mPortal

White Paper

Solving the Discovery Problem
A case study of the mobile application market

Josh Schoenwald
UI Design Lead
mPortal, Inc.
mportal.com
@mportal
facebook.com/mportalinc

January 2012

INTRODUCTION

The challenge of discovery has always been a difficult problem to solve. The proliferation of various forms of media has reached unprecedented levels over the last few decades. Where before, a typical consumer's biggest array of choices might have been what color paint to use or what brand of cereal to buy, today we sport music catalogs tens of millions of albums strong, hundreds of thousands of software titles, and thousands of movies and television shows—all available instantly at our fingertips.

This white paper explores the fundamental problems inherent in the problem of discovering new content, using the example of the mobile applications market. However, these same principles will apply to any other market.

THE CURRENT LANDSCAPE

A popular saying teaches us that more choice is better, but it's important to realize the context in which this statement is true. As the logic goes: greater choice is good because it's likely that amongst all choices there will be some which are truly exceptional, and it provides an environment where everyone can get what they want. While the veracity of that statement as it applies to consumer electronics can be debated, there's no doubt it holds true for media. As music, movies, and books have done in the past, so the world of software applications is now experiencing an incredible content explosion. Amidst this abundance are some true masterpieces—there *is* something for everyone. But the $64,000 question is: how do you find the ones which aren't in the Top Ten lists or featured in popular blogs?

Perhaps the best parallel to the current app market is the music industry. Songs and apps are very similar in a lot of important ways, namely:

- The price each ($1 for songs and just a bit higher for apps[1])
- Customers easily acquire them by the hundreds or thousands

Music	Average Price	Apps
$0.99	Average Price	$1.00+
Thousands	Quantity owned	Hundreds
Easy	Portability	Easy
Multiple times daily	Re-use	Multiple times daily

Figure 1: Similarities between music and mobie applications.

- They are with people all the time
- They are often used repeatedly

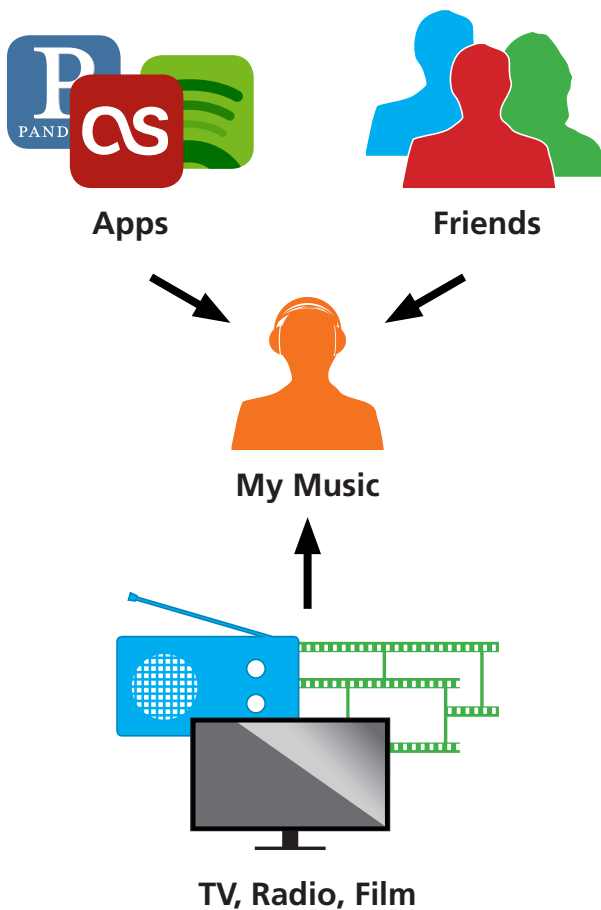**Apps**   **Friends**

My Music

TV, Radio, Film

Figure 2: Sources of music discovery and recommendation.

Although the scale is certainly not the same for both apps and music, and there as yet is no mainstream concept of an "album" of apps, the similarities are closer than for just about any other media category, so discussing them in terms of the problem of discovery is apt. The nut of the problem is that there is far more content available than is logical to sift through. And not only that, but there is often a great amount of similarity among related content, making it laborious to determine what you're really looking for. Music has some advantages in this regard, because it's everywhere (radio, television, movies, internet, stores, restaurants. etc…) and to experience it you need only your ears and your partial attention. Add to that the fact that it's easy to preview a song before you buy it, and we can see how the problem of app discovery becomes even more complicated.

But even for music, there's still the problem of finding out about great new content. Trolling through iTunes and sampling songs at random is not likely to be productive. So how has music solved this problem?

We hear about music from our friends. Everyone knows at least one person who's a music aficionado and who gives great recommendations. We read reviews of artists and albums that compare their music to other artists and albums. We use services like Pandora and Last.FM to help us discover new music. And we share our discoveries and favorites with our friends on various social networks. We even get music from commercials, movies, and television shows. Altogether we've created a fairly comprehensive network of sources that collaborate to give us the music we want. And while it's certain that we are missing out on at least some music that we'd surely love, that owes more to the fact that the available content is effectively infinite and growing faster than ever.

THE CORE OF THE PROBLEM

So too is the app market growing, though it still (and always will) pale in comparison to the music market. But where is our recommendation network for apps? Is it services like Appsfire and Chorus? Is it blogs like Engadget and Phandroid? What about social networks like Facebook and Twitter? Or maybe it's your cousin who jailbroke his phone and lets you try out apps for free? Maybe it's all of these?

To answer the question, we have to get to the core of discovery, which is the ability to find things that you would have ordinarily overlooked because there's too many items to search through, not enough information about each item, or you don't have enough access to the content. In the case of apps, it's clearly a problem of size as well as knowing enough about each app to understand if it's something you want. Access to apps is actually quite good, so that's not a large barrier, though unlike music which can be enjoyed passively, apps require our complete attention which means that access alone may not always be enough. So how do you overcome these obstacles?
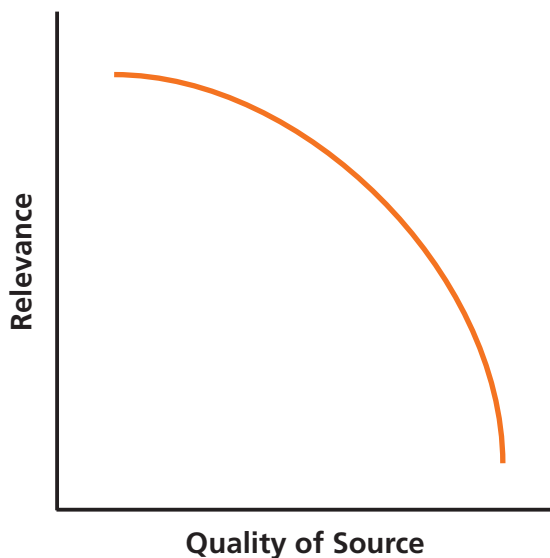
The answer lies in understanding the elements that determine whether we're interested enough in an app to download it, as well as the relationship between those elements. The two most important elements are the source and relevancy, and these have a very complementary relationship. Recognizing that an app is relevant means that it fulfills a need or a desire that you have. If you want to tweak your photos on the go, you'd likely download an app like Photoshop Express because it's highly relevant to you. The source refers to how you found out about the app, and how trustworthy that reference is perceived by you. You'd probably download an app built by your college roommate because you want to support him. The interplay of the source and relevancy is behind every app that you download. And they have an inversely proportional relationship when it comes to predicting whether you'll download a specific app.

Relevance

Quality of Source

Figure 3: The line represents conditions for which "downloadability" is maximized.

From Figure 3, as the intrinsic relevance (to you) of a specific app increases, there is less need for a trusted source to be involved in order to get you to download the app. If you need an app that can do vector calculus, and you find one that fits your needs exactly, the reference which led you to that app (be it a friend, website, or Google search) matters quite little. Conversely, if your best friend releases an app, there's a good chance that you'll download it just on that merit alone, regardless of whether it's actually an app that is relevant to your needs. These are the extremes of the spectrum, and most of the magic is found where these forces coexist: imagine a very specific app that you've been looking for which just happens to have been made by your best friend; the chance you'll download that app is virtually 100%. Basing recommendations solely on trusted sources or trying to find the one app that's exactly what you're looking for are needle-in-a-haystack problems that are destined to fail 99.5% of the time. The key is to optimize both ends of the spectrum as much as possible to increase the chances that you'll find something that appeals on more than one metric.

"Basing recommendations solely on trusted sources or trying to find the one app that's exactly what you're looking for are needle-in-a-haystack problems…"

Maximizing recommendations that come from a trusted source is actually rather easy, since the desired sources can be identified by the user. Maximizing relevance is tougher, and requires mathematical algorithms that can relate various applications across an aggregate of users to be able to suggest apps that are likely to be relevant based on other factors such as: previously-purchased apps, previously viewed apps, or apps the user explicitly dislikes. Many companies do this today: Apple's Genius feature and Amazon's "Customers who viewed this also viewed…" feature both go down this same road. But as becomes clear from casual use, this isn't enough to really deliver a steady stream of content that is highly relevant. The results are very hit-or-miss, and these methods are susceptible to quirks like diaper recommendations when you were just looking to buy a baby shower gift, etc. But if this ability to relate apps can be paired with known trusted sources, you can start making real progress towards understanding what people are likely to want.

Information

But even so, there are other factors that weigh in on this process. One such factor is information. The amount of information available about a given app will help inform its relevance or help to credit or de-credit a particular source. Information alone is not enough to generate downloads but it does provide value. If a user sees a single icon and a name representing an app, they're less likely to know if the source which referred them is credible, or whether the app is objectively relevant to them. This is a similar dynamic as for music; a song title alone isn't enough to know if it's what you want. Adding a description, price, author, and reviews will help the user in deciding the value of the app.

Access to Source

Another important factor is easy access to the source. There are many places on the Internet where someone can go to talk about their favorite this or that, from songs to burgers to car mechanics. Sites like Last.FM, Yelp, and Angie's List understand the value in leveraging your friends who are knowledgeable in different areas, but they require you to belong to their social network in order to access their data. While sign-up is usually free, it's the effort and the lack of integration that prevents real mainstream traction. Yelp has a lot of users, but only about 1% of them actually write reviews.[2] So this information ends up trapped or perhaps shared with a relatively small subset of people. While your music friend might be all over Pandora, Last.FM, iLike, and SoundCloud, that doesn't mean you should have to be on all those networks to take advantage of his expertise. The key to leveraging access for discovery is to provide information from an array of networks that people are *already* using, and mining those feeds for insights; there is no such thing as a

All possible apps from all App Stores

Apps appearing in a user's twitter feed

Sponsored Apps

Sweet Spot

Apps appearing in a user's facebook feed

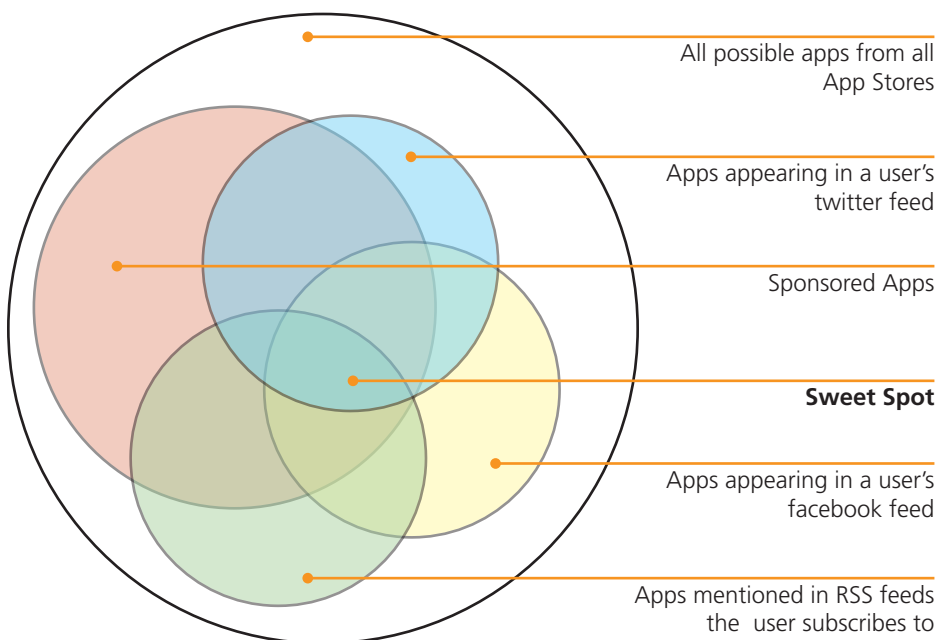Apps mentioned in RSS feeds the user subscribes to

Figure 4: This diagram illustrates the importance of integrating data from multiple feeds. The proportions are not to scale, but recognizing overlaps among sources is key to finding relevance.

"small enough" barrier to entry. If there is a niche for which no network currently exists, then creating a new network *is* likely to be successful, though these are increasingly difficult to find.

Price

A third factor to be concerned with is price. It might seem logical that people buy based mostly on price, but price is seldom the most important factor in making purchase decisions.[3] Exceptions are for perceived commodity items, where price is the only meaningful differentiator. This means that it's certainly possible that an app could be relevant and come from a reputable source and still be spurned because it's too expensive, however this isn't likely to happen if the app is priced appropriately compared to similar apps.

Likes, Dislikes, Activity   **Existing Sources**

Algorithm

+

**Contextual
Information**

↓

**Inviting, Exploratory
User Interface**

Figure 5: Recipe for solving the discovery problem.

PUTTING IT ALL TOGETHER

Synthesizing all these elements into a coherent strategy allows us to create a robust infrastructure for collecting information and delivering quality recommendations which allow people to discover relevant new applications. The first step is collecting information about the user's app tendencies to help get an idea of who they are. A listing of installed applications on the user's device is a good place to start, but offers relatively limited information. We can draw rudimentary conclusions about other likely relevant apps (indeed many recommendation engines do just that) but we can't do a good job predicting future downloads from new or unpopular apps—ones that won't show up on "New" or "Popular" lists—which is precisely the problem we're trying to solve.

The second step is to augment recommendations with sources that the user has already established as being important and with which they have an existing relationship. Some app recommendation engines try to create their own social network or allow users to share apps across existing networks. While this can be useful, the friction is in the wrong

place. In those models, a user gets value by receiving recommended apps from friends who also use the same service and share through the appropriate channels. The incentive to share your own purchases is to encourage others to share with you, but there's no guarantee that they will, and it's difficult to control. In other words, value is derived only when many others are sharing. Reversing this relationship aims to take advantage of existing streams of information—be they Twitter or Facebook timelines, blogs, or existing RSS feeds. This allows a user to subscribe to feeds they find interesting and use the data from those feeds to help inform the recommendation algorithm. These (public) feeds can also be used to create a better picture of the user, so even if a given feed doesn't yield many relevant apps, it can still help build the profile to seed future recommendations. And unlike the previous model, a user gets value for himself without relying on participation from others.

"Five hundred items in a list feels like a lot of work to go through, but presenting a portion of that same list in a nonlinear landscape becomes an intriguing exploration."

The third piece is to begin augmenting these rudimentary recommendations by providing an appealing landscape through which users can interact with the recommended apps. Typically, recommendations are displayed in lists in alphabetical order, sorted by recency, or even by relevance. All three of these approaches have important drawbacks. Alphabetical ordering is completely arbitrary, leaving the user with no choice but to scroll through the entire list. Relevancy has no correlation with the alphabet, so the list needs to be rather short to be of any use at all. Sorting by recency is moderately more helpful, since you can find new titles which haven't garnered much attention, or older gems that are languishing. However, this obscures the vast majority of titles which get modest but not generous attention and which are neither new nor have been available for a long time. Again, long lists become prohibitive and aren't well suited to the issue of relevance. Finally, sorting by perceived relevance is also flawed. The first few recommended apps might be great matches, but as the user scrolls down, the results will explicitly become less and less relevant, which encourages disengagement early on. It also provides a pretty bad experience at the outset when the system is still trying to learn, which will also discourage use. Arrangement aside, lists in and of themselves

aren't well suited to discovery. By definition lists are finite. Long lists make traversal more of a chore because they expose the entire set, so it feels like an almost impossible number of apps, which inevitably hinders apps at the bottom of the list. Short lists only display a partial list of what's recommended, arbitrarily leaving off some relevant content.

So rather than present results in a list, a nonlinear layout that encourages exploration and obscures relevancy ranking is a better way to entice users. Not only are you presenting them with apps they will hopefully like, but you're making the process of finding those apps an engaging one. It allows you to present a large number of applications without overwhelming the user. Five hundred items in a list feels like a lot of work to go through, but presenting a portion of that same list in a nonlinear landscape becomes an intriguing exploration. So the experience becomes truly one of discovery—the user feels as if *she* has chosen the apps—rather than a list which feels like the apps were chosen *for* her. Filter Squad's "Discovr" apps along with Thesixtyone's "Aweditorium" and the "HBO Go" app are excellent examples of this line of thinking.

CONCLUSION

Solving the problem of discovery lies in presenting a variety of choices culled from trustworthy sources which maximize relevance. Sources can be vetted by utilizing information streams from existing networks rather than requiring users to join Yet Another Social Network. Relevance can be attained by using an algorithm that combines links from trusted sources with existing profile information about likes, dislikes, and past behavior. Providing as much information as is sensible with each choice (description, reviews, price, etc) will help determine whether that choice is actually relevant to the user (though more information doesn't create relevance, it only identifies it). Similarly, additional information for each choice will help validate the trustworthiness of the source. These three elements work in concert to provide the best possible chance for people to discover new things. The goal should be to maximize relevance by incorporating trustworthy sources with information we already know about the likes and dislikes of the person. Doing this in a well-designed interface that invites exploration will complete the discovery experience and lead users to finding things they didn't know they wanted.

REFERENCES

1   **Data from: 148apps.biz**
http://148apps.biz/app-store-metrics/?mpage=appprice

2   **Holloway, Daniel (2011, June 6).** Yelp and the "1/9/90 Rule".
Retrieved from:
http://officialblog.yelp.com/2011/06/yelp-and-the-1990-rule.html

3   **Zeithaml, V. A. (1988).** Consumer perceptions of price, quality,
and value: a means-end model and synthesis of evidence. *The Journal of
Marketing Vol. 52, No. 3 (Jul., 1988), 2-22.*