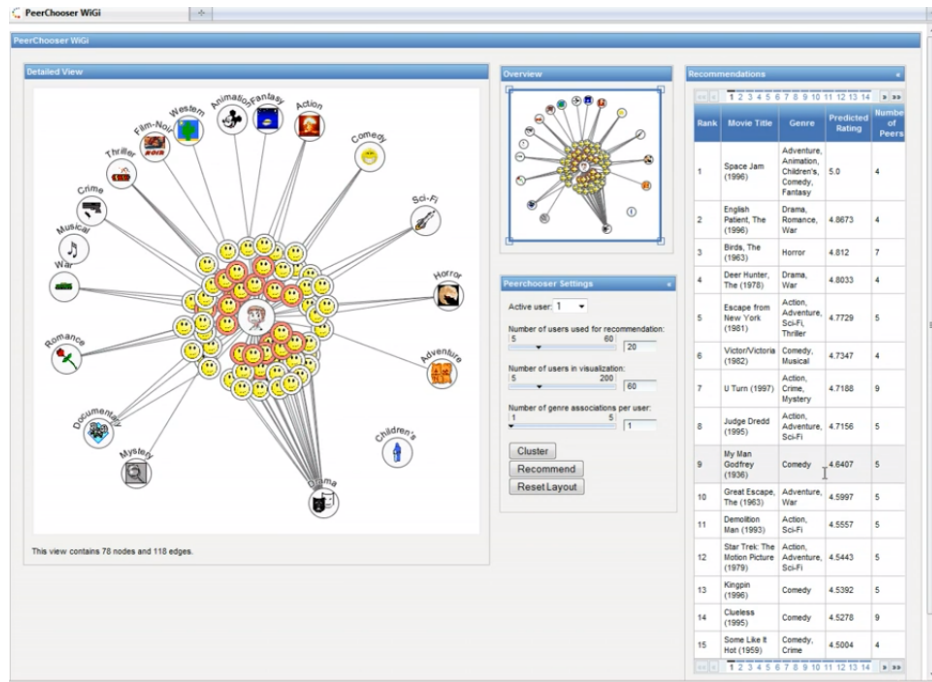


Figure 2.5: The PeerChooser interface.

recommendation partners.

2.4.6 Comparative study of visual explanation systems for item recommendation

The following listing is adapted from [30]:

- **Transparency:** explain how the system works;
- **Scrutability:** allow users to tell;
- **Trust:** increase users' confidence in the system;
- **Effectiveness:** help users make good decisions;
- **Persuasiveness:** convince users to try or buy;
- **Efficiency:** help users make decisions faster;
- **Satisfaction:** increase the ease of usability or enjoyment.

In table 1 an overview is presented of some the different recommender systems and some of their properties. So far most of the applications discussed here used collaborative filtering and some kind of visualization and allowed the user to interact with the system.

Chapter 3

Designing a white box model for collaborative filtering

In this chapter we will describe our approach to design a white box model for collaborative filtering. In order to provide explanations about the recommendation process, we will base our design on the characteristics of collaborative filtering.

The underlying structure of collaborative filtering can be interpreted as a dual graph. This is a graph $G(V, E)$ for which $V = U \cup I$ such that $U \cap I = \emptyset \wedge E \subseteq U \times I$ [7]. In this case the set of nodes U corresponds to the set of users, and the other set of nodes I is set of items. In plain language this means that there only exist edges of that go from an item to a user or from a user to an item. A quantification of the similarity between users can then be identified as the number of edges leading to common items within user profiles. This way we hope to be able to answer most of the questions raised in the previous paragraphs, describing Herlocker's white box model.

One of the challenges of this approach is overcoming the graph drawing problem, as defined earlier in section ???. In an application where easily millions of items may be involved, scalability becomes a significant constraint on the visualization design[13].

Based on a visualization design by Valdis Krebs, a dimensionality reduction can be performed on the graph, by keeping only one set of nodes and representing the other set of nodes as implicit information in the edges. In [29] Valdis Krebs kept the items, books purchased from the Amazon web store in this case, and represented the number of users that where linked to these items, i.e., purchased a particular book, by the thickness of the edges.

Chapter 4

Iterative development

4.1 Methodology

4.2 Iteration 1: paper prototype

4.3 Iteration 2: digital prototype

4.4 Iteration 3: application implementation

Chapter 5

Implementation: the SoundSuggest application

5.1 Software design and application architecture

5.2 Technologies

5.2.1 The Last.fm API

5.2.2 D3.js JavaScript Library

5.2.3 Chrome extensions

5.3 Implementation

Chapter 6

Conclusion and future work

Bibliography

- [1] C. Anderson. *The Long Tail: Why the Future of Business Is Selling Less of More*. Hyperion, 2006.
- [2] R. M. Bell and Y. Koren. Lessons from the netflix prize challenge. *SIGKDD Explor. Newsl.*, 9(2):75–79, Dec. 2007.
- [3] S. Bostandjiev, J. O’Donovan, and T. Höllerer. Tasteweights: a visual interactive hybrid recommender system. In *Proceedings of the sixth ACM conference on Recommender systems*, RecSys ’12, pages 35–42, New York, NY, USA, 2012. ACM.
- [4] M. Bostock. D3.js - data-driven documents. URL: <http://d3js.org/>, 2012. [Online; accessed 26-December-2012].
- [5] R. Burke. Hybrid recommender systems: Survey and experiments. *User Modeling and User-Adapted Interaction*, 12(4):331–370, Nov. 2002.
- [6] T. Crnovrsanin, I. Liao, Y. Wuy, and K.-L. Ma. Visual recommendations for network navigation. In *Proceedings of the 13th Eurographics / IEEE - VGTC conference on Visualization*, EuroVis’11, pages 1081–1090, Aire-la-Ville, Switzerland, Switzerland, 2011. Eurographics Association.
- [7] K. Dekimpe and B. Demoen. Fundamenten voor de informatica. URL: <http://people.cs.kuleuven.be/~bart.demoen/FVI/fundamenten.pdf>, 2007. [Online; accessed 9-February-2013].
- [8] S. Faridani, E. Bitton, K. Ryokai, and K. Goldberg. Opinion space: A scalable tool for browsing online comments. *CHI 2010 : Understanding Comments*, 2010.
- [9] L. Gou, F. You, J. Guo, L. Wu, and X. L. Zhang. Sfviz: interest-based friends exploration and recommendation in social networks. In *Proceedings of the 2011 Visual Information Communication - International Symposium*, VINCI ’11, pages 15:1–15:10, New York, NY, USA, 2011. ACM.
- [10] B. Gretarsson, S. Bost, C. Hall, and T. Höllerer. Smallworlds: Visualizing social recommendations. *Eurographics/ IEEE-VGTC Symposium on Visualization 2010*, 2010.
- [11] J. L. Herlocker, J. A. Konstan, and J. Riedl. Explaining collaborative filtering recommendations. In *Proceedings of the 2000 ACM conference on Computer supported cooperative work*, CSCW ’00, pages 241–250, New York, NY, USA, 2000. ACM.

- [12] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl. Evaluating collaborative filtering recommender systems. *ACM Trans. Inf. Syst.*, 22(1):5–53, Jan. 2004.
- [13] I. Herman, G. Melançon, and M. S. Marshall. Graph visualization and navigation in information visualization: A survey. *IEEE Transactions on Visualization and Computer Graphics*, 6(1):24–43, Jan. 2000.
- [14] D. A. Keim. Information visualization and visual data mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, Jan. 2002.
- [15] G. Klein, B. Moon, and R. R. Hoffman. Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems*, 21(4):70–73, July 2006.
- [16] G. Klein, B. Moon, and R. R. Hoffman. Making sense of sensemaking 2: A macrocognitive model. *IEEE Intelligent Systems*, 21(5):88–92, Sept. 2006.
- [17] KULeuven. ”masterproef t313 : Visualisatie van muziekaanbevelingen”. URL: <https://www.cs.kuleuven.be/cs/studenten/eindwerken/20122013/onderwerpen/individueel/T313.shtml>, 2008. [Online; accessed 10-October-2012].
- [18] M. Levy and K. Bosteels. Music recommendation and the long tail. URL: <http://womrad.org/2010/papers/1.pdf>, 2008. [Online; accessed 26-April-2013].
- [19] P. Melville, R. J. Mooney, and R. Nagarajan. Content-boosted collaborative filtering for improved recommendations. In *Eighteenth national conference on Artificial intelligence*, pages 187–192, Menlo Park, CA, USA, 2002. American Association for Artificial Intelligence.
- [20] J. Nielsen. *Usability Engineering*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 1993.
- [21] C. North. Toward measuring visualization insight. *IEEE Comput. Graph. Appl.*, 26(3):6–9, May 2006.
- [22] J. O’Donovan, B. Smyth, B. Gretarsson, S. Bostandjiev, and T. Höllerer. Peerchooser: visual interactive recommendation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI ’08, pages 1085–1088, New York, NY, USA, 2008. ACM.
- [23] M. J. Pazzani and D. Billsus. The adaptive web. In P. Brusilovsky, A. Kobsa, and W. Nejdl, editors, *The adaptive web*, chapter Content-based recommendation systems, pages 325–341. Springer-Verlag, Berlin, Heidelberg, 2007.
- [24] A. Rajaraman, J. Leskovec, and J. Ullman. *Mining of Massive Datasets*. Cambridge University Press, 2012.
- [25] D. M. Russell, M. J. Stefik, P. Pirolli, and S. K. Card. The cost structure of sense-making. In *Proceedings of the INTERACT ’93 and CHI ’93 Conference on Human Factors in Computing Systems*, CHI ’93, pages 269–276, New York, NY, USA, 1993. ACM.

- [26] scar Celma Herrada. Music recommendation and discovery in the long tail. http://mtg.upf.edu/static/media/PhD_celma.pdf, 2008.
- [27] G. Shani and A. Gunawardana. Evaluating recommendation systems. In F. Ricci, L. Rokach, B. Shapira, and P. B. Kantor, editors, *Recommender Systems Handbook*, pages 257–297. Springer US, 2011.
- [28] P. Shirley and S. Marschner. *Fundamentals of Computer Graphics*. A. K. Peters, Ltd., Natick, MA, USA, 3rd edition, 2009.
- [29] J. Steele and N. Iliinsky. *Beautiful Visualization: Looking at Data through the Eyes of Experts*. O’Reilly Media, Inc., 1st edition, 2010.
- [30] N. Tintarev and J. Masthoff. A survey of explanations in recommender systems. In *Proceedings of the 2007 IEEE 23rd International Conference on Data Engineering Workshop*, ICDEW ’07, pages 801–810, Washington, DC, USA, 2007. IEEE Computer Society.
- [31] K. P. Tripathi. Article: A study of interactivity in human computer interaction. *International Journal of Computer Applications*, 16(6):1–3, February 2011. Published by Foundation of Computer Science.
- [32] UsabilityNet. Usabilitynet: International standards. URL: http://www.usabilitynet.org/tools/r_international.htm#9241-11, 2006. [Online; accessed 20-February-2013].
- [33] C. Ware. *Information Visualization: Perception for Design*. Morgan Kaufmann Publishers Inc., San Francisco, CA, USA, 2004.
- [34] J. S. Yi, Y.-a. Kang, J. T. Stasko, and J. A. Jacko. Understanding and characterizing insights: how do people gain insights using information visualization? In *Proceedings of the 2008 Workshop on BEyond time and errors: novel evaluation methods for Information Visualization*, BELIV ’08, pages 4:1–4:6, New York, NY, USA, 2008. ACM.
- [35] S. Zhao, M. X. Zhou, Q. Yuan, X. Zhang, W. Zheng, and R. Fu. Who is talking about what: social map-based recommendation for content-centric social websites. In *Proceedings of the fourth ACM conference on Recommender systems*, RecSys ’10, pages 143–150, New York, NY, USA, 2010. ACM.

List of Figures

2.1	The long-tail: Items ordered by popularity are layed out against their popularity rating. Most of the items reside in the long tail of the graph. Companies such as Amazon can offer a vastly greater subset of the total item space.	11
2.2	The SmallWords interface.	18
2.3	The Pharos social map. Colours indicate activity within a certain group. .	18
2.4	The TasteWeights interface.	19
2.5	The PeerChooser interface.	20

List of Tables

Index

- black box, 14
- cold start, 14
 - new item, 14
 - new user, 14
- cost of knowledge, 10
- dimensionality, 15
- Fitt's law, 9
- focus-context problem, 9
- graph
 - edge, 15
 - node, 15
- gray sheep, 14
- Hick-Hyman law, 8
- information visualization, 15
- infovis, *see* information visualization
- insight, 5
- Long Tail, 11
- Netflix challenge, 12
- object file, 10
- Pharos, 17
- power law of practice, 9
- principle of transparency, 11
- ramp-up, *see* cold start
- recommendation algorithm, 11
 - collaborative filtering, 13
 - content-based filtering, 13
- recommender system, 11
- relational data, 15
- sensemaking, 6
- SmallWorlds, 17
- stimulus-response compatibility, 9
- usability, 7
- utility matrix, 13
- visual data exploration, *see* visual data mining
- visual data mining, 8
- visual query, 10
- visual thinking algorithm, 10
- visual working memory, 10
- visualization, 8
 - interactive visualization, 8
- white box, 14

AFDELING
Straat nr bus 0000
3000 LEUVEN, BELGIË
tel. + 32 16 00 00 00
fax + 32 16 00 00 00
@kuleuven.be
www.kuleuven.be

