**Abstract**

 Recommender systems are a key aspect of the e-commerce industry. With this research project, we sought to discover how recommender systems are working or not working from the users’ perspective, specifically in terms of privacy. A recommender system is an information filtering system that takes user information and presents recommendations for products to them. One of the biggest issues involved with recommender systems is user privacy. We asked users about recommender systems in general, specifically Netflix, Amazon, and Pandora, and found that users have a favorable opinion of recommender systems. When we asked users about privacy, we found that they were not likely to read the privacy policy, were concerned about privacy in general, but were not concerned about the privacy of their personal information when it was used in recommender systems. We concluded that despite the problems that some users had with recommender systems, they still enjoy them overall and are likely to use them. We also concluded that though users were concerned about privacy, the issues with privacy concerning recommender systems would not prevent them from continuing to use the recommender systems. Our recommendations for future work include researching issues involved with multiple users on one account and further research involving privacy and the individual privacy policies of the companies who utilize recommender systems.

**Introduction and Statement of the Research Problem**

 Computers along with the Internet have changed the way we as people do things in a society. One of the areas affected the most is the way we think about business and because of that, the way we consume goods and services. Businesses have begun the switch to an online model of business using the Internet to perform electronic based sales as the primary source of profit. Companies like Amazon, Google, Facebook, Netflix, and Pandora are among the most successful companies in the world and are all classified as being a part of the new industry known as e-commerce. These companies capitalize on their market by collecting, storing, and exchanging personal information that they have acquired through user input in order to increase income through marketing and increase customer satisfaction (Boritz and No, 2012). Users create accounts and enter in basic user information that is inputted into a massive database along with all other users. From there, companies design a recommendation system using unique algorithms to try to predict products that they think users will be interested in buying or using. After the user has created an account, the system draws in more information and gets more accurate by using previous purchases, things you like, and things that people similar to you have purchased or liked.

Recommendation systems created a new shopping experience that users have slowly been getting accustomed to over the past twenty years. As the industry has grown, so have the concerns over privacy. Personal information is spread all over the Internet, often times residing in massive user databases, including the ones used in recommender systems. Users’ personal information, which when used properly can benefit both the user and the company, can oftentimes end up in the wrong hands through hacking or illegal business transactions. Because of this, companies that do not do a good job of protecting their users personal data find themselves in lawsuits due to a violation of the privacy policy or even simply failing to secure private information. Netflix has experienced several lawsuits over the past few years including two lawsuits that resulted in a combined ten million in court fines (Mlot, 2012). One of those two cases portrayed just one of the many potential security issues involved in the sharing of personal information over the Internet. Netflix intentionally gave out private information for a contest aimed to improve their recommender system, which was used by two researchers to identify specific users based off of their reviews and history (Northrup, 2009). Using user data to identify specific users, while it can be beneficial for both parties, can also classify users by characteristics that are unwanted such as sexual orientation, like in the Netflix lawsuit, or economic status.

Understanding user perspectives on recommendation systems would provide some insight on what could be done to improve the systems regarding customer service and compatibility. What are the benefits of recommendation systems and what are the drawbacks? Recommender systems make usability much better for customers and increases profits for companies, but are companies wasting millions of dollars on something that users really don’t need or want? Additionally, what are the ethical implications concerning privacy of user personal information? Privacy policies are oftentimes lengthy and challenging to comprehend, which may be why, “many online consumers are unaware of the collection and use of their online data” (Bowman, 2012). Can companies find a way to make their policies more explicit to users? These are all questions that will be addressed as we attempt to determine what characteristics of recommender systems users find most appealing and what things could be changed to enhance user satisfaction and privacy.

**Background and Related Work**

A recommender system is an information filtering system used by companies to create higher usability and user satisfaction, as well as ultimately increasing profits for companies. They do this by managing large databases and filtering out what things consumers like from what they dislike and then suggesting them to the consumer. There are three main types of recommender systems: Collaborative, Content Based, and Hybrid filtering systems. In our research, we focused on three recommender systems: Amazon, Netflix, and Pandora. Amazon and Netflix use a collaborative filtering system to recommend their products to customers, while Pandora uses a content based filtering system to recommend their products to customers.

Content Based filtering analyzes comparisons between the content. Pandora analyzes comparisons between the style of music including the tempo, instruments used, and genre and recommends other songs based on the song or artist that the user inputs and liking or disliking songs (Grey, 2009). A Collaborative filtering system typically makes its recommendations based off of an algorithm that rates things you purchase/use, things you’ve liked, things you’ve rated positively with a higher score, and things you’ve disliked or rated poorly with a lower score. It then compares your information with thousands of others who have similar interests as you and begins to recommend things that it thinks you would be interested in. Amazon uses their recommendation system to recommend products to users for purchase. According to Amazon.com, they want users to get the maximum potential of their recommender system. Users have the ability to fill out questionnaires as well as the ability to rate products to build a more accurate representation of their interests. Netflix, in a similar fashion to Amazon, states on their website that they want users to have a better understanding of what they are doing so that they can build a better rapport with their clients, encouraging them to give more feedback effectively improving the recommendations that are provided for them (Amatriain and Basilico).

While both systems do have a encouraging policy on ways to improve their results, they also both have many drawbacks. Netflix, for example, uses a system that is intended to recommend for the whole family, providing a top ten list that consists of recommendations for all family members of various ages and genders (Amatriain and Basilico). This system, however can cause problems with parental controls giving the children in the family access to inappropriate content in addition to recommending irrelevant shows and movies simply because you can not personalize accounts within the main account. Amazon suffers similar problems with shared accounts between spouses and other family members who receive recommendations that do not apply to them. Both sites policies on privacy address it as a serious issue. On Amazon.com they offer suggestions on how to opt out of the recommendation process so that the user does not have to input any personal information other than that, which is required for purchasing and shipping. It also describes their policy on keeping record of your browsing history as, “our record of your activity on our site expires anywhere between a few days to a few weeks, depending on your level of activity.” Deleting old browsing history is an effective way to cut down the potential risk of having people’s personal data compromised.

**Research Methodology**

 At the beginning of our project we had previously just been aware of recommender systems and what they did for us individually and not on a larger scale. Starting out, we researched how they work and what other kinds of purposes they are intended for. As mentioned previously, we found out the basic information about recommender systems such as what types there are, how companies use them, how recommender systems are built, etc. After we gained some initial knowledge on what recommender systems consisted of, we started addressing what we going to search for in this project.

First, we decided that we needed to get some information from people just to see what knowledge people had about recommender systems and identify some potential aspects that we could focus on. Therefore we decided to send out the first survey (Appendix 1), which will be discussed in more detail later, with the intention of finding out what people know about recommender systems. We asked if they had encountered a recommender system, which ones had they used before, and then comments about the ones they had. Then we had a scale questionnaire for what we thought would be the four most encountered about different functions of a recommender system. To get the most results possible we decided to use social media as our main method of exposure. Facebook, Twitter, Google+, and some emails and texts were all used to bring these people in. After analyzing feedback from the first survey we decided to primarily focus on the security aspect of recommender systems, while primarily focusing on the three most popular recommender systems from the survey.

 Since we were narrowing our sphere of focus, we went back to researching more to find out information on ethical issues involving recommender systems. We read the privacy agreements for the three largest sites, and looked into cases involving issues with those agreements. Eventually, we sent out a second survey (Appendix 2) to gather more data about users on ethical/privacy aspects of recommender systems. We wanted to know what specific issues people had with the systems so we asked questions about different parts of recommender systems. At the beginning of the survey we asked scale questions ranging involving a one to five aspect in word form, and at the end of the survey we had statements that were based off of the previous questions. We had questions pertaining to the same aspect in different forms to see how much a conflict of interest there was involving different aspects of recommender systems compared to the overall opinion of recommender systems.

The overlapping questions were a key part in helping identify some of the larger issues involving recommender systems because the results were able to show how a company could improve the system that it uses to make users feel more comfortable. The larger than expected amount of responses we got for each survey further enhanced the issues that we were able to identify. However, a majority of our responses were limited to a certain demographic and that most likely resulted in similar answers for certain questions and influenced our conclusions to some degree.

**Analysis and Results**

In our first survey, we asked a series of questions about recommender systems in general and about specific recommender systems: Amazon, Netflix, Pandora, and Last.fm. From the first survey, we learned that 98.6% of users surveyed had encountered a recommender system online and Amazon, Netflix, Pandora, and Facebook Friends Suggestion were the most popular recommender systems. Thus, we decided to focus our research on recommender systems of Amazon, Netflix, and Pandora because the majority of users used those three.

Netflix users rated its recommender system as very good overall, excellent in terms of usefulness,, good in terms of accuracy, and said that they were very likely to purchase or use the suggestions. Amazon users rated its recommender system as good overall, excellent in terms of usefulness, very good in terms of accuracy, and said that they were likely to purchase or use the suggestions. Pandora users rated its recommender system as excellent overall, excellent in terms of accuracy, very good in terms of accuracy,, and said that they were very likely to purchase or use the suggestions.

It seemed like most users have experienced and are familiar with the recommender systems of Amazon, Netflix, and Pandora, and they were rated as “not the most accurate” but “very useful anyway” because the suggested items were what they usually needed recommendations of. However, recommender system can be improved in terms of accuracy.

In our second survey, we asked questions about the ethical issues involved with recommender systems, mainly how users' information is used to recommend items to them. From this, we have learned that 86.2% of users were concerned about giving up their private information to companies when creating an online account when signing up for websites like Amazon, Pandora, and Netflix; however, 83% of users are Not at all Likely to read the privacy policy.

For recommendation systems to be beneficial for both the users and the companies, there should be understanding and agreement between the user and the company about privacy policy of their rights, duties and limits. The majority of users didn't care or were neutral about websites such as Amazon, Netflix, and Pandora using their information to recommend products to them. We asked the users about which information they were uncomfortable about websites such as Amazon, Netflix, and Pandora using in order to make recommendations to them. The categories included gender, age, geographical location, race/nationality, occupation, previous purchases/activity, and interests. We learned that geographical location was the category of information that users were most concerned about.

The majority of users strongly agreed with the statement, "rental information is personal data," as well as, “personal preferences for items on websites such as Netflix, Amazon, and Pandora could be used to identify each user.” However, they also agreed that they are comfortable with websites such as Netflix, Amazon, and Pandora using their personal information to persuade them to purchase or view items, and were neutral about the statement that their personal information is safe and secure in a company's database.

It seems like they need to worry about privacy issues because the survey shows that most users are concerned about their privacy but were not aware of the privacy policy and did not read the terms when they signed up for a new account. Recommender systems work best and are kept secure when there is understanding and agreement between the user and the company about the privacy policy.

**Conclusions and Recommendations for Future Work**

Through our research, we were able to conclude that a large majority of users have encountered and used recommender systems online. In general, users have favorable opinions of recommender systems, especially Amazon, Netflix, and Pandora and in terms of usefulness and overall opinion. Despite this overall favorable opinion, however, users still believe that these recommender systems can be improved in terms of accuracy. But, despite this, users report still being likely to use or purchase items recommended to them.

 In terms of privacy, we can conclude that most users do not read the privacy policies of websites such as Amazon, Netflix, and Pandora when signing up for online accounts. Nevertheless, most do not have strong opinions about the general safety of their personal information with these websites. There are some users who are concerned about their privacy online, but the majority are comfortable with websites using their information to recommend products to them and will therefore probably continue to use the recommender systems.

 At the beginning of our research project, we sought to identify a recommender system that rose above the others in terms of user opinion and privacy. We found that Amazon was the most popular recommender system, as well as a favorite among a considerable amount of users. Nevertheless, there was not a clear winner when we asked users to rate Amazon, Netflix, and Pandora on their overall opinion of the recommender system, the system’s accuracy, the system’s usefulness, and the user’s likelihood to purchase or buy items recommended to them. Moreover, since our survey about privacy did not include any system-specific questions, it is not clear from our research whether or not Amazon, Netflix, or Pandora excels in terms of privacy. Future work including a sharper focus on this question could cause a clear answer about which recommender system is the best to emerge.

 In addition to the question of which recommender system rises above the rest, several considerations for future work became evident to us at the end of our research. First of all, in our surveys, we failed to ask about demographic details such as age. In future research about recommender systems from the user’s experience, details about age could reveal whether or not a user’s opinions regarding privacy had a correlation with that user’s age, specifically whether or not a user has concerns about their privacy. We could also look into specific information in privacy policies that users could be missing when they fail to read the policies and how not knowing this information could affect them and their data. Moreover, it would be helpful to ask questions about why the users who said that they weren’t likely to read the privacy policies. It is possible that some users do not read them because they feel as though the companies would not be able to keep their information safe regardless.

 When dealing with recommender systems like Netflix and Amazon, future work could take into account the issue of multiple users on a single account. For example, Netflix allows for multiple users with multiple rental queues. Some users in our research expressed concerns about other users within the same account being able to see that user’s preferences, purchase or rental history, etc. The issue of gift purchases on Amazon was a concern as well. If a user is purchasing a gift for another user within the same account, the user who is receiving the gift might be able to see it before he or she receives it for the holidays. Gift purchasing for users outside of the account could also affect preferences, since they would be taken into account alongside the user’s preferences for purchasing items for him- or herself. With Netflix, multiple queues raises the concern of multiple ratings and preferences being taken into account. What if a child’s preferences for animated children’s films affects the recommendations presented to an adult in the family? In future, research could be undertaken to analyze specific issues with multiple-user accounts and how these features could be improved in order for them not to affect the accuracy and usefulness of the recommender systems.

**Citations**

Amatriain, X., & Basilico, J. (2012, April 6). *Netflix recommendations: Beyond the 5 stars (part 1)*. Retrieved from http://techblog.netflix.com/2012/04/netflix-recommendations-beyond-5-stars.html

Boritz, J., & Won Gyun, N. (2011). E-Commerce and Privacy: Exploring What We Know and Opportunities for Future Discovery.*Journal Of Information Systems*, *25*(2), 11-45.

Grey, T. (2011). Pandora pulls back curtain on its magic music machine. *Fast Company.*

Retrieved from http://www.fastcompany.com/1718527/pandora-pulls-back-curtain-its-magic-music-machine

Haw, K., & Watermiller, A. Comparing content filtering and collaborative filtering in recommender systems: A problem analysis.

Herthner, N. K. (2012). Recommender systems: A true smorgasbord of options. *Searcher*, *20*(6).

Howe, M. (n.d.). *Pandora's music recommender*. Informally published manuscript, University of Washington, Seattle, WA, Retrieved from http://www.cs.washington.edu/education/courses/csep521/07wi/prj/michael.pdf

Krause, S. (2006, January 30). *Pandora and last.fm: Nature vs. nurture in music recommenders*. Retrieved from http://blog.stevekrause.org/2006/01/pandora-and-lastfm-nature-vs-nurture-in.html

Linden, G., Smith, B., & York, J. (2003). Amazon.com recommendations: Item-to-item collaborative filtering. *Internet Computing, IEEE*, *7*(1), 76-80.

MacManus, R. (2009, January 28). *5 problems of recommender systems*. Retrieved from http://readwrite.com/2009/01/28/5\_problems\_of\_recommender\_systems

Northrup, L. (2009, Dec. 19). Closeted lesbian sues netflix for potential outing. Retrieved from http://consumerist.com/2009/12/19/closeted-lesbian-sues-netflix-for-possible-accidental-outing/

*Popular recommender systems*. (n.d.). Retrieved from http://www.deitel.com/ResourceCenters/Web20/RecommenderSystems/PopularRecommenderSystems/tabid/1312/Default.aspx

Ramakrishnan, N., Keller, B. J., Mirza, B. J., Grama, A. Y., & Karypis, G. (2001). When being weak is brave: Privacy issues in recommender systems. *IEEE Internet Computing*.

*Recommendations*. (n.d.). Retrieved from http://www.amazon.com/gp/help/customer/display.html?ie=UTF8&nodeId=13316081

*Recommender systems*. (n.d.). Retrieved from http://recommender-systems.org

Wen, Z. (2008). *Recommendation system based on collaborative filtering*. Informally published manuscript, Stanford University, Stanford, CA, Retrieved from http://cs229.stanford.edu/proj2008/Wen-RecommendationSystemBasedOnCollaborativeFiltering.pdf

Appendix A:



Appendix B:



