ISCX
Information Security
Centre of Excellence

# Exploring Reverse Engineering Symptoms in Android apps

**Hugo Gonzalez, Andi A. Kadir, Natalia Stakhanova, Ali A. Ghorbani**
***Faculty of Computer Science, University of New Brunswick***

EST. 1785 UNB
UNIVERSITY OF NEW BRUNSWICK
Computer Science

## Motivation

- Rise of Android malware.
- What are the ways that adversaries create malware?
  - From scratch?
  - Repackaging other apps?
- Several work on detecting repackaged apps.
  - Expensive computations

## Proposed solution

- Mobile app repackaging is often an indicator for suspicious app.
- Triage to do more work on the suspicios apps.
- Instead of performing static or dynamic analysis, we focus on the layout of the .dex file.
- String Offset Order is an easy extractable attribute that is a signal for repackaging.
- We performed extensive evaluation of String Offset Order metric to assess its capabilities over 90,000 samples.
- AndroidSOO, a lightweight approach for the detection of repackaging (reverse engineering) symptoms on Android apps.
- Large scale evaluation

## Discussion

- AndroidSOO effectively detect apktool and dalvik-obfuscator.
- It missed manual repackaged which we consider proof-of-concept.
- AndroidSOO detect Adobe Air as repackaged symptoms. For sure the app was not created by Android Development Toolkit / Android Studio

## Focus on .dex file layout

| | |
|---|---|
| header | Structural information |
| string_ids | Offset list for strings |
| type_ids | Index list into the string_ids for types |
| proto_ids | Identifiers list of prototypes |
| field_ids | Identifiers list of fields |
| method_ids | Identifiers list of methods |
| class_defs | Structure list of classes |
| data | Code and data |
| link_data | Data in statically linked files |

## Validation Dataset

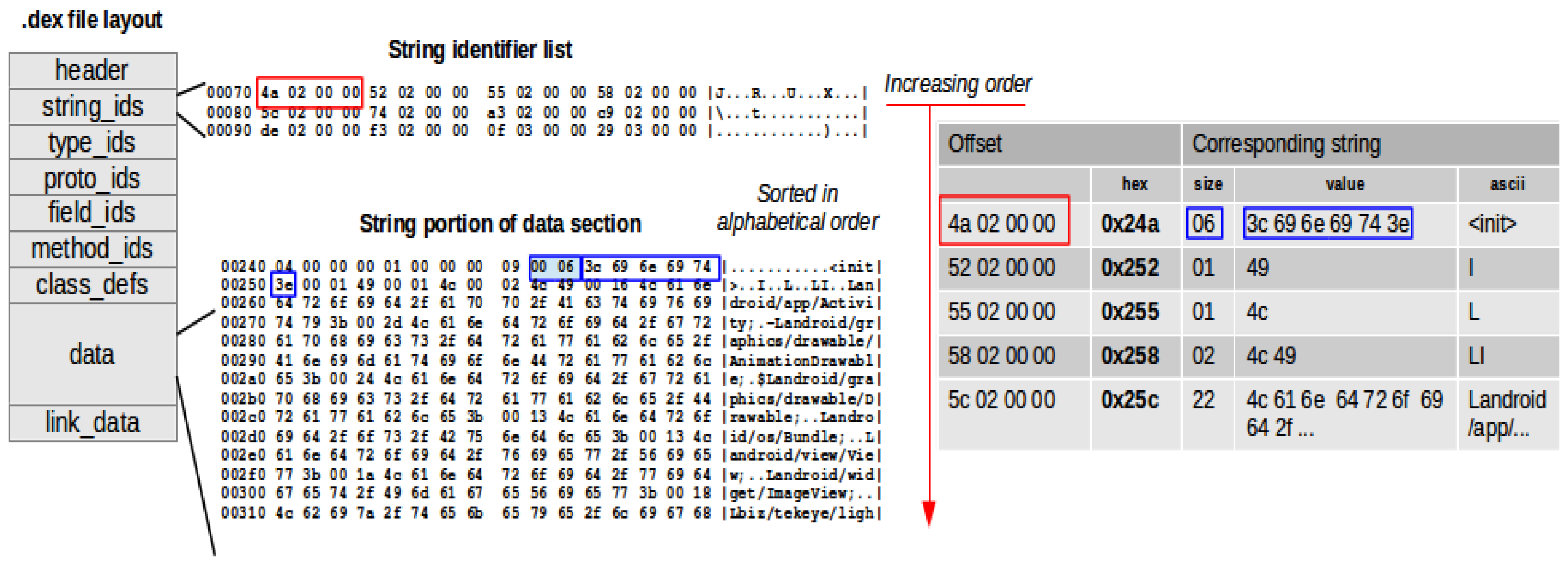
| Origin | # of apps |
|---|---|
| **Not-repackaged apps** | |
| Original apps from individual sources | 48 |
| Obfuscated/optimized with | |
| Proguard | 48 |
| Bangcle | 3 |
| HosedDex2jar | 3 |
| DashO | 3 |
| ApkProtector | 3 |
| Apps enhanced with Mobile Ad library SDKs | 5 |
| Application generators | |
| PhoneGap | 5 |
| AdobeAir | 5 |
| Titanium | 5 |
| Bizness Apps | 1 |
| Andromo | 1 |
| App Inventor | 2 |
| iBuildApp | 2 |
| Como (Mobile by Conduit) | 1 |
| Dot42 | 14 |
| DexGuard apps (GooglePlay) | 5 |
| DexGuard malware apps (VirusTotal) | 2 |
| Official apps from large open-source projects (optimized) | 14 |
| **Repackaged apps** | |
| akpTool | 156 |
| dalvik-obfuscator | 5 |
| manual repackaging | 3 |

| Apps | # of apps | Detected correctly | Missed |
|---|---|---|---|
| Apps without repackaging | 170 | 165 | 5 |
| Repackaged apps | 164 | 161 | 3 |
| Total | 334 | DR =98%, FPR = 2.9% | |

## Conclusions

- Even in the presence of obfuscation, we can detect repackaged apps.
- AndroidSOO does not need training.
- Its scalable.
- Its a complementary approach in more comprehensive analysis for existing apps.

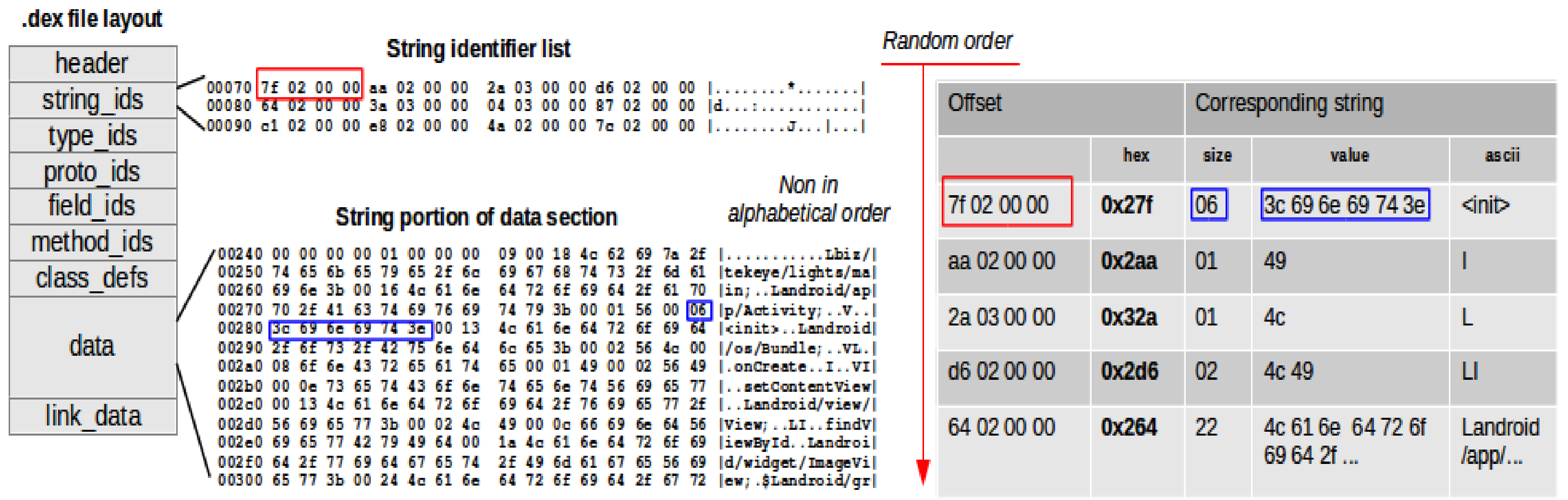
## String Offset Order (SOO)



```
00070 4a 02 00 00 52 02 00 00  55 02 00 00 58 02 00 00 |J...R...U...X...|
00080 5c 02 00 00 74 02 00 00  a3 02 00 00 c9 02 00 00 |\...t...........|
00090 de 02 00 00 f3 02 00 00  0f 03 00 00 29 03 00 00 |............)...|
```

```
00240 04 00 00 00 01 00 00 00  09 00 06 3c 69 6e 69 74 |...........<init|
00250 3e 00 01 49 00 01 4c 00  02 4c 49 00 16 4c 61 6e |>..I..L..LI..Lan|
00260 64 72 6f 69 64 2f 61 70  70 2f 41 63 74 69 76 69 |droid/app/Activi|
00270 74 79 3b 00 2d 4c 61 6e  64 72 6f 69 64 2f 67 72 |ty;.-Landroid/gr|
00280 61 70 68 69 63 73 2f 64  72 61 77 61 62 6c 65 2f |aphics/drawable/|
00290 41 6e 69 6d 61 74 69 6f  6e 44 72 61 77 61 62 6c |AnimationDrawabl|
002a0 65 3b 00 24 4c 61 6e 64  72 6f 69 64 2f 67 72 61 |e;.$Landroid/gra|
002b0 70 68 69 63 73 2f 64 72  61 77 61 62 6c 65 2f 44 |phics/drawable/D|
002c0 72 61 77 61 62 6c 65 3b  00 13 4c 61 6e 64 72 6f |rawable;..Landro|
002d0 69 64 2f 6f 73 2f 42 75  6e 64 6c 65 3b 00 13 4c |id/os/Bundle;..L|
002e0 61 6e 64 72 6f 69 64 2f  76 69 65 77 2f 56 69 65 |android/view/Vie|
002f0 77 3b 00 1a 4c 61 6e 64  72 6f 69 64 2f 77 69 64 |w;..Landroid/wid|
00300 67 65 74 2f 49 6d 61 67  65 56 69 65 77 3b 00 18 |get/ImageView;..|
00310 4c 62 69 7a 2f 74 65 6b  65 79 65 2f 6c 69 67 68 |Lbiz/tekeye/ligh|
```

| Offset | Corresponding string | | | |
|---|---|---|---|---|
| | hex | size | value | ascii |
| 4a 02 00 00 | **0x24a** | 06 | 3c 69 6e 69 74 3e | <init> |
| 52 02 00 00 | **0x252** | 01 | 49 | I |
| 55 02 00 00 | **0x255** | 01 | 4c | L |
| 58 02 00 00 | **0x258** | 02 | 4c 49 | LI |
| 5c 02 00 00 | **0x25c** | 22 | 4c 61 6e 64 72 6f 69 64 2f ... | Landroid /app/... |



```
00070 7f 02 00 00 aa 02 00 00  2a 03 00 00 d6 02 00 00 |........*.......|
00080 64 02 00 00 3a 03 00 00  04 03 00 00 87 02 00 00 |d...:...........|
00090 c1 02 00 00 e8 02 00 00  4a 02 00 00 7c 02 00 00 |........J...|...|
```

```
00240 00 00 00 00 01 00 00 00  09 00 18 4c 62 69 7a 2f |...........Lbiz/|
00250 74 65 6b 65 79 65 2f 6c  69 67 68 74 73 2f 6d 61 |tekeye/lights/ma|
00260 69 6e 3b 00 16 4c 61 6e  64 72 6f 69 64 2f 61 70 |in;..Landroid/ap|
00270 70 2f 41 63 74 69 76 69  74 79 3b 00 01 56 00 06 |p/Activity;..V..|
00280 3c 69 6e 69 74 3e 00 13  4c 61 6e 64 72 6f 69 64 |<init>..Landroid|
00290 2f 6f 73 2f 42 75 6e 64  6c 65 3b 00 02 56 4c 00 |/os/Bundle;..VL.|
002a0 08 6f 6e 43 72 65 61 74  65 00 01 49 00 02 56 49 |.onCreate..I..VI|
002b0 00 0e 73 65 74 43 6f 6e  74 65 6e 74 56 69 65 77 |..setContentView|
002c0 00 13 4c 61 6e 64 72 6f  69 64 2f 76 69 65 77 2f |..Landroid/view/|
002d0 56 69 65 77 3b 00 02 4c  49 00 0c 66 69 6e 64 56 |View;..LI..findV|
002e0 69 65 77 42 79 49 64 00  1a 4c 61 6e 64 72 6f 69 |iewById..Landroi|
002f0 64 2f 77 69 64 67 65 74  2f 49 6d 61 67 65 56 69 |d/widget/ImageVi|
00300 65 77 3b 00 24 4c 61 6e  64 72 6f 69 64 2f 67 72 |ew;.$Landroid/gr|
```

| Offset | Corresponding string | | | |
|---|---|---|---|---|
| | hex | size | value | ascii |
| 7f 02 00 00 | **0x27f** | 06 | 3c 69 6e 69 74 3e | <init> |
| aa 02 00 00 | **0x2aa** | 01 | 49 | I |
| 2a 03 00 00 | **0x32a** | 01 | 4c | L |
| d6 02 00 00 | **0x2d6** | 02 | 4c 49 | LI |
| 64 02 00 00 | **0x264** | 22 | 4c 61 6e 64 72 6f 69 64 2f ... | Landroid /app/... |

## Large scale evaluation

| Dataset | Total | SOO random | SOO intact |
|---|---|---|---|
| Genome Project | 1260 | 48.73% | 51.27% |
| Debrin | 5555 | 22.8% | 76.72% |
| DroidAnalytics | 2140 | 67.20% | 32.80% |
| Googleplay | 5,058 | 2.01% | 97.99% |
| VirusTotal .dex | 28,700 | 35.20% | 64.80% |
| VirusTotal .apk | 53,621 | 16.97% | 83.03% |
| Total | 96,334 | | |

45 Adobe Air
57 repackaged
- 30 adware by VirusTotal

| Dataset | Total (secs) | Unpack (secs) | Average total time per app (ms) |
|---|---|---|---|
| Genome Project | 12.665 | 9.259 | 2 |
| Debrin | 62.530 | 49.644 | 2 |
| DroidAnalytics | 32.405 | 23.120 | 4 |
| Googleplay | 153.412 | 127.050 | 5 |
| VirusTotal .dex | 85.442 | – | 3 |
| VirusTotal .apk | 1994.977 | 1672.251 | 6 |

## Code

***http://github.com/hugo-glez/androsoo***