الگوریتم جست و جوی اول عمق و هدفمند برای تست سیستمی و اصولی برنامه های اندروید

چکیده :

جست وجوی سیستمی برنامه های اندروییدیک توان مند ساز برای تحلیل طیف وسیعی از برنامه ها و فعالیت های آزمایشی است. انجام جست و جو در زمان کار کرد گوشی های واقعی، برای جست و جو کشف همه قابلیت های نرم افزار لازم است. با این حال، جست و جوی نرم افزار های دنیای واقعی بر روی گوشی های واقعی به دلیل عدم قطعیت، عدم جریان کنترل استاندارد، توسعه پذیری و محدودیت های سربار سخت است. با اتکا به کار بران نهایی برای انجام جست و جو ها ما یک مطالعه 7 کاربری را بر روی برنامه های محبوب اندروییدانجام داده و پی بردیم که پوشش 7 کاربری برای فیلتر برنامه ها برابر با 30.8 درصد و برای روش های نرم افزار و برنامه نویسی 6.46 درصد است. رویکرد های قبلی برای جست و جوی خودکار برنامه های اندروییدی دارای برنامه هایی در شبیه ساز بوده و یا این که بر برنامه های کوچک تر یمتمرکز هستند که کد منبع آن ها قابل دسترس است. برای حل این مسائل، ما از رویکرد A^3E استفاده کردیم که از طریق آن می توان به جست و جوی سیستمی برنامه های اجرا شده برر روی گوشی های اقعی بدون دسترسی به کد منبع پرداخت. اطلاعات کلیدی مربوط به رویکرد ما شامل استفاده از تحلیل جریان داده ها، استاتیک و سبک قدیمی بر روی بایت کد های برنامه به شکلی جدید است که می تواند طیف وسیعی از فعالیت ها را در نظر بگیرد. ما از این نمودار برای توسعه یک راهبرد جست و جویی موسوم به کشف و جست و جوی هدفمند استفاده می کنیم که امکان جست و جوی مستقیم از جمله فعالیت هایی که طی استفاده طبیعی سخت هستند را می دهد. ما از راهبرد موسوم به جست و جوی اول عمق نیز بهره می بریم که از عملیات کاربر برای فعالیت های اکتشافی و نیز اجزای سازنده آن به طور کند تر ولی سیستمی الگو برداری می کند: ما از دو شاخص استفاده می کنیم: پوشش فعالیت و پوشش روش. ازمایشات با استفاده از رویکرد ما بر روی 25 برنامه اندروییدشامل بی بی سی نیوز، گاز بادی، و غیره بود.

مقدمه

کاربران به طور روز افزونی برای انجام کار های محاسباتی به گوشی های هوشمند متکی هستند(1-2) و از این روی دغدغه هایی نظیر درستی نرم افزار ، عملکرد و امنیت آن دارند(6-8-38،31). تحلیل دینامیگ و پویا یک رویکرد جذاب برای مقابله با این نگرانی ها از طریق پروفیل بندی و پایش است و برای مطالعه خواص و ویژگی های مختلف از مصرف انرژی تا پروفیل بندی و امنیت استفاده کرده است. با این حال، تحلیل دینامیک بر قابلیت دسترسی به نمونه های آزمایشی متکی است که اطمینان از پوشش خوب را می دهد و به این ترتیب اجرای برنامه ها از طریق مجموعه ای از وضعیت های برنامه ای معروف مهم خواهد بود(36-37).

به منظور تسهیل ساخت و ساز و جست و جوی بسیاری از نرم افزار های گوشی های هوشمند چندین رویکر د ارایه شده است. ابزار مانکی(15) قادر به ارسال تصادفی استریم های رویداد به برنامه است و این موجب محدود شدن کارایی کشف می شود. جارچوب هایی نظیر مونوکریمر(24)، راباتیو(18) و تیرود(20) برای برنامه نویسی ارسال میشود و از این روی برنامه نویسی مدت زمان زیادی را طول می کشد. رویکرد های قبلی برای جست و جوی خودکار GUI محدودیت های زیادی را داشته است. این محدودیت ها مربوط به رویکرد برنامه نویسی، اجرای ناقص مدل ها و کشف فضای حالت است.

برای کشف بیشتر مسئله، اجرای یک سری ویژکی های االکتشافی نظیر آمازون تلفن همراه، کاسبودی، یوتیوب، شازم را دوباره به نواز دعوت، و یا سی ان ان را در نظر بگیرید که کد منبع آن ها قابل دسترس نیست. رویکرد ما این کار را به خوبی انجام می دهد زیرا ما به نرم افزار ها و برنامه های اچرا شونده بر روی گوشی تاکید داریم. با این حال رویکرد های موجود دارای مشکلاتی به دلیل نبود کد منبع و نیز اجرای برنامه بر روی شبیه ساز ها که در آن طیف وسیعی از ورودی ها و خروجی ها وجود دارند می باشند. برای مقابله با این چالش ها، ما نرم افزار (A^3E)، را بر اساس ابزار منبع باز و رویکرد خود برای کشف مستقیم نرم افزار های دنیای واقعی که بر روی تلفن های واقعی کار می کنند ارایه می دهیم. توسعه دهندکان می توانند از رویکرد ما به طور مکمل با شرایط ازمایشی موجود با یک سری نمونه های خود ساخته بهره ببرند که هدف آن ها جست و جوی سیستمی است. چون (A^3E)، نیازی به دسترسی به کد منبع ندارد، کاربرانی به جز توسعه دهنده ها می توانند بخش های اصلی برنامه را به طور خودکار اجرا کنند. A^3E از سنسور ها پشتیبانی کرده و از این روی نیاز به ابزار سطح چارچوب یا هسته ندارد و از این روی سربار ایجاد نرم افزار و و شبیه ساز یمحصول اجتناب می شود. از این روی

می توان باور داشت که محققان ومتخصصان می توانند از A^3E به عنوان اساس تحلیل دینامیک استفاده کنند(پایش، پروفیل بندی، مسیر یابی جریان اطلاعات)، تست و اشکال زدایی.

در این مقاله، رویکرد ما بر بهبود هم گرایی در دوسطح متکی است: فعالیت و روش/ فعالیت ها بخش های اصلی از برنامه های اندروییدی هستند- یک فعالیت با صفحه و پنجره در برنامه های مبتنی بر GUI ارتباط دارد. افزایش پوشش فعالیت به معنی جست و جوی صفحات بیشتر است. برای پشش روش ما بر روش های پوشش برنامه تاکید درایم که در بایت کود داولیک موجود است و بر روی داولیک VM بر روی گوشی واقعی اجرا می شود. پیاده سازی فعالیت متشکل از روش های بسیاری است و از این روی با بهبود پوشش روش می توان کارکرد هر یک از فعالیت های جست وجو شده را افزایش داد. در بخش2، ما مروری بر پلاتفرم و برنامه های اندروید داریم و از این روی گراف هایی را ایجاد کنیم که تعریفی از شاخص های پوششی را برای ما می دهند.

برای درک سطح کشف حاصله با کاربران برنامه اندروید، ما یک مطالعه کاربر محور انجام داده و پوشش را در طی تعامل منظم اندازه گیری کردیم. برای اهداف مطالعه 7 کاربر ثبت نام شدهف 28 برنامه اندروید محبوب را اجرا کردندو نتایج نشان داد که در همه برنامه ها و شرکت کننده ها، به طور متوسط، 30.8 درصد همه صفحات برنامه و 6.46 درصد همه روش ها کشف شدند. نتایج و دلایل این سطوح پوشش در بخش 3 ارایه شده است.

در بخش 4، ما رویکرد خود را برای جست و جوی خودکار ارایه می دهیم: با توجه به یک نرم افزار، ما مسیر های جست و جوی سیستمی را برای اهداف مختلف نظیر تحلیل دینامیک ارایه می کنیم. رویکرد ما متشکل از دو فن است: جست و جوی هدفمند یک روش مستقیم و هدفمند می باشد که از تحلیل بایت کد ساکن برای استخراج گراف انتقال فعالیت استاتیک استفاده کرده و سپس گراف را به طور خودکار در زمان اجرا بر روی تلفن انتخاب می کند. حست و چوی عمل اول یک رویکرد کاملا پویا بر اساس کشف خودکار فعالیت ها عناصر GUI می باشد/.

در بخش 5ما مروری بر پیاده سازی A^3E داریم: پلاتفرم سخت افزار، ابزار ها و روش های اندازه گیری. در بخش 6، یک ارزیابی از رویکرد مان بر روی 25 برنامه صورت می گیرد. ما نشان می دهیم که رویکرد ما موثر است: به طور متوسط 64.11 و 59.39 درصد پوشش فعالیت از طریق جست و جوی عمق اول و هدفمند ایجاد شد. این خود به پوشش روش به ترتیب 29.53 و 36.46 درصد از طریق این دو روش رسید. رویکرد ما

کارامد است و یک سری مقادیر به صورت 74 ثانیه برای ساخت گراف فعالیت استاتیک، برای 87 دقیقه برای جست و جوی هدفمند و 104 دقیقه در نظر کرفته می شود.

به طور خلاصه، این مطالعه پیامد های زیر را به دنبال دارد

- مطالعه کمی و کیفی از پوشش بدست آمده توسط 7 کاربر برای 28 برنامه اندروید
- دو رویکرد جست و جوی هدف مند و عمق اول،
- ارزیابی صحت دو جست و جوی هدفمند و عمق اول، بر روی 25 نرم افزار اندروید محبوب

2- فعالیت های اندروید، گراف ها و شاخص ها

ما اندروید را به صورت پلاتفرم هدفمند برای پیاده سازی A^3E انتخاب شده است زیرا امروزه یک پلاتفرم موبایل و سیار موجود در امریکا و سراسر جهان است(3). ما اکنون به توصیف ساختار پیشرفته پلاتفرم اندروید پرداخته و گراف های فعالیت را معرفی می کنیم که جریان کاری پیشرفته را در برنامه تعریف کرده و پوشش را براساس این گراف ها تعریف می کنیم.

2-1 ساختار برنامه اندروید

پلاتفرم و برنامه های اندروید: برنامه های اندروید معمولا در جاوا با برخی کد های بومی نوشته می شوند. کد جاوا به صورت فایل دکس نوشته می شود که دارای بایت کود فشرده است. بایت کد در ماشین مجازی دالویک اجرا می شود که بر روی یک نسخع تلفن هوشمند با برنامه لینوکس کار می کند. برنامه های اندروید به صورت فایل apk می باشند در حالی که با کد های dex یا نام فایل AndroidManifest.xml نوشته می شود.

جریان کار برنامه اندروید

یک چارچوب برنامه ای قوی موجب تسهیل ساختار برنامه اندروید می شود که یک سری منابع و نیز رابط پیش رفته را برای تعامل با ابزار های سطح پایین ارایه می کند. از این روی چارچوب برنامه موجب تنظیم جریان کار برنامه شده و به این ترتیب به ساخت برنامه ها در مورد جریان کنترل کمک می کند/

یک برنامه اندروید متشکل از صفحات مجزا است که به آن فعالیت می گویند. رویکرد ما بر بهبود هم گرایی در دوسطح متکی است: فعالیت و روش/ فعالیت ها بخش های اصلی از برنامه های اندرویدی هستند- یک فعالیت با صفحه و پنجره در برنامه های مبتنی بر GUI ارتباط دارد. افزایش پوشش فعالیت به معنی جست و جوی

صفحات بیشتر است. برای پشش روش ما بر روش های پوشش برنامه تاکید درایم که در بایت کود داولیک موجود است و بر روی داولیک VM بر روی گوشی واقعی اجرا می شود. پیاده سازی فعالیت متشکل از روش های بسیاری است و از این روی با بهبود پوشش روش می توان کارکرد هر یک از فعالیت های جست وجو شده را افزایش داد. در بخش2، ما مروری بر پلاتفرم و برنامه های اندرویید داریم و از این روی گراف هایی را ایجاد کنیم که تعریفی از شاخص های پوششی را برای ما می دهند.

برای درک سطح کشف حاصله با کاربران برنامه اندرویید، ما یک مطالعه کاربر محور انجام داده و پوشش را در طی تعامل منظم اندازه گیری کردیم. برای اهداف مطالعه 7 کاربر ثبت نام شدهف 28 برنامه اندرویید محبوب را اجرا کردندو نتایج نشان داد که در همه برنامه ها و شرکت کننده ها، به طور متوسط، 30.8 درصد همه صفحات برنامه و 6.46 درصد همه روش ها کشف شدند. نتایج و دلایل این سطوح پوشش در بخش 3 ارایه شده است.

یک فعالیت به صورت یک محفظه برای عناصر GUI نظیر پاپ اپ، تست باکس، نمایش متن، اسپینر، ایتم فهرست، پراگرس بار، چک باکس در نظر گرفته می شود. هنگام تعامل با نرم افزار، کاربران معمولا فهالیت های مختلف را هدایت می کنند. از این روی در فعالیت های رویکرد محور این فعالیت ها به کاربر نهایی ارایه می شوند. به همین دلیل، ما بر تابع فعالیت در طی اجرای برنامه متکی هستیم زیرا نقش آن در تست GUI بسیار مهم است.

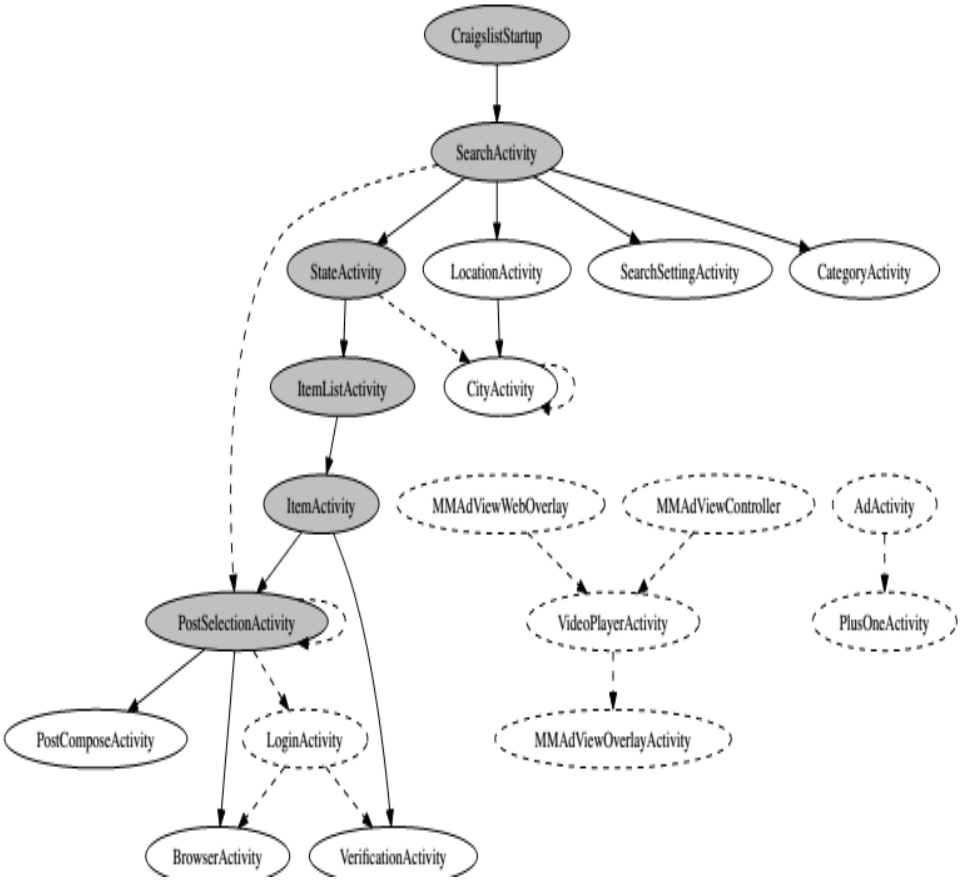
فعالیت ها می توانند اهداف مختلفی را دنبال کنند. برای مثال دریک نرم افزار جدید، فعالیت هوم اسکرین فعرستی از اخبار جدید را نشان می دهد و از این روی یک سری تیتر های خبری را نشان می دهد که موجب تحریک فعالیت های دیگری می شود که یک ایتم خبری کامل را نمایش می دهد. فعالیت ها در درون برنامه جست و چو شده و برخی فعالیت ها از خارج برنامه کنترل می شوند به خصوص اگر کل هاست یا میزبان امکان استفاده از این برنامه ها را بدهد

جدول 1: مروری بر برنامه های بررسی شده

App	Type	Category	Size		# Down-loads
			Kinst.	KBytes	
Amazon Mobile	Free	Shopping	146	4,501	58,745
Angry Birds	Free	Games	167	23,560	1,586,884
Angry Birds Space P.	Paid	Games	179	25,256	14,962
Advanced Task Killer	Free	Productivity	9	75	428,808
Advanced Task Killer P.	Paid	Productivity	3	99	4,638
BBC News	Free	News&Mag.	77	890	14,477
CNN	Free	News&Mag.	204	5,402	33,788
Craigslist Mobile	Free	Shopping	56	648	61,771
Dictionary.com	Free	Books&Ref.	105	2,253	285,373
Dictionary.com Ad-free	Paid	Books&Ref.	49	1,972	2,775
Dolphin Browser	Free	Communication	248	4,170	1,040,437
ESPN ScoreCenter	Free	Sports	78	1,620	195,761
Facebook	Free	Social	475	3,779	6,499,521
Tiny Flashlight + LED	Free	Tools	47	1,320	1,612,517
Movies by Flixster	Free	Entertainment	202	4,115	398,239
Gas Buddy	Free	Travel&Local	125	1,622	421,422
IMDb Movies & TV	Free	Entertainment	242	3,899	129,759
Instant Heart Rate	Free	Health&Fit.	63	5,068	100,075
Instant Heart R.-Pro	Paid	Health&Fit.	63	5,068	6,969
Pandora internet radio	Free	Music&Audio	214	4,485	968,714
PicSay - Photo Editor	Free	Photography	49	1,315	96,404
PicSay Pro - Photo E.	Paid	Photography	80	955	18,455
Shazam	Free	Music&Audio	308	4,503	432,875
Shazam Encore	Paid	Music&Audio	308	4,321	18,617
WeatherBug	Free	Weather	187	4,284	213,688
WeatherBug Elite	Paid	Weather	190	4,031	40,145
YouTube	Free	Media&Video	253	3,582	1,262,070
ZEDGE	Free	Personalization	144	1,855	515,369

شکل 1: مثالی از سناریوی انتقال فعالیت از برنامه محبوب اندروید، آمازون موبایل

گراف انتقال فعالیت استاتیک یگ گراف موسوم به $G_S = (V_S, E_S)$ می باشد که در آن مجموعه ای از رئوس ES نشان دهنده انتقال فعالیت است. ما STAGE را به طور خودکار از تحلیل استاتیگ استخراج می گنیم.



شکل 2: گراف گذار فعالیت استاتیک استخراج شده اتوماتیک با رویکرد برنامه موبایل کرایجیسلیت. گره های خاکستری و یال های مربوطه توسط کاربران کشف شده اند. گره های خطوط پر رنگ (خاکستری و سفید) و یال های خطوط پر رنگ به طور پویا با بررسی ما معکوس می شوند. گره های خط چین و یال ها کشف نشده اند. نام فعالین ها ساده شده اند.

شکل STAG2 را برای برنامه های خرید محبوب نظیر کریک لیست موبایل نشان می دهد. شما می توانید گره ها را در نظر نگرفته و تنها رنگ های حاشیه ای و نیز سبک خطی را در نظر بگیرید. توجه کنید که یک سری فعالیت ها را می توان به طور مستقل در نظر گرفت و از این روی STAG را می توان به شکل گرافت ترسیم کرد و STAG برای درک برنامه بسیار مفید است زیرا یک نمای کلی را از جریان کار نشان می دهد.

گراف انتقال فعالیت پویا

گراف انتقال فعالیت پویا یک گراف $G_D = (V_D, E_D)$ است که در آمئ مجموعه ای از رئوس VD نشان دهنده فعالیت های برنامه است که مجموعه ای از حاشیه های ED را نشان می دهد.

یک DATG می تواند حاصلجست و چوی دینامیک و یا تعامل کاربر به شکل شهودی یا غیر مستقیم باشد و این زیر گرافی از STAG است. شکل2 در بر گیرنده DATG برای طیف وسیعی از مسیر ها و گره ها است.مسیر ها در توالی عمل DATG برای دست یابی به یک وضعیت مطلوب لازم هستند و در ساخت موارد جدید می توان از آن ها استفاده کرد.

ما شاخص ها و سنجه های پوشش را به صورت معیار اندازه گیری و ارزیابی کارایی رویکرد خود استفاده کردیم: پوشش فعالیت و پوشش روش با این حال، جست و جوی نرم افزار های دنیای واقعی بر روی گوشی های واقعی به دلیل عدم قطعیت، عدم جریان کنترل استاندارد، توسعه پذیری و محدودیت های سربار سخت است. با اتکا به کار بران نهایی برای انجام جست و جو ها ما یک مطالعه 7 کاربری را بر روی برنامه های محبوب اندروییید انجام داده و پی بردیم که پوشش 7 کاربری برای فیلتر برنامه ها برابر با 30.8 درصد و برای روش های نرم افزار و برنامه نویسی 6.46 درصد است. رویکرد های قبلی برای جست و جوی خودکار برنامه های اندروییدی دارای برنامه هایی در شبیه ساز بوده و یا این که بر برنامه های کوچک تر یمتمرکز هستند که کد منبع آن ها قابل دسترس است. برای حل این مسائل، ما از رویکرد A^3E استفاده کردیم که از طریق آن می توان به جست و جوی سیستمی برنامه های اجرا شده برر روی گوشی های اقعی بدون دسترسی به کد منبع پرداخت. اطلاعات کلیدی مربوط به رویکرد ما شامل استفاده از تحلیل جریان داده ها، استاتیک و سبک قدیمی بر روی بایت کد های برنامه به شکلی جدید است که می تواند طیف وسیعی از فعالیت ها را در نظر بگیرد. ما از این نمودار برای توسعه یک راهبرد جست و جویی موسوم به کشف و جست و جوی هدفمند استفاده می کنیم که امکان جست و جوی مستقیم از جمله فعالیت هایی که طی استفاده طبیعی سخت هستند را می دهد. ما از راهبرد موسوم به جست و جوی اول عمق نیز بهره می بریم که از عملیات کاربر برای فعالیت های اکتشافی و نیز اجزای سازنده آن به طور کند تر ولی سیستمی الگو برداری می کند: ما از دو شاخص استفاده می کنیم: پوشش فعالیت و پوشش روش. ازمایشات با استفاده از رویکرد ما بر روی 25 برنامه اندروییید شامل بی بی سی نیوز، گاز بادی، و غیره بود.

پوشش فعالیت: پوشش فعالیت به صورت نسبت فعالیت ها طی اجرا به تعداد کل فعالیت های تعریف شده در برنامه $AC = \frac{AR}{AT}$ تعریف می شود. هر چه AC بالاتر باشد، صفحات بیشتری جست و جو می شود و جست

و جوی جامع تری صورت می گیرد. ما به طور دینامیک به AR دست پیدا کردیم و AT آماری در بخش5-2 توصیف شده است.

پوشش روش: پوشش فعالیت به صورت غیر مستقیم است و نشان دهنده درصد صفحات حاصل شده است. به علاوه، کاربران علاقه مند به یک جست و جوی جامع با سطح پایین تر روش هستند. از این روی ما از سنجه دقیق تر یعنی درشد روش ها برای کمی سازی این استفاده کردیم. پوشش روش نسبتی از روش های اجرا شده در زمان اجرا به کل تعداد روش های تعریف شده در برنامه MT می باشد یعنی: $MC = \frac{ME}{MT}$.

پی برده شد که همه برنامه های بررسی شده به جز برنامه Advanced Task Killer، کشتی با گره کتاب خانه شخص ثالث در فایل APK دسته بندی می شوند. ما روش های ثالث را از محاسبات ME-MT خارج کردیم زیرا این روش ها توسط تولید کننده های نرم افزار تعریف نشده اند و از این روی ما استفاده از آن ها غلط انداز است. ما ME را با استفاده از اطلاعات پروفیل بندی زمان اجرا و MT را با تحلیل استاتیک انجام دادیم که در بخش 5-2 بررسی شده اند.

3-مطالعه کاربر: پوشش در طی استفاده منظم

یک رویکرد احتمالی برای کشف، اتکا به اجرای واقعی است یعنی با مشاهده شیوه تعامل کاربران نهایی با برنامه. متاسفانه، این روش سیستماتیک نیست: همان طور که اندازه گیری ها نشان می دهد، در زمان تعامل طبیعی کاربر، پوشش پایین است زیرا کاربران تنها مجموعه کوچکی را از ویژگی ها و کارکرد آن ها کشف می کنند. از این روی،اتکا به کاربران دارای مطلوبیت کم تری هستند. به منظور اندازه گیری کمی پوشش واقعی حاصل شده توسط کاربران نهایی، ما یک مطالعه کاربر محور را انجام دادیم.

مجموعه داده های برنامه: تا مارس 2013، گوپل پلی، که بازار اصلی برنامه اندروید بود، بیش از 600000 برنامه را داشت. ما مجموعه ای از 28 برنامه را برای مطالعه انتخاب کردیم. برنامه ها و ویژکی های آن ها در جدول 1 نشان داده شده است. اولا، ما می خواهیم تا برنامه های رایگان و غیر رایگان را ترکیب کنیم که به این ترتیب 7 برنامه رایگان و غیر رایکان انتخاب شد. دوم، هدف ما مدل سازی مقوله های مختلف نظیر تولید، بازی ها، سرگرمی، اخبار بود و به طور کلی، مجموعه داده های ما دارای برنامه هایی از 17 مقوله مختلف است. سوما، هدف ما برنامه های اصلی بود: اندازه برنامه های انتخاب شده از هزار بایت گد دستور العمل و KBوجود داشت که در

ستون 4 و5 نشان داده شده است. در نهایت هدف ما بررسی برنامه های محبوب بود که در ستون آخر تعداد دانلود ها در کوگل پلی تا 28 مارس 2013 نشان داده شده است که از 2775 تا 6499521 متغیر است. اعتقاد ما بر این است که این مجموعه، طیف وسیعی از برنامه های موبایل محبوب و جهان واقعی را پوشش می دهد.

روش شناسی: ما 7 کاربر مختلف را در مطالعه ثبت کردیم: یک کاربر با پوشش بالا کاربر 1 و شش کاربر مرتب 2تا 7. هر برنامه توسط هر کاربر به مدت 5 دقیقه اجرا شد که طولانی تر از زمان متوسط برنامه به میزان 71.56 ثانیه است. برای استفاده از برنامه واقعی، شش کاربر منظم مورد اموزش قرار گرفتند تا از برنامه به صورت طبیعی عمل کنند. یعنی، کاربران منظم سعی می کردند تا به پوشش بالایی برسند. کاربر 1 بسیار ویژه بود زیرا هدف کاربر دست یابی به پوشش ماکزیمم در محدوده زمانی بود. برای هر دوره، اطلاعات جمع اوری شدند و از این روی ما 192 دور را برای کمی سازی پوشش فعالیت و پوشش روش در نظر گرفتیم.

3-1 پوشش فعالیت

اکنون ما در مورد سطوح پوشش فعالیت بحث می کنیم که بر اساس پوشش کاربر نهایی برای هر شاخص بدست آمده است.

پوشش تجمعی: از آن جا که کاربران مختلف می توانند از ویژکی های مختلف استفاده می کنند به طور کلی با توجه به DATG's G1 و G2، می توان گراف $G = G_1 \cup G_2$ را در نظر گرفت که گره ها و یال های G1-G2 را نشان میدهد. ما از این پوشش تجمعی مبتنی بر پوشش تجمعی به عنوان اساس مقایسه کشف دستی با کشف خودکار استفاده می کنیم.

نتایج: در جدول 2، ما تعداد فعالیت و خلاصه ای از پوشش فعالیت دستی با 7 کاربر را نشان می دهیم. ستون 2 تعداد فعالیت ها را در هر برنامه نشان می دهیم از جمله تبلیغات.

ستون 3 تعداد فعالیت ها به جز تبلیغات را نشان می دهد. ستون 4 نشان دهنده پوشش فعالیت تجمعی است به خصوص زمانی که پوشش از طریق یک گراف ترکیب شود. درصد ها با توجه به ستون 3 محاسبه می شوند و از این روی ما مجموعه ای کامل را در جدول 5 در پیوست ارایه می کنیم.

می توان دید که در پوشش تجمعی، کاربران نسبتا کمی هستند: پوشش تجمعی میانگین 30.08 درصد است. اکنون می توان دلیل این موضوع را یافت.

چرا تعداد فعالیت های کمی کشف شده است؟ گروه فعالیت های از دسته رفته از ستون ها در جدول 2 نشان می دهد که فعالیت ها همگی توسط کاربران از دست می روند و به همین دلیل این فعالیت ها از دست می روند. می توان این فعالیت ها را به مقوله های زیر طبقه بندی کرد:

- ویژگی های کشف نشده: ویژگی های خاص به این دلیل از بین می روند که در نظر گرفته نمی شوند. برای مثال برنامه هایی نظیر دیکشنری دات کام یا Tiny Flashlight + LED ویژگی گوه ای را ارایه می کنند یعنی رابط برنامه عرض تر از ایکون دسکتاپ که امکان دسترسی زیادی را فراهم می کند. مثال دیگر جست و جوی صوتی است که در بروزر دلفین دیده می شود که زمانی کشف می شود که کاربر با صوت اقدام به سرچ کند.
- یکپارچگی شبکه اجتماعی: بسیاری از برنامه ها گزینه ای برای اشتراک اطلاعات در سایت های شبکه های اجتماعی ارایه می کنند به خصوص فیس بوک و تویتر و یا شبکه های شخصی نظیر شازم. درطی استفاده از برنامه های طبیعی، کاربران لزوما مجبور به اشتراک اطلاعات نمی باشند. این انواع فعالیت ها در ستون اجتماعی ظاهر می شوند.
- حساب: بسیاری از برنامه ها می توانند بدون حساب کاربری نیز کارکنند. در صورتی که کاربر با حساب خود وارد شودف می تواند پروفیل خود را دیده و به دیگر فعالیت ها دسترسی داشته باشد. در مواردی که کاربران فاقد حساب باشند، فعالیت های مربوط به حساب اجرا نمی شوند.
- خرید: برنامه های تجارت الکترونیک نظیر آمازون موبایل، گزینه های خرید و فروش را ارایه می کند. اگر کاربر این عملیات را انجام ندهد، این فعالیت ها کشف نخواهند شد.
- گزینه ها: وقتی کاربران محتوی با تنظیمات پیش فرض برنامه هستند و تمظیمات را تغییر نمی دهند با دسترسی به گزینه ها، فعالیت های گزینه اعمال نمی شوند.
- تبلیغات: بسیاری از برنامه های رایگان حاوی فعالیت های مربوط به تبلیغات می باشند. برای مثال، در انگری بیرد، همه فعالیت ها به جز یکی مربوط به تبلیغات هستند. از این روی، به طور کلی برنامه های رایگان حاوی

فعالیت های بیشتری نسبت به انواع هزینه ای می باشند. وقتی که کاربر بر روی تبلیغات کلیک نکند، فعالیت های مربوط به تلیغات کشف نمی شوند

2-3 پوشش روش

چون پوشش فعالیت به طور متوسط 30 درصد بود، انتظار می رود که پوشش روش کم تر از روش های مربوط به فعالیت های کشف نشده باشد. اخرین گروه از ستون ها در جدول 2 ، تعداد کل روش ها را برای هر برنامه و نیز درصد روش های تحت پوشش کاربر 1 و 2-7 را نشان میدهد. می توان دید که پوشش روش کاملا پایین است:6.46 درصد به طور تجمعی برای کاربران 1-7 است. مجموعه داده های کامل در جدول 6 در پیوست نشان داده شده است. در بخش 6-1، گزارش کاملی از دلیل پایین بودن پوشش شبکه وجود دارد.

برنامه	فعالیت ها		درصد پوشش فعالیت	فعالیت های ازدست رفته							روش ها	پوشش روش
	Total #	Excluding ads	Users 1–7 (cumulative)	# Missed	Features	Social	Account	Purchase	Options	Ads		Users 1–7 (cumulative)
Amazon Mobile	39	36	25.64	30	•		•	•			7,154	4.93
Angry Birds	8	1	100	6						•	6,176	10.98
Angry Birds Space Premium	1	1	100	0							7,402	0.68
Advanced Task Killer	7	6	70	3	•				•		3,836	11.46
Advanced Task Killer Pro	6	6	57	2	•				•		427	21.32
BBC News	10	10	52.34	3	•				•		257	7.69
CNN	42	39	19.05	10	•		•				7,725	4.97
Craigslist Mobile	17	15	42	35	•	•	•				2,095	10.76
Dictionary.com	22	18	61	11	•	•				•	2,784	13.83
Dictionary.com Ad-free	15	15	73.33	4	•						1,272	19.10
Dolphin Browser	56	56	12.5	49	•	•					13,800	13.26
ESPN ScoreCenter	5	5	60	2					•		4,398	1.35
Facebook	107	107	5.60	95	•				•		21,896	1.69
Tiny Flashlight + LED	6	4	66.67	4	•					•	1,578	15.91
Movies by Flixster	68	67	23.3	48	•	•				•	7,490	5.32
Gas Buddy	38	33	30.2	29	•	•	•			•	5,792	9.13
IMDb Movies & TV	39	37	25.64	30		•				•	8,463	4.60
Instant Heart Rate	17	14	29.4	15	•	•				•	2,002	4.60
Instant Heart Rate - Pro	17	16	13.2	16	•	•				•	1,927	5.13
Pandora internet radio	32	30	12.5	30			•		•	•	7,620	3.21
PicSay - Photo Editor	10	10	10	9					•	•	1,580	4.39
PicSay Pro - Photo Editor	10	10	33.33	9					•		-[a]	-
Shazam	38	37	15.8	36	•		•			•	9,884	9.43
Shazam Encore	38	37	22.3	33	•	•	•			•	9,914	9.32
WeatherBug	29	24	29	24	•	•				•	7,948	8.15
WeatherBug Elite	28	28	14.30	24	•	•					8,194	6.39
YouTube	18	18	27.77	17	•		•				11,125	5.13
ZEDGE	34	34	38.9	18	•					•	6287	9.27
Mean			*30.08*									*6.46*

جدول 2: نتایج مطالعه کاربر

4-رویکرد

اکنون به بررسی رویکرد خود میپردازیم. اکتشاف هدف که منظور از آن دست یابی به کشف فعالیت است و جست و جوی عمق اول که هدف آن کشف سیستماتیک برنامه ها می باشد، دو راهبرد غیر مکمل می باشند: بلکه ما آن ها را برای دست یابی به اهداف خاص طراحی کرده ایم. جست و جوی عمق اول، GUI تست کرد و با کلیک بر روی تست باکس ها، فعالیت های جدید تر می توان از دکمه بک استفاده کرد. کشف هدف برای مدیریت شرایط خاص لازم است. راهبرد هدف، نیاز است زیرا همه فعالیت ها از طریق تعامل کاربر شروع می شود. هر دو راهبرد کشف را در هر نقطه ورودی شروع کرده، فعالیت های GUI کاربر مانند را تززریق کرده و تولید کال بک هایی برای فعالیت های خاص می کنند

شکل 4: عبور تعمدی از ابر کلاس ها در NPR نیوز

```
//class NewsListActivity extends TitleActivity
public void onItemClick (...)
{
  Intent localIntent = new Intent(this, NewsStoryActivity.class);
  ...
  startActivityWithoutAnimation( localIntent );
}

//class TitleActivity extends RootActivity
public startActivityWithoutAnimation( Intent paramIntent)
{ super.startActivityWithoutAnimation(paramIntent); }

//class RootActivity
protected void startActivityWithoutAnimation( Intent paramIntent)
{ startActivity (paramIntent );...}
```

برای اثبات اصول اصلی این راهبرد ها، از جریان برنامه آمازون موبایل نشان داده شده در شکل 1 نشان داده شده است. در کشف هدفمند، SATG، از طریق تحلیل استاتیک ساخته می شود و کشف ما بر فعالیت های موقت توسط SATG تاکید می کند. در جست و جوی عمق اول ما از نقاط ورودی از SATG برای شروع کشف استفاده می کنیم. سپس در هر فعالیت، ما عناصر GUI را باز یابی کرده وآن ها را به طور سیستمی تعیین می کنیم. در آمازون موبایل، ما با Main activity شروع می کنیم. ما ابتدا به بررسی شیوه ساخت satg پرداخته و سپس کشف هدف را مطالعه می کنیم.

1-4 ساخت SATG

تعیین ترتیب صحیح جست و جوی عناصر GUI باید چالش بر انگیز باشد. مسئله اصلی این است که جریان کنترل برنامه های اندروید غیر استاندارد است. هیچ گونه main وجود ندارد بلکه برنامه ها حول کال بک هایی با چارچوب برنامه در پاسخ به واکنش کاربر متمرکز است. به این ترتیب استدلال در مورد کنترل چریان پیش فرض است.برای مثال اگر فعالیت فعلی A باشد و به سمت B انتقال یابد، روش های مربوط به A ارتباط مستقیم با B ندارد. در عوض بر اساس منطقی می توان عمل گرد. هدف ما استفاده از تحلیل جریان داده دها می باشد. یک قسمت کلیدی این است که ساخت SATG به مسئله پایش کاهش می یابد.

بازگشت به مثال قبلی با فعالیت A-B، استفاده از تحلیل راه اندازی را نشان میدهد به طوری که اگر به درخواست اصلی از A برسیم، فعالیت B ازA قابل دسترس است. اهداف اصلی را می توان برای شروع فعالیت ها با استدلال startActivity در نظر گرفت. این اهداف برای شروع خدمات استفاده شده و پیام هایی به برنامه های دیگر ارسال می شوند. از این روی، درک فعالیت های جریان باید در نظر گرفته شود.

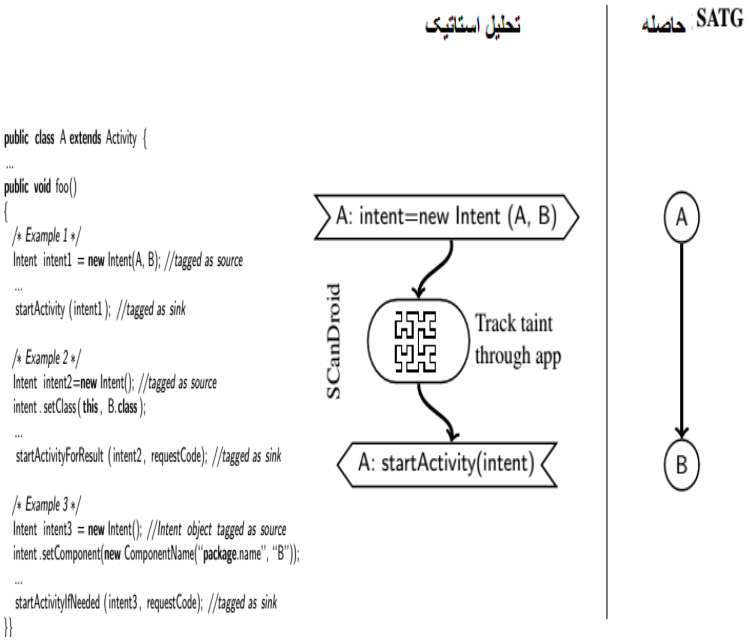
اکنون مثال هایی را در مورد شیوه استفاده از ساختار SATG را با تحلیل تینت نسبت به منطق اصلی ارایه می کنیم. در شکل 3، در قسمت چپ، کد جاوای اندروید برای کلاس A در نظر گرفته شده است. برنامه نویسی را در نظر بگیرید که کلاس فعالیت B را نشان می دهد. این سه مثال با بررسی بخش های A تعیین می شوند. در مثال 2، ارتباط با تظنیم کلاس B به صورت کلاس هدف انجام می شود. در مثال 3، B به صورت ملوفه ای از هدف باشد. تحلیل ما علایم را به صورت منابع در نظر می گیرد.مخازن به صورت startActivityForResult و startActivityIfNeeded می باشند. اگرچه ما کد جاوا را برای شفافیت نشان می دهیم تحلیل ما بر روی بیت کد صورت می گیرد. پایش در شکل 3 تشان داده شده است. از این روی اصل کلی ساخت و ساز STAG، شناسایی نقاط ساخت هدف به صورت منبع می باشد و فعالیت درخواست را به صورت محازن شروع می کند.

یک مثال دنیای واقعی پیشرفته تر از شیوه تحلیل ما یک سلسله مراتب در شکل 4 نشان داده شده است. و از این روی هدف در NewsListActivity .onItemClick (...) به صورت منبع مد نظر قرار می گیرد. وقتی تحلیل تمام شد، بر اساس مقدر الودگی شناسایی شده، حاشیه ای NewsListActivity to NewsStoryActivity در STAG در نظر گرفته می شود.

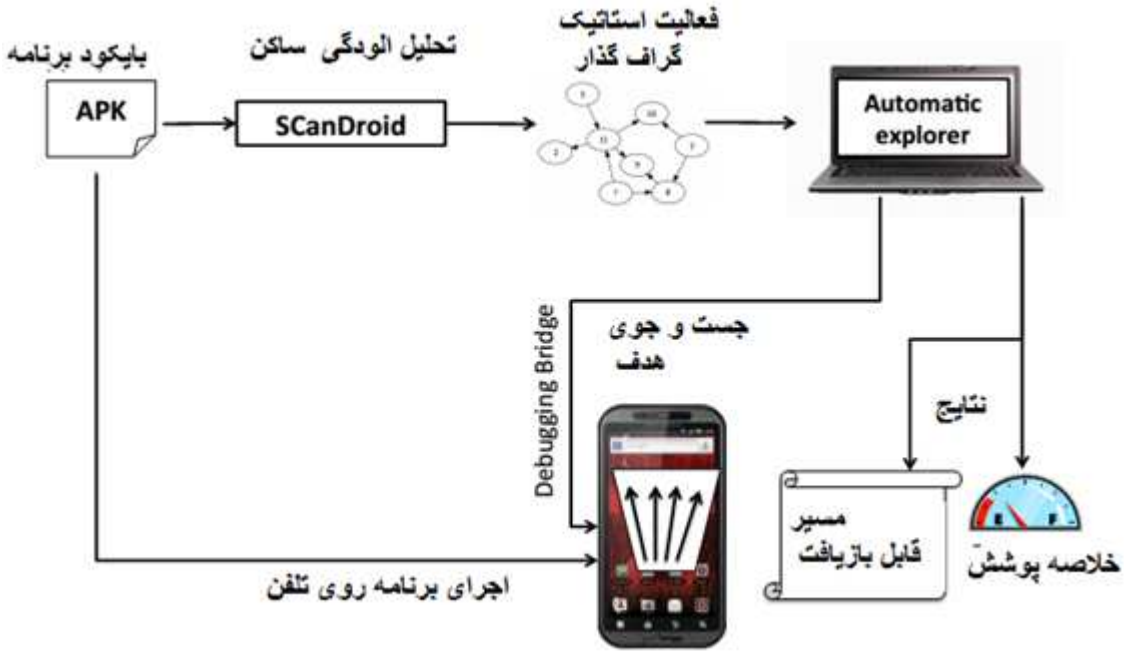
2-4 کشف هدفمند

ما شیوه انجام جست و جوی هدف را با STAG توصیف می کنیم. جست و جوگر حودکار بر روی لپ تاپ در نظرفته می شود و کشف را هماهنگ می سازد.

منابع و مخازن تگ



شکل 3: ساخت STAG با تحلیل آلودگی:منابع و مخازن به طور خودکار نشان گذاری می شوند و الودگی با SCanDroid و STAG تعیین می شود



شکل 5: مروری بر اکتشاف هدف در A3E

اولین جست و جوگر ،STAG ساخته شده را با SCanDroid می خواند. الگوریتم ساخت STAG فهرست همه فعالیت های را نشان می دهد. فعالیت های جست و جو شده فهالیت هایی هستند که مستقل از برنامه هستند. همه فعالیت ها را می توان از فرایند طبیعی بدست اورد. برای مثال، وقتی که فعالیت پارامتر هایی را از فعالیت قبلی بگیرد، پارامتر ها حاوی اطلاعات امنیتی است که محدود به حوزه کاربرد است.

سپس، جست و جو گر بر روی الگوریتم جست و جوی هدف اجرا می شود که به طور کوتاه توصیف می شود. جست و جوگر اجرای برنامه را کنترل کرده و با تلفن از طریق پل دباگینک اندروید ارتباط برقرار می کند. نتیجه جست و جو متشکل از مسیر قابل بازیافت است.

اکنون به توصیف الکوریتم جست و جوی هدف می پردازیم: بخشی از الکوریتم روی تلفن اجرا می شو.د. در یک نات شل، STAG حاوی یال های A-B است و. این نشان دهده کذار فعالیت حقوقی است. با فرض کشف فعالیت A، دو مورد برای B وجود دارد 1-B صادرات پذیر است یعنی، از A قابل دسترس است ولی ناشی از تعامل GUI در A است و 2- B از A به دلیل تعامل GUI در A حاصل می شود.

الگوریتم A: جست و جوی هدف

Input: SATG $G_S = (V_S, E_S)$

1: **procedure** TARGETEDEXPLORATION(G_S)
2: **for** all nodes A_i in V_S that are entry points **do**
3: Switch to activity A_i
4: $currentActivity \leftarrow A_i$
5: **for** all edges $A_i \rightarrow A_j$ in E_S **do**
6: **if** A_j is exportable **then**
7: Switch to activity A_j
8: $currentActivity \leftarrow A_j$
9: $G'_S \leftarrow$ subgraph of G_S from starting node A_j
10: TARGETEDEXPLORATION(G'_S)
11: **end if**
12: **end for**
13: $guiElementSet \leftarrow$ EXTRACTGUIELEMENTS($currentActivity$)
14: **for** each $guiElement$ in $guiElementSet$ **do**
15: exercise $guiElement$
16: **if** there is an activity transition to not-yet-explored activity A_n **then**
17: $G'_S \leftarrow$ subgraph of G_S from starting node A_n
18: $currentActivity \leftarrow A_n$
19: TARGETEDEXPLORATION(G'_S)
20: **end if**
21: **end for**
22: **end for**
23: **end procedure**

الگوریتم 1 توصیف دقیقی از رویکرد جست و جوی هدف ارایه می کند. این الکوریتم با STAG به صورت ورودی شروع می شود. ابتدا فعالیت ورودی برنامه از STAG شروع می شود. ما همه فعالیت های A_j را جست و جو می کنیم که دارای یال ورودی از Aj میباشد. وقتی که به هر فعالیت سوییچ کنیم، فعالیت A_n قابل خروج شدن نمی باشند ولی از Ai قابل دسترس است. و سپس الکوریتم را از A_n بدست می اوریم

مزیت جست و جوی هدف این است که به پوشش فعالیت از رویداد های GUIمی رسد.

شکل6: مروری بر جست و جوی عمق اول

الکوریتم 2: جست و جوی عمق اول

```
Input: Entry point activities |A|
 1: procedure DFE(|A|)
 2:     for all nodes A_i in |A| do
 3:         Switch to activity A_i
 4:         DepthFirstExploration(A_i)
 5:     end for
 6: end procedure
 7:
 8: procedure DepthFirstExploration(A_i)
 9:     guiElementSet ← ExtractGuiElements(A_i)
10:     for each guiElement in guiElementSet do
11:         excercise guiElement
12:         if there is an activity transition to not-yet-
    explored activity A_n then
13:             DepthFirstExploration(A_n)
14:             Switch back to activity A_i
15:         end if
16:     end for
17: end procedure
```

3-4 جست وجوی عمق اول

اکنون جست و جوی عمق اول را ارایه می کنیم که زمان بیشتری را می برد. چون این روش پویایی است، جست و جوی عمق اول را می توان انجام داد به خصوص زمانی که تستر فاقد اطلاعات کذار است.

شکل 5 یک مرور اجمالی را ارایه کرده است. در این رابطه STAG استفاده نمی شود ولی حست و جوگر هدف متفاوت است و الگوریتم جست و جوی عمق اول توصیف می شود. بقیه عملیات مشابه با جست و جوی هدف است یعنی جست و جو گر، جست و جو را تنظیم می کند و نتایج اطلاعات پوشش و قابل بازیافت در نظر گرفته می شود.

الگوریتم 2 توصیف دقیقی از روش جست و حوی عمق اول ارایه می کندو مشابه با جست و جوی هدف، ما از فعالیت های نقطه ورودی از APK استفاده می کنیم. این فعالیت ها نقطه شروعی برای کشف هستند. و سپس ما به بررسی نقطه ورودی A_i و جست و حوی عمق اول از نقطه 1-5 می پردازیم. به ازای هر فعالیت A_i ، همه عناصر GUI استخراج می شوند. سپس به طور سیستمی عناصر GUI با پاسخ به کنترل گر اجرا می شوند. وقتی که ما تغیرات فعالیت جدید را یافتیم، از الگوریتم مشابه بر روی A_n استفاده می کنیم. این فرایند به صورت جست و جوی عمق اول استفاده می شود.

4-4 اشکال زدایی و موارد ازمایشی قابل پاسخ

در طی جست و جو، A^3E به طور خودکار ، شاخه رویداد را با استفاده از RERAN ثبت می کند به طوری که کشف پاسخ داده می شود. این ویژگی به کاربران در ساخت نمونه های آزمایشی کمک می کند به طوری که بعد ها توسط RERAN اجرا می شود. به علاوه، ترکیب با RERAN موجب تسهیل اشکال زدایی می شود و از این روی شاخه ای را استخراج می کنیم که منجر به خرابی می شود

5-پیاده سازی

اکنون شرایط ازمایشی، جزییات و روش های اندازه گیری برای ساخت و ارزیابی A^3E در نظر گرفته می شود

5-1 شرایط و ابزار

تلفن های هوشمندمورد استفاده برای ازمایش شامل بیونیک درواید موتورلا با نسحه اندروید 2.3.4 می باشد. تلفن دارای ویژکی Dual Core ARM Cortex-A9 CPU می باشد که بر روی یک گیگاهرتز اجرا می شود.آزمایشات MacBook Pro بر روی Mac OS X 10.8.3 استفاده شد.

برای مطالعه کاربر از REREAN به عنوان ابزاری برای ثبت تعامل کاربر استفاده شد.

SCanDroid ابزاری برای تحلیل ساکن بر روی بیت کود دلویگ توسط محققان دیکر است. برای این منظور به توسعه SCanDroid در دو جهت می پردازیم 1- برای نشان گذاری اهداف و روش های سبک زندگی به صورت منبع و محزن 2- فهرست همه روش ها در نظر گرفته میشود

5-2 اندازه گیری پوشش

پوشش فعالیت: جست و جو گر خودکار AR را پایش می کند و این از طریق دستگاه logat توسط اندروید دیباک بریج ارایه می شود.

تعداد کل فعالیت های AT، به صورت افلاین بدست آمد: ما از `apktool` برای استخراج فایل حاصل از APK منبع باز استفاده کرده و مانیفست را برای فهرست بندی فعالیت ها تجزیه می شوند. از AT و AR، ما فعالیت های خروجی را کنار می گذاریم و از این روی این بخشی از پایگاه کد برنامه نیست. مثال هایی از فعالیت های خروجی فعالیت های مربوط به تبلیغات بوده و فعالیت های سیسستم های خارجی نظیر بروزر، میوزیک پلیر و دوربین را نشان می دهد.

پوشش روش: اندروید OS یک مدیریت برنامه را ارایه می کند که ایجاد پروفایل های آنی می کند در حالی که برنامه اجرا می شود.برای اندازه گیری ME، تعداد روش ها در طی اجرا از داده های پروفیل بندی کزارش شده توسط am استخراج می شود. MT اندازه گیری شد که تعداد کل روش ها در برنامه می باشد که فهرست همه درخواست روش های مجازی را در برنامه می دهد. توجه داشته باشید کد ثالث در محاسبات ME-MT در تظر گرفته نمی شود.

5-3 جست و گر خودکار

استخراج عنصر GUI برای هر دو جست و جوی عمق اول و هدفمند نیاز است. برای کشف عناص GUI، A^3E ا استفاده می شود. روبوتیوم قادر به استخراج هندلر برای مجموعه غنی از عناصر GUI است. این مجموعه شامل

فهرست ها، دکمه ها، جک باکس ها، دکمه های تاگل، تصاویر، تکست ویو، دکمه تصویر و غیره هستند. تیود به توسعه دهنده ها امکان رایت اسکریپت های رابی را بر روی درایو برنامه می دهد.

A^3E بر روی تیود اجرا می شود. ما تیود را طوری اصلاح کردیم که امکان هدایت خودکار را بر روی برنامه بدهد. هر یک از صفحات اندروید متشکل از عناصر GUI هستند که از طریق هندلر ها لینک می شوند. به این ترتیب تقشه سیستماتیک از رویداد تعیین شده و کاربر واقعی تقلید می شود. این دانش زمان اجرا برای کارکرد حست و چوگر خودکار لازم است. همان طور که در بخش 3-1 گفته شد، ما به تست ویژگی های کاربران نظیر گزینه ها، تبلیغات، تنظیمات و اشتراک در شبکه های اجتماعی می پردازیم. برای پوشش دادن این فعالیت ها و برنامه ها، A^3E از راهبرد های مختلف استفاده می کند. A^3E به طور خودکار قادر به شناسایی فعالیت های مربوط به مسئولیت پذیری خاص نظیر ورود به صفحه و شبکه های اجتماعی است. ما مجموعه ای از اطلاعات.. A^3E را ایجاد کرده و برنامه را به کاربر ارسال می کنیم.

ما رویکرد را بر روی چارچوب تست روبوتیوم اجرا کرده و سپس محدودیت ها را برطرف می کنیم. یک محدودیت، ناتوانی روبوتیوم برای ایجاد و ارسال حرکات است که می تواند به تنهایی بر روی روبوتیوم اجرا شود. برای حل این محدودیت، کتابخانه ای از حرکات را می توان نوشت. علاوه بر ایجاد کتابخانه و نیز کارکرد های ورودی، A^3E نیز از میکروفون، جی پی اس، پرگار و نیز شتاب سنج استفاده می کند. با این حال برنامه های خاص نیازمند ورودی های پیچیده نظیر پردازش فایل انتخابی توسط کاربر است. با این کتابخانه از رویداد های ورودی و راهبرد های جدید GUI، A^3E از الگوریتم کشف بسته به نوع کشف استفاده می کند.

4-5 جست و جوی هدف

بخش 4-1 به بررسی منطق عبور عمدی از میان فعالیت های برنامه های داخلی پرداخته و جست و جوی هدف از STAGپیش ساخته برای کشف این مسیر ها بهره می برد. با این حال اندروید امکان درخواست فعالیت ها را از برنامه های خارجی از طریق پیام رسانی می دهد: برنامه ها قادر به تعریف فیلتر هدف برای اطلاع رسانی OS هستند که فعالیت های برنامه قادر به شناسایی هدف می باشند. از این روی، در زنان انجام سیستماتیک عناصر GUI، این احتمال وجود دارد که عناصر GUI و فعالیت های خارجی از بین بروند. علاوه بر این فیلتر ها، توسعه

دهنده ها قادر به شناسایی فعالیت ها با تعریف android:exported="true' می باشد. در زمان اجرای کشف هدف، فعالیت های خارجی تعیین می شود.

5-5 جست و جوی عمق اول

با زیر ساخت استحراج پویای GUI و رویداد های موجود، ما جست و جوی عمق اول را با راهبرد عمق اول در نظر گرفتیم. هر زمان یک تغییر به فعالیت جدید یافت شد. این فرایند تا زمانی ادامه دارد که فعالیت جدید یافت نشود. در این نقطه باید به فعالیت اخر اشاره کرد.

6- ارزیابی

ما اکنون نتایج آزمایشی را ارایه می کنیم. از 28 برنامه بررسی شده در در بخش 3، 3 مورد قابل بررسی نبودند زیرا با کد منفی نوشته می شوند.

ما ابتدا به ارزیابی روش های کشف خودکار بر روی این برنامه ها از حیث کارایی و اثر بخشی پرداخته و در مورد ویژگی های برنامه ها صحبت کردیم.

6-1 اثر بخشی و کارایی

پوشش فعالیت:ما نتایج پوشش فعالیت را در جدول 3 ارایه کرده ایم. ستون 2 تعدادی از گره ها را در STAG نشان می دهد یعنی تعداد فعالیت ها ها در هر برنامه به جز تبلیغات. ستون های گروه بندی شده 3-5 پوشش فعالیت را بر اساس درصد می دهد و در هر سه سناریو، پوشش تجمعی برای کاربران 1-7پوشش از طریق جست و جوی هدف حاصل می شود و پوشش از طریق جست و جوی عمق اول حاصل می شود. در این جا چندین مشاهده وجود دارد. اول، جست و جوی سیستمی موجب افزایش پوشش فعالیت تا ضریب 2 از 30.08 توسط 7 کاربر تا 64.11 و 59.39 درصد حاصل از حست و حوی هدف و عمق اول می شود. از این روی رویکرد ما در فعالیت های حست و جوی سیستمی موثر تر است. دوم، STAG نیازمند هزینه است زیرا باید از تغییرات کشف شده آماری استفاده کند.

پوشش روش: نتایج پوشش روش در ستون آخر جدول 3 نشان داده شده است. ستون روش تعداد روش های تعریف شده را در برنامه نشان می دهد از جمله کد ثالث. ستون بعدی پوشش فعالیت را بر حسب درصد نشان می دهد: پوشش تجمعی برای کاربران 1-7، پوشش از جست و جوی هدفمند صورت می گیرد. ستون های گروه بندی

شده 3-5 پوشش فعالیت را بر اساس درصد می دهد و در هر سه سناریو، پوشش تجمعی برای کاربران 1-7پوشش از طریق جست و جوی هدف حاصل می شود و پوشش از طریق جست و جوی عمق اول حاصل می شود. در این جا چندین مشاهده وجود دارد. اول، جست و جوی سیستمی موجب افزایش پوشش فعالیت تا ضریب 4.5 از 6.46 توسط 7 کاربر تا 29.53 و 36.43 درصد حاصل از جست و جوی هدف و عمق اول می شود. از این روی رویکرد ما در فعالیت های جست و جوی سیستمی موثر تر است. دوم، STAG نیازمند هزینه است زیرا باید از تغییرات کشف شده آماری استفاده کند.

زمان اکتشاف: در جدول 4 زمان مورد نیاز برای جست و جو را نشان می دهد. ستون 2 دارای زمان تحلیل ساکن است که برای ساخت STAG لازم است. این بسیار موثر است زیرا حداقل 10 دقیقه زمان لازم است. زمان جست و جو با هر دو جست و جوی هدف و عمق اول بررسی شد.

برنامه	فعالیت	پوشش فعالیت (%)			روش	پوشش روش (%)		
		Users 1–7 (cumulative)	هدفمند	عمق اول		Users 1–7 (cumulative)	هدفمند	عمق اول
Amazon Mobile	36	25.64	63.90	58.30	7,154	4.93	28.1	45.10
Angry Birds	-	100	-	-	-	10.98	-	-
Angry Birds Space Premium	-	100	-	-	-	0.68	-	-
Advanced Task Killer	6	70	83.33	83.33	420	11.46	59.76	62.86
Advanced Task Kill. P.	6	57	83.30	83.30	257	21.32	39.30	73.20
BBC News	10	52.34	80.00	80.00	3,836	7.69	31.80	37.40
CNN	39	19.05	69.23	61.54	9,269	4.97	29.88	29.97
Craigslist Mobile	15	42	73.30	66.70	2,095	10.76	30.50	41.10
Dictionary.com	18	61	83.33	72.22	3,881	13.83	44.29	44.62
Dictionary.com Ad Free	15	73.33	100	80	1,846	19.10	47.72	49.13
Dolphin Browser	56	12.50	42.86	37.50	17,007	13.26	42.92	43.37
ESPN ScoreCenter	5	60	80.00	80.00	4,398	1.35	16.10	31.20
Facebook	107	5.60	-	-	-	1.69	-	-
Tiny Flashlight + LED	4	66.67	75	75	1,837	15.91	28.03	47.63
Movies by Flixster	67	23.30	77.60	61.20	10,151	5.32	29.50	31.80
Gas Buddy	33	30.20	72.70	63.60	5,792	9.13	31.40	47.80
IMDb Movies & TV	37	25.64	54.10	62.10	11,950	4.60	29.80	32.40
Instant Heart Rate	14	29.40	42.86	35.71	2,407	4.60	20.40	23.18
Instant Heart Rate - Pro	16	13.20	37.50	37.50	2,514	5.13	26.05	26.21
Pandora internet radio	30	12.50	80.0	76.70	7,620	3.21	21.10	31.70
PicSay - Photo Editor	10	10	50	40	1,458	4.39	25.58	27.43
PicSay Pro - Photo Editor	10	33.33	50	40	-	-	-	-
Shazam	37	15.80	45.95	45.95	12,461	9.43	34.74	35.67
Shazam Encore	37	22.30	45.90	51.40	9,914	9.32	29.10	36.30
WeatherBug free	24	29	54.17	45.83	7,744	8.15	40.05	40.33
WeatherBug Elite	24	14.30	91.70	87.50	7,948	6.39	17.20	25.70
YouTube	18	27.77	55.56	50	14,550	5.13	26.95	26.99
ZEDGE	34	38.90	67.60	67.60	6,287	9.27	16.60	24
Mean		*30.08*	*64.11*	*59.39*		*6.46*	*29.53*	*36.46*

شکل3: نتایج ارزیابی: پوشش روش و فعالیت

ما محدوده زمانی را تحمیل نکردیم. ستون 3-4 زمان جست و جوی دینامیک را نشان می دهد و 18-236 دقیقه برای جست و جوی هدف و 39-239 دقیقه برای جست و جوی عمق اول در نظر گرفته شد از این روی این روش به طور کارامد از جست و جوی سیستمی استفاده می کند. همان طور که انتظار می رفت، جست و جوی هدف سریع تر بود و بعد از افزودن زمان STAG، این زمان ها قابل قبول می باشند. جست و جوی هدف نیازمند تحلیل جریان داده های ساکن اولیه برای ساخت STAG است. با این حال STAG به صورت سریعی ساخته می شود و این طوری است که می توان به پوشش فعالیت بالا با سوییچینک فعالیت ها دست پیدا کرد.

جست و جوی عمق اول نیازمندSTAG نیست با این حال مرحله جست و جو کند تر است و به طور سیستمی همه عناصر GUI نیازمند جست و جوی طولانی است. با این حال، این جست و جوی طولانی نیازمند پوشش روش بالاتر است.

در این جا مثال هایی وجود دارد که موجب تشریح نتایج مطلوب تر می شود. برای برنامه های Advanced Task Killer وESPN ScoreCenter ، می توان پوشش فعالیت مشابه را بدست اورد ولی پوشش روش ESPN ScoreCenter به طور معنی داری پایین تر است دلیل اصلی ساختار نرم افزار است.

App	Time		
	SATG (seconds)	Targeted (minutes)	Depth-first (minutes)
Amazon Mobile	222	123	131
Advanced Task Killer	39	41	47
Advanced Task Kill. P.	24	27	58
BBC News	68	18	52
CNN	14	158	161
Craigslist Mobile	43	83	91
Dictionary.com	66	113	131
Dictionary.com Ad Free	45	153	156
Dolphin Browser	595	171	179
ESPN ScoreCenter	42	22	44
Tiny Flashlight + LED	52	33	39
Movies by Flixster	53	228	219
Gas Buddy	157	109	124
IMDb Movies & TV	107	135	126
Instant Heart Rate	56	47	51
Instant Heart Rate - Pro	50	48	49
Pandora internet radio	92	89	111
PicSay - Photo Editor	36	119	121
PicSay Pro - Photo Ed.	40	112	129
Shazam	64	236	239
Shazam Encore	248	188	230
WeatherBug free	120	69	107
WeatherBug Elite	119	115	124
YouTube	200	131	135
ZEDGE	124	97	114
Mean	*74*	*87*	*104*

جدول 4: نتایج ارزیابی: زمان ساخت و جست و جوی STAG. به واحد های زمانی ستون ها نگاه کنید.

ESPN ScoreCenter از فعالیت های پیچیده استفاده می کنند و آرایش های مختلف در یک فعالیت با ویژگی های زیاد بررسی نی شوند. الکوریتم هدف، فعالیت ها را بدون جست و جوی عناصر عمقی بررس می کند. به همین دلیل، جست و جوی هدفمند به طور معنی داری برای Advanced Task Killer Pro وBBC News سریع تر است.

بسیاری از برنامه ها پوشش فعالیت بهتر را برای جست و جوی هدف به جای عمق اول نشان می دهند. زیرا عمدتا دارای نقاط ورودی چند گانه هستند و یا دارای فعالیت هایی برای تحریک عملکرد هستند. برای مثال امازون موبایل دارای فعالیت جست و جوی بارکد است که در طی جست و جوی عمق اول دیده نمی شود. یک استثنا IMDb Movies & TV و Shazam Encore می باشد: هر دو برنامه دارای پوشش فعالیت پایین در جست و جوی هدف به جای عمق اول است. بعد از بررسی مشخص شد که برخی از فعالیت ها با جست و جوی هدفمند با استفاده از فیلتر با پارامتر ها تحریک می شود ولی جست و جوی هدفمند قادر به اجرای فعالیت نیست. این ناشی از پارامتر های ورودی است که تولید جست و جوی هدفمند می کند. جست و جوی عمق اول برنامه راهمانند کاربر اجرا می کند و بر روی صفحات خاص قرار می دهد. زمان حست و جو در جدول 4 نیز برخی از نقاط ورودی را نشان می دهد. زمان جست و جو بستکی به اندازه و پیچیدگی GUI دارد به طور طبیعی جست و جوی عمق اول کند تر از جست و جوی هدف است. جست و جوی عمق اول به فعالیت ها بر اساس پارامتر های هدف رسید در حالی که جست و جوی هدفمن با فعالیت های تحریک کننده قبل از ایجاد پارامتر ها تعیین شدند.

2-6 باز دارنده ها و کاتالیزور های جست و جوی خودکار

ما اکنون تجربه را با 28 برنامه بررسی کرده و در مورد ویژگی هایی صحبت می کنیم که تسهیل کننده یا باز دارنده جست و جو از طریق A^3E است. دلیل اصلی دست یابی به پوشش 100 درصدی از لحاظ فنی و بنیادین است. دلایل فنی شامل حضور فعالیت هایی است که به طور خودکار به دلیل ماهیت آن ها جست و جو نمی شوند.

حرکات و ورودی های پیچیده: روش های ما می تواند یک مسیر خوب را از برنامه های ساخته شده ارایه کند برای درک سطح کشف حاصله با کاربران برنامه اندرویید، ما یک مطالعه کاربر محور انجام داده و پوشش را در طی

تعامل منظم اندازه گیری کردیم. برای اهداف مطالعه 7 کاربر ثبت نام شدهف 28 برنامه اندروید محبوب را اجرا کردندو نتایج نشان داد که در همه برنامه ها و شرکت کننده ها، به طور متوسط، 30.8 درصد همه صفحات برنامه و 6.46 درصد همه روش ها کشف شدند. نتایج و دلایل این سطوح پوشش در بخش 3 ارایه شده است. یک فعالیت به صورت یک محفظه برای عناصر GUI نظیر پاپ اپ، تست باکس، نمایش متن، اسپینر، ایتم فهرست، پراگرس بار، چک باکس در نظر گرفته می شود. هنگام تعامل با نرم افزار، کاربران معمولا فهالیت های مختلف را هدایت می کنند. از این روی در فعالیت های رویکرد محور این فعالیت ها به کاربر نهایی ارایه می شوند. به همین دلیل، ما بر تابع فعالیت در طی اجرای برنامه متکی هستیم زیرا نقش آن در تست GUI بسیار مهم است. فعالیت ها می توانند اهداف مختلفی را دنبال کنند. برای مثال دریک نرم افزار جدید، فعالیت هوم اسکرین فعرستی از اخبار جدید را نشان می دهد و از این روی یک سری تیتر های خبری را نشان می دهد که موجب تحریک فعالیت های دیگری می شود که یک ایتم خبری کامل را نمایش می دهد. فعالیت ها در درون برنامه جست و چو شده و برخی فعالیت ها از خارج برنامه کنترل می شوند به خصوص اگر کل هاست یا میزبان امکان استفاده از این برنامه ها را بدهد.

برنامه های عرضه خدمات:برخی برنامه ها به صورت عرضه کننده خدمات با اجرای برنامه هایی عمل می کنند که به صورت پیش زمینه اجرا می شوند. برای مثال WeatherBug در پیش زمینه هوا شناسی اجرا می شود و Advanced Task Kille در پیش زمینه مربوط به زمان اندازه گیری تعیین می شود. از این روی برای این برنامه ها می توان یک سری برنامه های مربوط به سرویس را اجرا کرد

کد بومی: در نهایت، تحلیل استاتیک ما بر روی بایت کد دلویک VM کار می کند. اگر برنامه از کد بومی به جای دلویک استفاده کند، می توان از روش های خاصی استفاده کرد. برای مثال، دو برنامه انگری بیرد از کد OPEN Gl استفاده می کند تا مدیریت تصویر را انجام دهد و استحراج عنصر gui به اطلاعات GUI بومی دسترسی ندارد.

7-کار نسبی

مطالعات پاستوگی و همکاران(10) با کار های ما هم خوانی دارد. سیستم آن ها موسوم به پلی گراند است که در یک امولاتر بر روی نرم افزار های اندروید اجرا می شود. هدف نهایی پایش الودگی پویا است. مطالعه آن ها یک

روش جست و جو را ارایه کرده است که موسوم به اجرای هوشمند است. اجرای هوشمند نیازمند شروع برنامه به طور پویا با عناصر GUI و جست و جوی آن ها بر اساس سیاست های توالی و دنباله ای است. پلی گراند بر روی 3986 برنامه اجرا شد و پوشش کد به طور متوسط 33 درصد بود. اولا، برنامه ها روی نرم افزاری در اندروید انجام شدند. دوما، پلی گراند همانند جست و جوی عمق اول برخی از فعالیت ها رانادیده می گیرد. سوما، راهبرد جست و جوی GUI بر اساس اصول اکتشافی است.چهارما، چون ما ازمایشات را بر وی تلفن های واقعی انجام دادیم، VM قادر به جمع اوری پوشش اطلاعاتی نیست.

ممون و همکاران بر روی تست GUI در برنامه های دسکتاپ کار کردند و به مدل سازی برنامه GUI پرداختند. روش آن ها برای مدل سازی GUI به صورت گراف تعاملی رویداد استفاده شد.

EIG توالی هایی از فعالیت های کاربر را نشان می دهد که بر روی GUIاجرا می شود. اگرچه روش EIG برای طراحی راهبرد های جست و جو برای تست GUI در برنامه مناسب است، چندین عامل در استفاده از برنامه های لمسی نقش دارند. اولا، تغیرات مربوط به مولفه های غیر فعال ارتباطی با برنامه های هوشمند ندارند و از این روی اگر GUI متشکل از گوه هایی بر روی کانواس باشد، هر یک به صورت گراف مدل سازی می شوند. سپس گراف ها می توانند قابل پایش باشند به علاو.ه عمل کاربر بعدی بر حالت کانواس اثر دارد. رای مثال فعالیت com.aws.android.lib.location.LocationListActivity در برنامه WeatherBug حاوی لایه هایی است که هر کدام چندین گوه دارند و به این ترتیب EIG از گراف های دو بخشی با مجموعه ای از یال ها استفاده می کند. سوما، حالت GUI از خارج از برنامه اجرا می کند. رفتار درخواست تماس موجب اصلاح وضعیت های GUI می شود. از این روی فهرستی از توالی های عمل ایجاد می شود که مجموعه ای از حالت های GUI استUITAR یک چارچوب GUI بر روی EIG است. اندروید GUITAR با توسعه اندروید Android SDK’s MonkeyRunner را ارایه می کند.

یانگ و همکاران(11) ابزاری را برای جست و جوی خودکار موسوم به اربیت ارایه کرده اند رویکرد آن ها از تحلیل اسکن بر روی کد منبع جاوا برای تشخیص فرایند های رمبوط به حالت GUI استفاده کرده و سپس از مرمر گر پویا یا روبتیوم بهره می برد. ما از تحلیل ساکن روی بایت کد برنامه برای استخراج STAG به صورت فعالیت های پایدار استفاده کرده این ولی از جست و جوی GUI برای مقابله با ارایش پویای درون فعالیت استفاده می کنیم.آن

ها به پوشش معنی دار 63-91 درصد در مجموعه ای از 8 برنامه منبع باز دست پیدا کردند. ما بر حوزه مسئله تاکید داریم و برنامه های دنیای واقعی برکه برای آن ها کد منبع وجود دارد، می توان زمان جست و جو و پوشش به طور مستقیم قابل مقایسه نیست.

آناند و همکاران(34) یک رویکرد موسوم به ACTEve را برای تولید رویداد ها برای تست برنامه های اندروید ارایه کرده اند که کد منبع آن ها موجود است. تاکید اصلی بر پوشش شاخه ها ضمن اجتناب از مسئله انفجار مسیر است. ACTEve تولید ورودی های ازمایشی برای 5 متن باز کوچک در 0.3-2.7 می کند. به طور مشابه، جانسن و همکاران(5) برنامه ای را برای مشتق کردن توالی های منجر به حالت هدف در برتامه اندروید ارایه کردند. تاکید اصلی ما مربوط به جست و جوی سنسور محور و GUI برای برنامه های محبوب به حای تاکید بر مسیر های خاص است. استفاده از اجرای کانکولیک به ما امکان افزایش پوشش را می دهد.با این حال نیازمند یک موتور جست و چو گر نمادین برای کار بر روی APK از برنامه های دنیای واقعی است.

مانکی(15) از ازمایش مطلوبیت SDK برای توالی یابی رویداد های قطعی و تصادفی بر روی برنامه استفاده کردند. رویداد های تصادفی برای تست تنش و تست فاز موثر است و ولی برای جست و جوی سیستمی مناسب نیست: رویداد های قطعی باید برنامه نویسی شوند در حالی که در جست و جوی سیستمی به صورت خودکار است. مانکی رانر(24) یکAPI اندروید SDK را ارایه کردند که به برنانویسان برای نوشتن برنامه فیتون کمک می کند. روبوتیوم(18) یک چارچوب ازمایشی برای اندروید است که از تست وایت باکس و بلک باکس استفاده می کند. روبوتیوم موجب تسهیل تعامل با اجزای GUI می شود که از این روی برای کشف رویداد مربوطه استفاده شده و موارد ازمایشی برای اجرای برنامه ها به کار گرفته می شود.

تریود(20) یک ابزار ازمایشی و پاسخ بروی روبوتیوم است که برای استخراچ گوه GUI ، ثبت رویداد های GUI و اسکریپت استفاده می کند. از بخش های تریود در رویکرد استفاده می کنیم. تریود برای جست و جوی هدفمند و عمق اول استفاده می کند زیرا مستلزم اسکریپت های ورودی برای انجام عناصر GUI است. به علاوه، در شکل غیر اصلاحی ، تیود دارای سربار عملکردی است که موجب کاهش سرعت جست و جو می شود و از این روی امکان کاهش سربار عملکرد می شود.

TEMA(41) مجموعه ای از ابزار های تست مدل محور برای اندروید است. GUI تشکیل یک ماشین حالت داده و رویداد های GUI به صورت لغات کلیدی در نظر کرفته می شودو در این چارچوب، اسکریپت های تست را می توان طراحی و اجرا کرد؟ برعکس، می توان مدل را به طور استاتیک یا دینامیک با ساختار های ازمایشی مدل سازی کرد.

آندروید ریپر(9) یک ابزار تست دینامیک و استاتیک مبتنی بر GUI برای اندروید است/ و از رویکرد مبتنی بر حالت برای تحلیل رویداد های GUI استفاده کرده و برای اتومات سازی تست با هر نوع نمونه آزمایشی به کار می رود. اندروید ریپر وضعیت برنامه را حفظ می کند به طور یکه برنامه می تواند یکئ GUI خاص باشد. یک ورودی می تواند موجب تغییر در حالت شود و کاربران قادر به نوشتن برنامه های آزمایشی بر اساس وظایف خاص هستند. رویکرد تنها روی شبیه ساز اندروید کار می کند و از این روی این از رویداد های حساس به سنسور نظیر برنامه دنیای واقعی پیروی نمی کند.

چندین ابزار تجاری یک سری کارکرد های مناسب را برای این منظور ارایه می کنند. تستدرویید(26) قادر به ثبت و اجرای ازمایش بر روی ابزار های مختلف باشد. رانورگکس(27) یک چارچوب اتوماسیون ازمایشی است. اگپلنت(29) موجب تسهیل برنامه های خودکار برای تست برنامه اندروید می شود. چارچوب تست برنامه خود کار قادر به خود کار سازی فرایند تست برنامه های اندروید در ابزار های مختلف است.

در نهایت، طیف وسیعی از ابزار های تحلیل استاتیک و دینامیک برای اندروید وجود دارد و این ابزار ها با کار ما ارتباط نزدیک دارند. ما هم چنین از تحلیل استاتیک برای ساخت STAG استفاده می کنیم ولی هدف نهایی ما انجام یک تحلیل استاتیک نیست. با این حال، یک سری پایش ها می توانند مطابق با سناریوی تحلیل پویایی بانشند که پوشش زیادی را ارایه می کند.

نتیجه گیری

ما A^3E را به صورت رویکرد و ابزاری در نظر کرفته ایم که امکان جست و جوی خود کار برنامه های اندروییدی را می دهد و به این ترتیب نتایج مطالعه نشان داد که کاربران قادر به جست و جوی مجموعه کوچکی از ویژگی های تعاملی با برنامه های اندروید هستند. ما الگوریتم جست و جوی هدفمند را ارایه می کنیم که یک روش جدیدی است که تحلیل استاتیک را اهرم کرده و موجب تسهیل جست و جوی موثر فعالیت های اندروییدی می

شود. رویکرد های قبلی برای جست و جوی خودکار برنامه های اندروییدی دارای برنامه هایی در شبیه ساز بوده و یا این که بر برنامه های کوچک تر یمتمرکز هستند که کد منبع آن ها قابل دسترس است. برای حل این مسائل، ما از رویکرد A^3E استفاده کردیم که از طریق آن می توان به جست و جوی سیستمی برنامه های اجرا شده برر روی گوشی های اقعی بدون دسترسی به کد منبع پرداخت. اطلاعات کلیدی مربوط به رویکرد ما شامل استفاده از تحلیل جریان داده ها، استاتیک و سبک قدیمی بر روی بایت کد های برنامه به شکلی جدید است که می تواند طیف وسیعی از فعالیت ها را در نظر بگیرد. ما از این نمودار برای توسعه یک راهبرد جست و جویی موسوم به کشف و جست و جوی هدفمند استفاده می کنیم که امکان جست و جوی مستقیم از جمله فعالیت هایی که طی استفاده طبیعی سخت هستند را می دهد. ما از راهبرد موسوم به جست و جوی اول عمق نیز بهره می بریم که از عملیات کاربر برای فعالیت های اکتشافی و نیز اجزای سازنده آن به طور کند تر ولی سیستمی الگو برداری کردیم همه این الگوریتم ها بر روی 25 اندروییید اجرا شدند و نشان داده شد که روش های ما به افزایش پوشش منجر می شود/. رویکرد می تواند یک مبنایی برای طیف وسیعی از تحلیل های پویا و نیز کار های ازمایشی باشد.