A Measurement Study of Google Play

Nicolas Viennot
Edward Garcia
Jason Nieh

Columbia University
Android is increasingly popular
Android Dominates the Market

World-Wide Smartphone Sales (%)

- Android
- iOS
- Symbian
- RIM
- Bada
- Windows Phone
- Windows Mobile
- Other
Google Play

Google Play APP INSTALLS

- 1 BILLION
- 3 BILLION
- 10 BILLION
- 15 BILLION
- 25 BILLION APP DOWNLOADS

2008 2009 2010 2011 2012
Uploading Content to Google Play is Easy

- Very low barrier to entry:
  - $25 developer account
  - Upload as many apps as you want
  - Once uploaded, app is immediately available to a huge user base
  - No review process
Who Knows What is Really Uploaded?

• Very easy to upload anything, bad or good

• Once installed, apps have access to users’ private life, permissions checks are ineffective

• Despite Google Play popularity, and the risks associated with downloading apps, very little is known on an aggregate level.
Our Study of Google Play

- First large scale measurement of Google Play
- We built PlayDrone to answer many questions
Questions

• How does Google Play content evolve over time?
  Quickly

• How many apps are clones of other apps?
  25%

• How do ratings correlate to popularity?
  Not necessarily as you would expect

• How does native experience correlate with popularity?
  Strongly

• Do developers protect their secrets?
  No

• How many apps have their code obfuscated?
  15%

• Many more in the paper
Questions

- **How does Google Play content evolve over time?**
  Quickly

- How many apps are clones of other apps?
  25%

- **How do ratings correlate to popularity?**
  Not necessarily as you would expect

- How does native experience correlate with popularity?
  Strongly

- **Do developers protect their secrets?**
  No

- How many apps have their code obfuscated?
  15%

- Many more in the paper
PlayDrone
Google Play Crawler

• Fast
  • Can crawl Google Play on a daily basis
  • Easily scales horizontally
• Simple - 2000 lines of Ruby
• Versatile
  • Extensible analysis framework and search engine
  • Decompilation and source code analysis
  • Tracks application changes over time
How does PlayDrone works?

• Interface with the Google Play API at scale
• Acquire content (apps metadata + APK)
• Process APKs
• Index all the results
Architecture

Google Play

PlayDrone
2k LOC in Ruby
Architecture

Google Play

Jobs (Sidekiq)

PlayDrone
2k LOC in Ruby
Architecture

Google Play

Jobs (Sidekiq)

Bookkeeping (Redis)

PlayDrone
2k LOC in Ruby
Architecture

- Google Play
- Jobs (Sidekiq)
  - Bookkeeping (Redis)
  - Repositories (Git)

PlayDrone
2k LOC in Ruby
Architecture

Google Play

Jobs
(Sidekiq)

Bookkeeping
(Redis)

Repositories
(Git)

Analytics
(Elasticsearch)

PlayDrone
2k LOC in Ruby
Architecture

Google Play

Jobs (Sidekiq)

- Bookkeeping (Redis)
- Repositories (Git)
- Analytics (Elasticsearch)

PlayDrone
2k LOC in Ruby

Frontend (Rails)
Gmail 4.5-694836

Git Repository

```bash
git clone git://node02.googleplaywith.me/com/google/android/gm.git com.google.android.gm
```

What's New

- New inbox: if enabled, your mail will be grouped into categories so you can see what's new at a glance and decide which emails to read when
- Streamlined user interface including swipe-down to refresh and sliding drawer with labels and account switcher (phones and 7" tablets)
- Sender images shown alongside messages - tap on images to select multiple emails
- Improved readability for many emails
- Emptying trash now supported
- Bug fixes and speed-ups

Android 4.0 (Ice Cream Sandwich) and up.

What's Really New

```bash
commit d966374c190f50c46bc05388049c78360b3c6c46
Author: Google Inc. <crawler@googleplaywith.me>
Date: Wed Jun 5 00:00:00 2013 -0400

[DecompileApk] Processed v4.5-694836
Version Code: 974
752 files changed, 32847 insertions(+), 23213 deletions(-)

| src/com/android/mail/AccountSpinnerAdapter.java | 389 --------- |
| src/com/android/mail/ContactInfo.java           | 17 +          |
| src/com/android/mail/EmailAddress.java          | 98 +          |
| src/com/android/mail/MailIntentService.java     | 39 -          |
| src/com/android/mail/MailLogService.java        | 126 +++        |
```
Deployment

• 10 servers: quad-cores at 3.8Ghz, 32GB of RAM, and 2x2TB drives

• Two crawls: May/June 2013 and Nov 2013
Crawl Day in May 2013

Time

Throughput (req/s)

Details

Search
How does Google Play content evolve over time?

Question #1
## Number of Applications
### 5-Month Evolution

<table>
<thead>
<tr>
<th></th>
<th>June 22, 2013</th>
<th>Nov. 30, 2013</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Free Apps</strong></td>
<td>691,517</td>
<td>884,217 (+28%)</td>
</tr>
<tr>
<td><strong>Paid Apps</strong></td>
<td>192,703</td>
<td>223,259 (+14%)</td>
</tr>
<tr>
<td><strong>All Apps</strong></td>
<td>887,220</td>
<td>1,107,476 (+25%)</td>
</tr>
</tbody>
</table>
Apps Breakdown with Download Counts

Number of Apps

Download Counts

- <500
- 500-1k
- 1k-5k
- 5k-10k
- 10k-50k
- 50k-100k
- 100k-500k
- 500k-1M
- 1M-5M
- 5M-10M
- 10M-50M
- >50M

Free
Paid

- 305376
- 172044
- 109477
- 41514
- 72969
- 19244
- 21229
- 3822
- 3594
- 524
- 392
- 55
Question #2

How do ratings correlate to popularity?
Average Average Rating vs Downloads

<table>
<thead>
<tr>
<th>Rating</th>
<th>&lt;500</th>
<th>500-1k</th>
<th>1k-5k</th>
<th>5k-10k</th>
<th>10k-50k</th>
<th>50k-100k</th>
<th>100k-500k</th>
<th>500k-1M</th>
<th>1M-5M</th>
<th>5M-10M</th>
<th>10M-50M</th>
<th>&gt;50M</th>
</tr>
</thead>
<tbody>
<tr>
<td>Free Apps</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
</tr>
<tr>
<td>Paid Apps</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
<td>Blue</td>
<td>Green</td>
</tr>
</tbody>
</table>

Download Counts:
- <500
- 500-1k
- 1k-5k
- 5k-10k
- 10k-50k
- 50k-100k
- 100k-500k
- 500k-1M
- 1M-5M
- 5M-10M
- 10M-50M
- >50M
Maximum Average Rating vs Downloads

- **Free Apps**
- **Paid Apps**

<table>
<thead>
<tr>
<th>Rating</th>
<th>&lt;500</th>
<th>500-1k</th>
<th>1k-5k</th>
<th>5k-10k</th>
<th>10k-50k</th>
<th>50k-100k</th>
<th>100k-500k</th>
<th>500k-1M</th>
<th>1M-5M</th>
<th>5M-10M</th>
<th>10M-50M</th>
<th>&gt;50M</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
<td>5</td>
</tr>
</tbody>
</table>

Download Counts:
- <500
- 500-1k
- 1k-5k
- 5k-10k
- 10k-50k
- 50k-100k
- 100k-500k
- 500k-1M
- 1M-5M
- 5M-10M
- 10M-50M
- >50M
Minimum Average Rating vs Downloads

Rating vs Download Counts

- <500
- 500-1k
- 1k-5k
- 5k-10k
- 10k-50k
- 50k-100k
- 100k-500k
- 500k-1M
- 1M-5M
- 5M-10M
- 10M-50M
- >50M

Free Apps vs Paid Apps
## Top5 Best Rated Apps with >1M Downloads

<table>
<thead>
<tr>
<th>App Name</th>
<th>Downloads</th>
<th>#Ratings</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>TvQuran</td>
<td>1M-5M</td>
<td>13,675</td>
<td>4.93</td>
</tr>
<tr>
<td>Билеты ПДД 2013 РФ</td>
<td>1M-5M</td>
<td>15,738</td>
<td>4.92</td>
</tr>
<tr>
<td>Holy Quran Maher Moagely</td>
<td>1M-5M</td>
<td>6,341</td>
<td>4.91</td>
</tr>
<tr>
<td>Slots Deluxe - Slot Machines</td>
<td>1M-5M</td>
<td>108,431</td>
<td>4.90</td>
</tr>
<tr>
<td>ﺍﺩﻌﺍﺓ ﻭﺃﺫﻜﺍﺭ ﺡﺹﺫึก ﺍﻟﻤﺴﻠﻡ</td>
<td>1M-5M</td>
<td>19,567</td>
<td>4.89</td>
</tr>
</tbody>
</table>
### Top5 Worse Rated Apps with >1M Downloads

<table>
<thead>
<tr>
<th>App</th>
<th>Downloads</th>
<th>#Ratings</th>
<th>Rating</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wet Lesbian</td>
<td>1M-5M</td>
<td>2,865</td>
<td>2.23</td>
</tr>
<tr>
<td>Ameba</td>
<td>1M-5M</td>
<td>35,933</td>
<td>2.21</td>
</tr>
<tr>
<td>HRS App</td>
<td>1M-5M</td>
<td>5,778</td>
<td>1.99</td>
</tr>
<tr>
<td>T-Mobile More For Me</td>
<td>5M-10M</td>
<td>1,763</td>
<td>1.84</td>
</tr>
<tr>
<td>DroidScale</td>
<td>1M-5M</td>
<td>5,450</td>
<td>1.67</td>
</tr>
</tbody>
</table>
DroidScale
DroidMob Technologies - December 21, 2009
Tools

Install  Add to Wishlist

This app is compatible with all of your devices.

⭐⭐⭐⭐⭐ (6,022)
8.1  +1,155  Recommend this on Google

Description

OVER 1 MILLION DOWNLOADS!!!! AD FREE PRO VERSION ALSO AVAILABLE!
This dynamic app turns your phone into a digital scale that you can take with you wherever you go!
Weigh normal household items with your friends, and see who can get the closest guess! Weighs in grams or ounces!
Follow the in-app instructions for use.
Note: DroidScale does not guarantee exact measurements.
DroidScale Code Sample
```java
public void onCreate(Bundle paramBundle) {
    super.onCreate(paramBundle);
    rgen = new Random();
    curr = 0.0D;
    conversion = 28.34952310;
    new AlertDialog.Builder(this).setMessage("Warning: DroidScale is intended for lightweight objects and will not accurately display heavy weights. Do not place too much weight on the scale for the safety of your screen! \nPlace a finger on the scale and place the object you desire to weigh on top of your finger.").setPositiveButton("OK", null).setTitle("Instructions").setIcon(0).show();
    
    TextView localTextView = (TextView) findViewById(R.id.units);
    localView.setOnTouchListener(new View.OnTouchListener())
    {
        public boolean onTouch(View paramAnonymousView, MotionEvent paramAnonymousMotionEvent)
        {
            if (paramAnonymousMotionEvent.getAction() == 1)
            {
                touchDown = false;
                curr = 0.0D;
                masterTextView.setText("0.0000");
                return true;
            }
            for (curr = rgen.nextDouble(); curr = rgen.nextDouble())
            { //_code
                if (curr >= 0.2D)
                {
                    if (inGrams)
                    {
                        DroidScale localDroidScale = DroidScale.this;
                        curr *= conversion;
                    }
                    String str = Double.toString(curr).substring(0, 6);
                    masterTextView.setText(str);
                    touchDown = true;
                    fluctuateWeights();
                    return true;
                }
            }
```
Question #3

Do developers protect their secrets?
Auth Tokens

- Used to authenticate a 3rd party app (e.g. AirBnB) to a service provider (e.g. Facebook)

- With a root level Amazon AWS token, you may access and launch EC2 servers.

- With a Facebook token, you may access users’ private information, write on their walls.
public class OAuthUtil {

    public static final String FACEBOOK_ACCESS_URL = "https://graph.facebook.com/oauth/access_token";
    public static final String FACEBOOK_AUTHORIZE_URL = "https://graph.facebook.com/oauth/authorize?scope=offline_access,read_stream,profile,email";
    public static final String FACEBOOK_CONSUMER_KEY = "63988480cab54990b906ee33ea83ef";
    public static final String FACEBOOK_CONSUMER_SECRET = "7b00de68c14c17b9e129a6ec075a4298";
    public static final String FACEBOOK_REDIRECT_URI = "http://www.taptu.com/streams/oauth";

    public static final String GOOGLE_READER_ACCESS_URL = "https://accounts.google.com/o/oauth2/token";
    public static final String GOOGLE_READER_CONSUMER_KEY = "67565137267-4l52e2pdnomd3gruzj9v97919h1cbvdro.apps.googleusercontent.com";
    public static final String GOOGLE_READER_CONSUMER_SECRET = "GnTTzQ99jaXOzw1RzeZuTeq9W";
    public static final String GOOGLE_READER_REDIRECT_URI = "https://www.taptu.com/oauth/google";

    public static final String GPLUS_ACCESS_URL = "https://accounts.google.com/o/oauth2/token";
    public static final String GPLUS_CONSUMER_KEY = "043467224689-t7110b2hl7f53e7jff26a957jontt.apps.googleusercontent.com";
    public static final String GPLUS_CONSUMER_SECRET = "H6IE-XzM2P0D-00xNQvKLeY";
    public static final String GPLUS_REDIRECT_URI = "http://localhost/";

    public static final String LINKEDIN_ACCESS_URL = "https://api.linkedin.com/uas/oauth/requestToken";
    public static final String LINKEDIN_AUTHORIZE_URL = "https://api.linkedin.com/uas/oauth/authorize";
    public static final String LINKEDIN_CONSUMER_KEY = "loCSO6hNuf1flgQ4-10PD0JSH329y7L5yOeEADHpt3VilhmLa_Ylz_yG38";
    public static final String LINKEDIN_CONSUMER_SECRET = "yu26Y8K-67b2EB6QYpinPq9q7soCNhV66UZgwjWj52C32xZU_vz3x9RvUyIq4X";
    public static final String LINKEDIN_REQUEST_URL = "https://api.linkedin.com/uas/oauth/requestToken";

    public static final String NETEASEWEIBO_ACCESS_URL = "";
    public static final String NETEASEWEIBO_AUTHORIZE_URL = "";
    public static final String NETEASEWEIBO_CONSUMER_KEY = "rmQe1gtqznKFA6l";
    public static final String NETEASEWEIBO_CONSUMER_SECRET = "foeTHs2tGwEeW7qYofDT9ohbV8hf";
    public static final String NETEASEWEIBO_REDIRECT_URI = "http://www.taptu.com/";
    public static final String REDREN_ACCESS_URL = "https://graph.renren.com/oauth/token";
    public static final String REDREN_AUTHORIZE_URL = "https://graph.renren.com/oauth?scope=read_user_feed read_user_status p";
    public static final String REDREN_CONSUMER_KEY = "aa83d67b74141ab8da076985d35b99c";
    public static final String REDREN_CONSUMER_SECRET = "269d048344359ea4e4atf344f7c94a";
    public static final String REDREN_REDIRECT_URI = "http://www.taptu.com/";

    public static final String SINAWEIBO_ACCESS_URL = "https://api.weibo.com/oauth2/access_token";
    public static final String SINAWEIBO_AUTHORIZE_URL = "https://api.weibo.com/oauth2/authorize?display=mobile&scope=statuses_to_me_read";
    public static final String SINAWEIBO_CONSUMER_KEY = "384815873";
    public static final String SINAWEIBO_CONSUMER_SECRET = "2ac91e3b8d77b4995eaa65a53d92a654";
    public static final String SINAWEIBO_REDIRECT_URI = "http://www.taptu.com/";

    public static final String TENCENTWEIBO_ACCESS_URL = "https://open.t.qq.com/cgi-bin/oauth2/access_token";
    public static final String TENCENTWEIBO_AUTHORIZE_URL = "https://open.t.qq.com/cgi-bin/oauth2/authorize";
    public static final String TENCENTWEIBO_CONSUMER_KEY = "98125326";
    public static final String TENCENTWEIBO_CONSUMER_SECRET = "b734d4848a937bb30db3db255s5d";
    public static final String TENCENTWEIBO_REDIRECT_URI = "http://www.taptu.com/";

    public static final String TWITTER_ACCESS_URL = "http://api.twitter.com/oauth/access_token";
    public static final String TWITTER_AUTHORIZE_URL = "http://api.twitter.com/oauth/authorize";
    public static final String TWITTER_CONSUMER_KEY = "m47r1ne70ji7uq6j5j7q";
    public static final String TWITTER_CONSUMER_SECRET = "HxXSBPjD80AAXn038HmT5xuNNxzzXebVxw4d0q9ge8";
    public static final String TWITTER_REQUEST_URL = "http://api.twitter.com/oauth/request_token";
# Regular Expressions

<table>
<thead>
<tr>
<th></th>
<th>Client ID</th>
<th>Secret Key</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon AWS</td>
<td>AKIA[0-9A-Z]{16}</td>
<td>[0-9a-zA-Z/+]{40}</td>
</tr>
<tr>
<td>Bitly</td>
<td>[0-9a-zA-Z_]{5,31}</td>
<td>R_[0-9a-f]{32}</td>
</tr>
<tr>
<td>Facebook</td>
<td>[0-9]{13,17}</td>
<td>[0-9a-f]{32}</td>
</tr>
<tr>
<td>Flickr</td>
<td>[0-9a-f]{32}</td>
<td>[0-9a-f]{16}</td>
</tr>
<tr>
<td>Foursquare</td>
<td>[0-9A-Z]{48}</td>
<td>[0-9A-Z]{48}</td>
</tr>
<tr>
<td>Google</td>
<td>[0-9a-zA-Z._-]*/?_app</td>
<td>[0-9a-zA-Z_]{24}</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>[0-9a-z]{12}</td>
<td>[0-9a-zA-Z]{16}</td>
</tr>
<tr>
<td>Twitter</td>
<td>[0-9a-zA-Z]{18,25}</td>
<td>[0-9a-zA-Z]{35,44}</td>
</tr>
</tbody>
</table>

Note: Additional criteria apply to reduce false positives
## Auth Tokens

<table>
<thead>
<tr>
<th></th>
<th>Total Candidates</th>
<th>Unique Candidates</th>
<th>Unique % Valid</th>
</tr>
</thead>
<tbody>
<tr>
<td>Amazon</td>
<td>1,241</td>
<td>308</td>
<td>93.5%</td>
</tr>
<tr>
<td>Facebook</td>
<td>1,477</td>
<td>460</td>
<td>71.7%</td>
</tr>
<tr>
<td>Twitter</td>
<td>28,235</td>
<td>6,228</td>
<td>95.2%</td>
</tr>
<tr>
<td>Bitly</td>
<td>3,132</td>
<td>616</td>
<td>88.8%</td>
</tr>
<tr>
<td>Flickr</td>
<td>159</td>
<td>89</td>
<td>100%</td>
</tr>
<tr>
<td>Foursquare</td>
<td>326</td>
<td>177</td>
<td>97.7%</td>
</tr>
<tr>
<td>Google</td>
<td>414</td>
<td>225</td>
<td>96.0%</td>
</tr>
<tr>
<td>LinkedIn</td>
<td>1,434</td>
<td>181</td>
<td>97.2%</td>
</tr>
<tr>
<td>Titanium</td>
<td>1,914</td>
<td>1,783</td>
<td>99.8%</td>
</tr>
</tbody>
</table>

Tokens found June 2013, validated Nov 2013
### Facebook and Twitter

<table>
<thead>
<tr>
<th></th>
<th>Facebook</th>
<th>Twitter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tokens Found</td>
<td>460</td>
<td>6,228</td>
</tr>
<tr>
<td>Corresponding Library Found</td>
<td>92,495</td>
<td>6,990</td>
</tr>
</tbody>
</table>

Facebook relies on their SDK to authenticate 3rd party applications through their Facebook app with Android Intents.
Create a place to store our constants Right-click the "src", "com.example.twittertutorial" package and select "Class". In the "Name" field, enter "Constants". In this file, we'll add constants that we'll reference in various aspects of our demo.

```java
package com.example.twittertutorial;

public class Constants {
    public final static String CONSUMER_KEY = "";
    public final static String CONSUMER_SECRET = "";
    public final static String OAUTH_TOKEN = "oauth_token";
    public final static String OAUTH_VERIFIER = "oauth_verifier";
    public final static String OAUTH_DENIED = "denied";
    public final static String OAUTH_CALLBACK = "twittertutorial://oauth";
}
```
Notified all service providers

- Service providers have since disabled all tokens that were security risks
- Various approaches for resolving security issue
  - Amazon - notify and work with customers directly
  - Facebook - immediately revoke access
Making Google Play Safer

- Notified and worked with Google
- Provided Google with PlayDrone token finder mechanism
- Google has integrated mechanism into Bouncer to automatically scan for tokens and notify developers
Hello,

If you embed Amazon Web Services (AWS) credentials in your applications, those credentials become publicly available. This exposure of your credentials could lead to unauthorized access to your AWS account, which may include associated excessive charges, and potentially unauthorized access to your data and your users’ data.

Please see this [AWS post](http://aws.amazon.com) on what to do if your AWS Access Key is exposed.

This post will give you guidance on what to do next, including:

1. Rotate your credentials, in order to invalidate your publicly exposed credentials.
2. Review your AWS account for unauthorized access.
3. Review Amazon’s tutorials for using short term credentials and/or identity providers, such as Amazon, Google, or Facebook.

If you have questions or require assistance, please contact [Amazon support](http://aws.amazon.com/support).

Regards,

Google Play Team
Conclusion

- First large scale study of Google Play
- PlayDrone provides answers to many questions
- Made Google Play safer
Source Code
http://github.com/nviennnot/playdrone

Contact
twitter: @nviennnot
email: nicolas@viennnot.com

Questions?
Backup Slides
How many apps obfuscate their sources?
Obfuscation Rate over Time

% of Applications

April 27, 2013

June 22, 2013

All Market
New Apps
Updated Apps
How many apps are clones of other apps?
Detecting Clones

• Existing approaches do complicated things with code analysis

• We take a simple approach:

• Similar apps have similar assets (images, sounds)

• Hash them to build app signatures: 45M signatures

• Reject common signatures (seen in >300 apps)

• 5% of false positives (sample of 400 apps)
Clone Study Result

At least 25% of apps are clones of other apps
How does native experience correlate with popularity?
Developing an App

- App generator (a few clicks)
- Cross platform frameworks (html/javascript)
- Use the regular Android SDK (java)
- With native libraries (compiled down to ARM)
## App Generators

<table>
<thead>
<tr>
<th>App Generators</th>
<th>Non-popular Apps (&lt;50k downloads)</th>
<th>Popular Apps (&gt;50k downloads)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Business Apps</td>
<td>10,011 (1.59%)</td>
<td>3 (0.01%)</td>
</tr>
<tr>
<td>App Inventor</td>
<td>9,560 (1.52%)</td>
<td>152 (0.29%)</td>
</tr>
<tr>
<td>Andromo</td>
<td>6,294 (1.00%)</td>
<td>156 (0.30%)</td>
</tr>
<tr>
<td>iBuildApp</td>
<td>4,149 (0.66%)</td>
<td>25 (0.05%)</td>
</tr>
<tr>
<td>Mobile by Conduit</td>
<td>3,989 (0.63%)</td>
<td>21 (0.04%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>34,003 (5.39%)</td>
<td>357 (0.68%)</td>
</tr>
</tbody>
</table>
## Cross-platform Frameworks

<table>
<thead>
<tr>
<th>Frameworks</th>
<th>Non-popular Apps (&lt;50k downloads)</th>
<th>Popular Apps (&gt;50k downloads)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PhoneGap</td>
<td>36,915 (5.85%)</td>
<td>606 (1.16%)</td>
</tr>
<tr>
<td>Adobe Air</td>
<td>12,761 (2.02%)</td>
<td>619 (1.18%)</td>
</tr>
<tr>
<td>Titanium</td>
<td>8,316 (1.32%)</td>
<td>138 (0.26%)</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>57,991 (9.20%)</strong></td>
<td><strong>1,363 (2.60%)</strong></td>
</tr>
</tbody>
</table>
What’s up with the removal of apps?
Removed Apps vs Time

Number of Apps

Other Categories

Personalization

April 27, 2013

June 22, 2013
Top Occurring Words in the Personalization Category

<table>
<thead>
<tr>
<th>Word</th>
<th>Personalization Category</th>
<th>Rest of the Market</th>
</tr>
</thead>
<tbody>
<tr>
<td>wallpaper</td>
<td>69%</td>
<td>4%</td>
</tr>
<tr>
<td>please</td>
<td>39%</td>
<td>12%</td>
</tr>
<tr>
<td>like</td>
<td>29%</td>
<td>12%</td>
</tr>
</tbody>
</table>
Who leads the Ads market?
Advertising Platforms
Market Share over Time
(Among Apps with Ads Libs)

% of Applications

April 27, 2013 to June 22, 2013
How to discover apps?
Discover Applications

• There is no way to get an exhaustive list of apps

• Results are capped to 500 apps (cannot “click on next” indefinitely when browsing the market)

• Dictionary based exploration: do a search for each of the 1,000,000 words from 10 languages.

  • **search** API endpoint

• We also look at the *related apps* of each app
Crawl Day in May 2013

Throughput (req/s)

Latency (s)

Time

Details | Search | Purchase

04:00 | 06:00 | 08:00 | 10:00 | 12:00 | 14:00 | 16:00 | 18:00 | 20:00
How to search in sources?
Figure 9: PLAYDRONE’s web interface to search decompiled sources showing Amazon Web Service tokens found in 130 ms.
Regular Expressions

tokens:amazon, :access_key_id => '[=\"](AKIA[0-9A-Z]{16})[&"]',
    :secret_access_key => '[=\"]([0-9a-zA-Z+/]{40})[&"]',
    :proximity => 5

tokens:facebook, :app_id => { :matcher => '[=\"]([0-9]{13,17})[&"]', :line_must_have => /facebook/i },
    :app_secret => { :matcher => '[=\"]([0-9a-f]{32})[&"]', :line_must_have => /facebook/i }

tokens:flickr, :api_key => { :matcher => '[=\"]([0-9a-f]{32})[&"]', :line_must_have => /flickr/i },
    :api_secret => { :matcher => '[=\"]([0-9a-f]{16})[&"]', :line_must_have => /flickr/i }

tokens:foursquare, :client_id => '(?:id|ID|token|TOKEN).*?[=\"]+(0-9A-Z){48})[&"]',
    :client_secret => '(?:secret|SECRET).*?[=\"]+(0-9A-Z){48})[&"]',
    :random_threshold => 0.2

tokens:google_oauth, :client_id => '[=\"]([0-9a-zA-Z.-]*\apps\googleusercontent\.com)[&"]',
    :client_secret => '[=\"]([0-9a-zA-Z.-]{24})[&"]'

tokens:linkedin, :api_key => '[=\"]([0-9a-z]{12})[&"]',
    :secret_key => { :matcher => '[=\"]([0-9a-zA-Z]{16})[&"]', :must_have => /[g-zA-Z]/ },
    :random_threshold => 0.6

tokens:twitter, :consumer_key => { :matcher => '[=\"]([0-9a-zA-Z]{18,25})[&"]', :line_must_have => /twitter/i },
    :consumer_secret => { :matcher => '[=\"]([0-9a-zA-Z]{35,44})[&"]', :line_must_have => /twitter/i }

tokens:bitlyv1, :login => { :matcher => '[=\"]([0-9a-zA-Z]{5,31})[&"]', :must-have => /[a-zA-Z]/ },
    :api_key => '[=\"](R_[0-9a-f]{32})[&"]'