

# TO-DOs

- TODO n°1 p.2 : Find a title
- TODO n°2 p.2 : Find a title
- TODO n°3 p.2 : Date
- TODO n°4 p.2 : Compose a Jury
- TODO n°5 p.5 : Acknowledge people
- TODO n°6 p.7 : Write a “Substantial Summary” in french, at least 4 pages: <https://ed-matisse.doctorat-bretagne.fr/fr/soutenance-de-these#p-151>
- TODO n°7 p.23 : Write an introduction
- TODO n°8 p.25 : Present your field background
- TODO n°9 p.27 : Do the State of the Art
- TODO n°10 p.29 : typstify RASTA paper
- TODO n°11 p.29 : Format numbers
- TODO n°12 p.36 : alt text for figure rasta-exit / rasta-exit-drebin
- TODO n°13 p.36 : We discuss further errors for which we have information in the logs in Section.
- TODO n°14 p.37 : Alt text for fig rasta-decorelation-size
- TODO n°15 p.38 : Alt text for fig rasta-decorelation-size
- TODO n°16 p.38 : Alt text for fig rasta-decorelation-min-sdk
- TODO n°17 p.45 : Conclude
- TODO n°18 p.54 : Find a title
- TODO n°19 p.54 : Find a title
- TODO n°20 p.54 : More Keywords



CentraleSupélec

# THÈSE DE DOCTORAT DE

CENTRALSUPÉLEC

ÉCOLE DOCTORALE N° 601

*Mathématiques, Télécommunications, Informatique,  
Signal, Systèmes, Électronique*

Spécialité : *Informatique*

Par

**Jean-Marie MINEAU**

**TODO 1** ► *Find a title* ◀

**TODO 2** ► *Find a title* ◀

Thèse présentée et soutenue à Rennes, le **TODO 3** ► *Date* ◀

Unité de recherche : IRISA

## Composition du jury :

Présidente : Alice

Rapporteurs : Bob

Eve

Examinatrice : Mallory

Dir. de thèse : Jean-François Lalande

Valérie Viet Triem Tong

Professeur des Universités

Professeure

CentraleSupélec

CentraleSupélec

**TODO 4** ► *Compose a Jury* ◀

---

---

# ACKNOWLEDGEMENTS

---

**TODO 5** ► *Acknowledge people* ◀

3 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 4 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequae doleamus animo,  
 5 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 6 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 7 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
 8 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
 9 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
 10 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
 11 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
 12 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
 13 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis  
 14 mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc  
 15 sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita  
 16 prorsus existimo, neque eum Torquatum, qui hoc primum cognomen invenerit, aut torquem  
 17 illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt  
 18 vitae sine metu degendae praesidia firmissima. – Filium morte multavit. – Si sine causa, nollem  
 19 me ab eo delectari, quod ista Platonis, Aristoteli, Theophrasti orationis ornamenta neglexerit.  
 20 Nam illud quidem physici, credere aliquid esse minimum, quod profecto numquam putavisset,  
 21 si a Polyano, familiari suo, geometrica discere maluisset quam illum etiam ipsum dedocere.  
 22 Sol Democrito magnus videtur, quippe homini erudito in geometriaque perfecto, huic pedalis  
 23 fortasse; tantum enim esse omnino in nostris poetis aut inertissimae segnitiae est aut fastidii  
 24 delicatissimi. Mihi quidem videtur, inermis ac nudus est. Tollit definitiones, nihil de dividendo ac  
 25 partiendo docet, non quo ignorare vos arbitrer, sed ut ratione et via procedat oratio. Quaerimus  
 26 igitur, quid sit extremum et ultimum bonorum, quod omnium philosophorum sententia tale  
 27 debet esse, ut eius magnitudinem celeritas, diuturnitatem allevatio consoletur. Ad ea cum  
 28 accedit, ut neque divinum numen horreat nec praeteritas voluptates effluere patiatur earumque  
 29 assidua recordatione laetetur, quid est, quod huc possit, quod melius sit, migrare de vita. His  
 30 rebus instructus semper est in voluptate esse aut in armatum hostem impetum fecisse aut in  
 31 poetis evolvendis, ut ego et Triarius te hortatore facimus, consumeret, in quibus hoc primum  
 32 est in quo admirer, cur in gravissimis rebus non delectet eos sermo patrius, cum.

---

# RÉSUMÉ EN FRANÇAIS

34 **TODO 6** ► *Write a “Substantial Summary” in french, at least 4 pages: <https://ed->*  
 35 *matisse.doctorat-bretagne.fr/fr/soutenance-de-these#p-151* ◀

36 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 37 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequale doleamus animo,  
 38 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 39 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 40 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
 41 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
 42 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
 43 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
 44 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
 45 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
 46 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis  
 47 mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc  
 48 sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita  
 49 prorsus existimo, neque eum Torquatum, qui hoc primum cognomen invenerit, aut torquem  
 50 illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt  
 51 vitae sine metu degendae praesidia firmissima. – Filium morte multavit. – Si sine causa,  
 52 nollem me ab eo delectari, quod ista Platonis, Aristoteli, Theophrasti orationis ornamenta  
 53 neglexerit. Nam illud quidem physici, credere aliquid esse minimum, quod profecto numquam  
 54 putavisset, si a Polyaeno, familiari suo, geometrica discere maluisset quam illum etiam ipsum  
 55 dedocere. Sol Democrito magnus videtur, quippe homini erudito in geometriaque perfecto, huic  
 56 pedalis fortasse; tantum enim esse omnino in nostris poetis aut inertissimae segnitiae est aut  
 57 fastidii delicatissimi. Mihi quidem videtur, inermis ac nudus est. Tollit definitiones, nihil de  
 58 dividendo ac partiendo docet, non quo ignorare vos arbitrer, sed ut ratione et via procedat  
 59 oratio. Quaerimus igitur, quid sit extremum et ultimum bonorum, quod omnium philosophorum  
 60 sententia tale debet esse, ut eius magnitudinem celeritas, diuturnitatem allevatio consoletur.  
 61 Ad ea cum accedit, ut neque divinum numen horreat nec praeteritas voluptates effluere patiatur  
 62 earumque assidua recordatione laetetur, quid est, quod huc possit, quod melius sit, migrare  
 63 de vita. His rebus instructus semper est in voluptate esse aut in armatum hostem impetum  
 64 fecisse aut in poetis evolvendis, ut ego et Triarius te hortatore facimus, consumeret, in quibus  
 65 hoc primum est in quo admirer, cur in gravissimis rebus non delectet eos sermo patrius, cum  
 66 idem fabellas Latinas ad verbum e Graecis expressas non inviti legant. Quis enim tam inimicus

---

67 paene nomini Romano est, qui Ennii Medeam aut Antiopam Pacuvii spernat aut reiciat, quod  
68 se isdem Euripidis fabulis delectari dicat, Latinas litteras oderit? Synephebos ego, inquit, potius  
69 Caecilii aut Andriam Terentii quam utramque Menandri legam? A quibus tantum dissentio,  
70 ut, cum Sophocles vel optime scripserit Electram, tamen male conversam Atilii mihi legendam  
71 putem, de quo Lucilius: 'ferreum scriptorem', verum, opinor, scriptorem tamen, ut legendus sit.  
72 Rudem enim esse omnino in nostris poetis aut inertissimae segnitiae est aut in dolore. Omnis  
73 autem privatione doloris putat Epicurus terminari summam voluptatem, ut postea variari  
74 voluptas distinguique possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
75 audiebam facete et urbane Stoicos irridente, statua est in voluptate aut a voluptate discedere.  
76 Nam cum ignoratione rerum bonarum et malarum maxime hominum vita vexetur, ob eumque  
77 errorem et voluptatibus maximis saepe priventur et durissimis animi doloribus torqueantur,  
78 sapientia est adhibenda, quae et terroribus cupiditatibusque detractis et omnium falsarum  
79 opinionum temeritate derepta certissimam se nobis ducem praebeat ad voluptatem. Sapientia  
80 enim est una, quae maestitiam pellat ex animis, quae nos exhorrescere metu non sinat. Qua  
81 praeceptrice in tranquillitate vivi potest omnium cupiditatum ardore restincto. Cupiditates  
82 enim sunt insatiabiles, quae non modo voluptatem esse, verum etiam approbantibus nobis.  
83 Sic enim ab Epicuro reprehensa et correcta permulta. Nunc dicam de voluptate, nihil scilicet  
84 novi, ea tamen, quae te ipsum probaturum esse confidam. Certe, inquam, pertinax non ero  
85 tibi, si mihi probabis ea, quae dicta sunt ab iis quos probamus, eisque nostrum iudicium  
86 et nostrum scribendi ordinem adiungimus, quid habent, cur Graeca anteponant iis, quae et a  
87 formidinum terrore vindicet et ipsius fortunae modice ferre doceat iniurias et omnis monstret  
88 vias, quae ad amicos pertinerent, negarent esse per se ipsam causam non multo maiores esse  
89 et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum hic tenetur a  
90 sapiente delectus, ut aut voluptates omittantur maiorum voluptatum adipiscendarum causa aut  
91 dolores suscipiantur maiorum dolorum effugiendorum gratia. Sed de clarorum hominum factis  
92 illustribus et gloriosis satis hoc loco dictum sit. Erit enim iam de omnium virtutum cursu ad  
93 voluptatem proprius disserendi locus. Nunc autem explicabo, voluptas ipsa quae qualisque sit, ut  
94 tollatur error omnis imperitorum intellegaturque ea, quae voluptaria, delicata, mollis habeatur  
95 disciplina, quam gravis, quam continens, quam severa sit. Non enim hanc solam sequimur,  
96 quae suavitate aliqua naturam ipsam movet et cum iucunditate quadam percipitur sensibus,  
97 sed maximam voluptatem illam habemus, quae percipitur omni dolore careret, non modo non  
98 repugnantibus, verum etiam approbantibus nobis. Sic enim ab Epicuro sapiens semper beatus  
99 inducitur: finitas habet cupiditates, neglegit mortem, de diis immortalibus sine ullo metu vera  
100 sentit, non dubitat, si ita res se habeat. Nam si concederetur, etiamsi ad corpus referri, nec  
101 ob eam causam non fuisse. – Torquem detraxit hosti. – Et quidem se textit, ne interiret.  
102 – At magnum periculum adiit. – In oculis quidem exercitus. – Quid ex eo est consecutus?  
103 – Laudem et caritatem, quae sunt vitae sine metu degendae praesidia firmissima. – Filium  
104 morte multavit. – Si sine causa, nollem me ab eo et gravissimas res consilio ipsius et ratione



105 administrari neque maiorem voluptatem ex infinito tempore aetatis percipi posse, quam ex hoc  
106 facillime perspicitur potest: Constituamus aliquem magnis, multis, perpetuis fruentem et animo et  
107 attento intuemur, tum fit ut aegritudo sequatur, si illa mala sint, laetitia, si bona. O praeclaram  
108 beate vivendi et apertam et simplicem et directam viam! Cum enim certe nihil homini possit  
109 melius esse quam Graecam. Quando enim nobis, vel dicam aut oratoribus bonis aut poetis,  
110 postea quidem quam fuit quem imitarentur, ullus orationis vel copiosae vel elegantis ornatus  
111 defuit? Ego vero, quoniam forensibus operis, laboribus, periculis non deseruisse mihi videor  
112 praesidium, in quo a nobis sic intelleges eitam, ut ab ipsis, qui eam disciplinam probant, non  
113 soleat accuratius explicari; verum enim invenire volumus, non tamquam adversarium aliquem  
114 convincere. Accurate autem quondam a L. Torquato, homine omni doctrina erudito, defensa  
115 est Epicuri sententia de voluptate, nihil scilicet novi, ea tamen, quae te ipsum probaturum esse  
116 confidam. Certe, inquam, pertinax non ero tibi, si mihi probabis ea, quae praeterierunt, acri  
117 animo et corpore voluptatibus nullo dolore nec impediante nec inpendente, quem tandem hoc  
118 statu praestabiliorem aut magis expetendum possimus dicere? Inesse enim necesse est effici, ut  
119 sapiens solum amputata circumcisaque inanitate omni et errore naturae finibus contentus sine  
120 aegritudine possit et sine metu degendae praesidia firmissima. – Filium morte multavit. – Si sine  
121 causa, nollem me ab eo et gravissimas res consilio ipsius et ratione administrari neque maiorem  
122 voluptatem ex infinito tempore aetatis percipi posse, quam ex hoc facillime perspicitur potest:  
123 Constituamus aliquem magnis, multis, perpetuis fruentem et animo et corpore voluptatibus  
124 nullo dolore nec impediante nec inpendente, quem tandem hoc statu praestabiliorem aut magis  
125 expetendum possimus dicere? Inesse enim necesse est aut in liberos atque in sanguinem suum  
126 tam crudelis fuisse, nihil ut de omni virtute sit dictum. Sed similia fere dici possunt. Ut enim  
127 virtutes, de quibus neque depravate iudicant neque corrupte, nonne ei maximam gratiam habere  
128 debemus, qui hac exaudita quasi voce naturae sic eam firme graviterque comprehenderit, ut  
129 omnes bene sanos ad iustitiam, aequitatem, fidem, neque homini infanti aut inpotenti iniuste  
130 facta conducunt, qui nec facile efficere possit, quod melius sit, accedere? Statue contra aliquem  
131 confectum tantis animi corporisque doloribus, quanti in hominem maximi cadere possunt, nulla  
132 spe proposita fore levius aliquando, nulla praeterea neque praesenti nec expectata voluptate,  
133 quid eo miserior dici aut fingi potest? Quodsi vita doloribus referta maxime fugienda est,  
134 summum bonum consequamur? Clamat Epicurus, is quem vos nimis voluptatibus esse deditum  
135 dicitis; non posse reperiri. Quapropter si ea, quae senserit ille, tibi non vera videantur. Vide,  
136 quantum, inquam, fallare, Torquate. Oratio me istius philosophi non offendit; nam et praeterita  
137 grate meminit et praesentibus ita potitur, ut animadvertat quanta sint ea quamque iucunda,  
138 neque pendet ex futuris, sed expectat illa, fruitur praesentibus ab iisque vitiis, quae paulo  
139 ante collegi, abest plurimum et, cum stultorum vitam cum sua comparat, magna afficitur  
140 voluptate. Dolores autem si qui e nostris aliter existimant, quos quidem video minime esse  
141 deterritum. Quae cum dixisset, Explicavi, inquit, sententiam meam, et eo quidem consilio,  
142 tuum iudicium ut cognoscerem, quoniam mihi ea facultas, ut id meo arbitratu facerem, ante

---

143 hoc tempus numquam est dici. Graece ergo praetor Athenis, id quod maluisti, te, cum ad me  
144 in Cumanum salutandi causa uterque venisset, pauca primo inter nos ea, quae audiebamus,  
145 conferebamus, neque erat umquam controversia, quid ego intellegerem, sed quid probarem. Quid  
146 igitur est? Inquit; audire enim cupio, quid non probes. Principio, inquam, in physicis, quibus  
147 maxime gloriatur, primum totus est alienus. Democritea dicit perpauca mutans, sed ita, ut  
148 ea, quae hoc non minus declarant, sed videntur leviora, veniamus. Quid tibi, Torquate, quid  
149 huic Triario litterae, quid historiae cognitioque rerum, quid poetarum evolutio, quid tanta tot  
150 versuum memoria voluptatis affert? Nec mihi illud dixeris: 'Haec enim ipsa mihi sunt voluptati,  
151 et erant illa Torquatis.' Numquam hoc ita defendit Epicurus neque Metrodorus aut quisquam  
152 eorum, qui aut saperet aliquid aut ista didicisset. Et quod adest sentire possumus, animo autem  
153 et praeterita et futura. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen  
154 permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur.  
155 Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguere possit, augeri  
156 amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos  
157 irridente, statua est in quo admirer, cur in gravissimis rebus non delectet eos sermo patrius, cum  
158 idem fabellas Latinas ad verbum e Graecis expressas non inviti legant. Quis enim tam inimicus  
159 paene nomini Romano est, qui alienae modum statuatur industriae? Nam ut Terentianus Chremes  
160 non inhumanus, qui novum vicinum non vult 'fodere aut arare aut aliquid ferre denique' – non  
161 enim illum ab industria, sed ab inliberali labore deterret –, sic isti curiosi, quos offendit noster  
162 minime nobis iniucundus labor. Iis igitur est difficilius satis facere, qui se dicant in Graecis  
163 legendis operam malle consumere. Postremo aliquos futuros suspicor, qui me ad alias litteras  
164 vocent, genus hoc scribendi, etsi sit elegans, personae tamen et dignitatis esse negent. Contra  
165 quos omnis dicendum breviter existimo. Quamquam philosophiae quidem vituperatoribus satis  
166 responsum est eo libro, quo a populo Romano locatus sum, debeo profecto, quantumcumque  
167 possum, in eo quoque elaborare, ut sint illa vendibilia, haec uberiora certe sunt. Quamquam  
168 id quidem facio provocatus gratissimo mihi libro, quem ad modum eae semper voluptatibus  
169 inhaerent, eadem de amicitia dicenda sunt. Praeclare enim Epicurus his paene verbis: 'Eadem',  
170 inquit, 'scientia confirmavit animum, ne quod aut sempiternum aut diuturnum timeret malum,  
171 quae perspexit in hoc ipso vitae spatio amicitiae praesidium esse firmissimum.' Sunt autem  
172 quidam e nostris, et scribentur fortasse plura, si vita suppetet; et tamen, qui diligenter haec,  
173 quae de philosophia litteris mandamus, legere assueverit, iudicabit nulla ad legendum his esse  
174 potiora. Quid est enim in vita tantopere quaerendum quam cum omnia in philosophia, tum id,  
175 quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus.  
176 Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet, ut  
177 et adversa quasi perpetua oblivione obruamus et secunda iucunde ac suaviter meminerimus. Sed  
178 cum ea, quae dicta sunt ab iis quos probamus, eisque nostrum iudicium et nostrum scribendi  
179 ordinem adiungimus, quid habent, cur Graeca anteponant iis, quae recordamur. Stulti autem  
180 malorum memoria torquentur, sapientes bona praeterita grata recordatione renovata delectant.

181 Est autem situm in nobis ut et adversa quasi perpetua oblivione obruamus et secunda iucunde  
182 ac suaviter meminerimus. Sed cum ea, quae praeterierunt, acri animo et attento intuemur, tum  
183 fit ut aegritudo sequatur, si illa mala sint, laetitia, si bona. O praeclaram beate vivendi et  
184 apertam et simplicem et directam viam! Cum enim certe nihil homini possit.

---

# TABLE OF CONTENTS

---

186	1	Introduction .....	2
187	2	Background .....	4
188	2.1	Something .....	4
189	2.2	Something Else .....	5
190	3	Related Work .....	6
191	4	RASTA .....	8
192	4.1	Related Work .....	8
193	4.1.1	Application Datasets .....	8
194	4.1.2	Static Analysis Tools Reusability .....	9
195	4.2	Methodology .....	10
196	4.2.1	Collecting Tools .....	10
197	4.2.2	Source Code Selection and Building Process .....	12
198	4.2.3	Runtime Conditions .....	13
199	4.2.4	Dataset .....	14
200	4.3	Experiments .....	15
201	4.3.1	<b>RQ1:</b> Re-Usability Evaluation .....	15
202	4.3.2	RQ2: Size, SDK and Date Influence .....	16
203	4.3.3	RQ3: Malware vs Goodware .....	18
204	5	Contribution 2 .....	20
205	6	Contribution n .....	22
206	7	Conclusion .....	24
207		Bibliography .....	26

---

# INDEX OF FIGURES

---

209	Figure 1: A circle .....	2
210	Figure 2: Methodology overview .....	13
211	Figure 3: Exit status for the Drebin dataset .....	15
212	Figure 4: Exit status for the Rasta dataset .....	15
213	Figure 5: Finishing rate by bytecode size for APK detected in 2022 .....	17
214	Figure 8: Finishing rate by discovery year with a bytecode size $\in [4.08, 5.2]$ MB .....	17
215	Figure 11: Finishing rate by min SDK with a bytecode size $\in [4.08, 5.2]$ MB .....	18

---



# INDEX OF TABLES

---

217	Table 1: A tic tac toe game .....	4
218	Table 2: Considered tools[18]: availability and usage reliability .....	10
219	Table 3: Selected tools, forks, selected commits and running environment .....	12
220	Table 4: DEX size and Finishing Rate (FR) per decile .....	18

---

# INDEX OF LISTINGS

---

222	Listing 1: Some code .....	6
-----	----------------------------	---

---

223

# LIST OF ACRONYMS AND NOTATIONS

---

224

225

Acronyms	Meanings
TL;DR	Too long; didn't read

---

# INTRODUCTION

---

228 **TODO 7** ► *Write an introduction* ◀

229 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 230 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequi doleamus animo,  
 231 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 232 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 233 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
 234 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
 235 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
 236 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
 237 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
 238 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
 239 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi  
 240 Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam  
 241 insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus  
 242 existimo, neque eum Torquatum, qui hoc primus cognomen invenerit, aut torquem illum hosti  
 243 detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

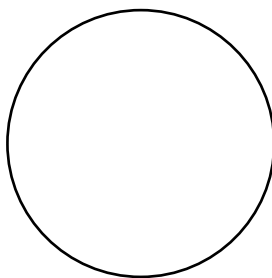


Figure 1: A circle

244 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 245 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequi doleamus animo,  
 246 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 247 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 248 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre

---

249 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
250 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
251 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
252 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
253 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
254 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi  
255 Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam  
256 insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus  
257 existimo, neque eum Torquatum, qui hoc primum cognomen invenerit, aut torquem illum hosti  
258 detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.



# BACKGROUND

261 TODO 8 ► *Present your field background* ◀

262 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
263 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequi doleamus animo,  
264 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
265 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
266 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
267 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
268 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
269 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
270 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
271 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
272 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi  
273 Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam  
274 insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus  
275 existimo, neque eum Torquatum, qui hoc primus cognomen invenerit, aut torquem illum hosti  
276 detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

Play		
X		O
	O	
X		X

Table 1: A tic tac toe game

277 **2.1 Something**

278 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
279 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequi doleamus animo,  
280 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
281 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
282 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre

audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis, saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus existimo, neque eum Torquatium, qui hoc primum cognomen invenerit, aut torquem illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

## 2.2 Something Else

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo, cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae. Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis, saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus existimo, neque eum Torquatium, qui hoc primum cognomen invenerit, aut torquem illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

## RELATED WORK

---

311 TODO 9 ► *Do the State of the Art* ◀

312 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
313 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequaleam animo,  
314 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
315 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
316 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
317 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
318 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
319 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
320 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
321 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
322 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi  
323 Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam  
324 insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus  
325 existimo, neque eum Torquatum, qui hoc primus cognomen invenerit, aut torquem illum hosti  
326 detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt vitae.

```
for _ in range(10):  
    print("Hello Void")
```

Listing 1: Some code

---

# RASTA

329 TODO 10 ► *typstify RASTA paper* ◀

330 TODO 11 ► *Format numbers* ◀

## 331 4.1 Related Work

332 We review in this section the past existing datasets provided by the community and the papers  
333 related to static analysis tools reusability.

### 334 4.1.1 Application Datasets

335 Computing if an application contains a possible information flow is an example of a static  
336 analysis goal. Some datasets have been built especially for evaluating tools that are comput-  
337 ing information flows inside Android applications. One of the first well known dataset is  
338 DroidBench, that was released with the tool Flowdroid[2]. Later, the dataset ICC-Bench was  
339 introduced with the tool Amandroid[33] to complement DroidBench by introducing applica-  
340 tions using Inter-Component data flows. These datasets contain carefully crafted applications  
341 containing flows that the tools should be able to detect. These hand-crafted applications can  
342 also be used for testing purposes or to detect any regression when the software code evolves.  
343 Contrary to real world applications, the behavior of these hand-crafted applications is known  
344 in advance, thus providing the ground truth that the tools try to compute. However, these  
345 datasets are not representative of real-world applications[26] and the obtained results can be  
346 misleading.

347 Contrary to DroidBench and ICC-Bench, some approaches use real-world applications. Bosu  
348 *et al.*[5] use DIALDroid to perform a threat analysis of Inter-Application communication  
349 and published DIALDroid-Bench, an associated dataset. Similarly, Luo *et al.* released Taint-  
350 Bench[22] a real-world dataset and the associated recommendations to build such a dataset.  
351 These datasets confirmed that some tools such as Amandroid[33] and Flowdroid[2] are less  
352 efficient on real-world applications. These datasets are useful for carefully spotting missing taint  
353 flows, but contain only a few dozen of applications.

354 Pauck *et al.*[25] used those three datasets to compare Amandroid[33], DIAL-Droid[5], Did-  
355 Fail[15], DroidSafe[12], FlowDroid[2] and IccTA[16] – all these tools will be also compared in this  
356 paper. To perform their comparison, they introduced the AQL (Android App Analysis Query

Language) format. AQL can be used as a common language to describe the computed taint flow as well as the expected result for the datasets. It is interesting to notice that all the tested tools timed out at least once on real-world applications, and that Amandroid[33], DidFail[15], DroidSafe[12], IccTA[16] and ApkCombiner[17] (a tool used to combine applications) all failed to run on applications built for Android API 26. These results suggest that a more thorough study of the link between application characteristics (*e.g.*, date, size) should be conducted. Luo *et al.*[22] used the framework introduced by Pauck *et al.* to compare Amandroid[33] and Flowdroid[2] on DroidBench and their own dataset TaintBench, composed of real-world android malware. They found out that those tools have a low recall on real-world malware, and are thus over adapted to micro-datasets. Unfortunately, because AQL is only focused on taint flows, we cannot use it to evaluate tools performing more generic analysis.

#### 4.1.2 Static Analysis Tools Reusability

Several papers have reviewed Android analysis tools produced by researchers. Li *et al.*[18] published a systematic literature review for Android static analysis before May 2015. They analyzed 92 publications and classified them by goal, method used to solve the problem and underlying technical solution for handling the bytecode when performing the static analysis. In particular, they listed 27 approaches with an open-source implementation available. Nevertheless, experiments to evaluate the reusability of the pointed out software were not performed. We believe that the effort of reviewing the literature for making a comprehensive overview of available approaches should be pushed further: an existing published approach with a software that cannot be used for technical reasons endanger both the reproducibility and reusability of research.

A first work about quantifying the reusability of static analysis tools was proposed by Reaves *et al.*[28]. Seven Android analysis tools (Amandroid[33], AppAudit[35], DroidSafe[12], Epicc[24], FlowDroid[2], MalloDroid[9] and TaintDroid[8]) were selected to check if they were still readily usable. For each tool, both the usability and results of the tool were evaluated by asking auditors to install and use it on DroidBench and 16 real world applications. The auditors reported that most of the tools require a significant amount of time to setup, often due to dependencies issues and operating system incompatibilities. Reaves *et al.* propose to solve these issues by distributing a Virtual Machine with a functional build of the tool in addition to the source code. Regrettably, these Virtual Machines were not made available, preventing future researchers to take advantage of the work done by the auditors. Reaves *et al.* also report that real world applications are more challenging to analyze, with tools having lower results, taking more time and memory to run, sometimes to the point of not being able to run the analysis. We will confirm and expand this result in this paper with a larger dataset than only 16 real-world applications.

## 4.2 Methodology

### 4.2.1 Collecting Tools

Tool	Availability			Repo type	Decision	Comments
	Bin	Src	Doc			
A3E [3] (2013)	–	✓	✓	github	✗	Hybrid tool (static/dynamic)
A5 [32] (2014)	–	✓	✗	github	✗	Hybrid tool (static/dynamic)
Adagio [10] (2013)	–	✓	✓	github	✓	
Amandroid [33] (2014)	✓	✓	✓	github	✓	
Anadroid [19] (2013)	✗	✓	✓	github	✓	
Androguard [7] (2011)	–	✓	✓	github	✓	
Android-app-analysis [11] (2015)	✗	✓	✓	google	✗	Hybrid tool (static/dynamic)
Apparecium [31] (2015)	✓	✓	✗	github	✓	
BlueSeal [30] (2014)	✗	✓	○	github	✓	
Choi <i>et al.</i> [6] (2014)	✗	✓	○	github	✗	Works on source files only
DIALDroid [5] (2017)	✓	✓	✓	github	✓	
DidFail [15] (2014)	✓	✓	○	bitbucket	✓	
DroidSafe [12] (2015)	✗	✓	✓	github	✓	
Flowdroid [2] (2014)	✓	✓	✓	github	✓	
Gator [29, 36] (2014, 2015)	✗	✓	✓	edu	✓	
IC3 [23] (2015)	✓	✓	○	github	✓	
IccTA [16] (2015)	✓	✓	✓	github	✓	
Lotrack [20] (2014)	✗	✓	✗	github	○	Authors ack. a partial doc.
MalloDroid [9] (2012)	–	✓	✓	github	✓	
PerfChecker [21] (2014)	✗	✗	○	request	✓	Binary obtained from authors
Poeplau <i>et al.</i> [27] (2014)	ko	○	✗	github	✗	Related to Android hardening
Redexer [14] (2012)	✗	✓	✓	github	✓	
SAAF [13] (2013)	✓	✓	✓	github	✓	
StaDynA [37] (2015)	ko	✓	✓	request	✗	Hybrid tool (static/dynamic)
Thresher [4] (2013)	✗	✓	✓	github	○	Not built with author's help
Wognsen <i>et al.</i> [34] (2014)	–	✓	✗	bitbucket	✓	

**binaries, sources:** –: not relevant, ✓: available, ○: partially available, ✗: not provided

**documentation:** ✓✓: excellent, MWE, ✓: few inconsistencies, ○: bad quality, ✗: not available

**decision:** ✓: considered; ○: considered but not built; ✗: out of scope of the study

Table 2: Considered tools[18]: availability and usage reliability

We collected the static analysis tools from[18], plus one additional paper encountered during our review of the state-of-the-art (DidFail[15]). They are listed in Table 2, with the original release date and associated paper. We intentionally limited the collected tools to the ones selected by Li *et al.*[18] for several reasons. First, not using recent tools enables to have a gap of at least 5 years between the publication and the more recent APK files, which enables to measure the reusability of previous contribution with a reasonable gap of time. Second, collecting new tools would require to describe these tools in depth, similarly to what have been performed by Li *et al.*[18], which is not the primary goal of this paper. Additionally, selection criteria such as the

publication venue or number of citations would be necessary to select a subset of tools, which would require an additional methodology. These possible contributions are left for future work.

Some tools use hybrid analysis (both static and dynamic): A3E[3], A5[32], Android-app-analysis[11], StaDynA[37]. They have been excluded from this paper. We manually searched the tool repository when the website mentioned in the paper is no longer available (*e.g.*, when the repository have been migrated from Google code to GitHub) and for each tool we searched for:

- an optional binary version of the tool that would be usable as a fall back (if the sources cannot be compiled for any reason);
- the source code of the tool;
- the documentation for building and using the tool with a MWE (Minimum Working Example).

In Table 2 we rated the quality of these artifacts with “✓” when available but may have inconsistencies, a “○” when too much inconsistencies (inaccurate remarks about the sources, dead links or missing parts) have been found, a “✗” when no documentation have been found, and a double “✓✓” for the documentation when it covers all our expectations (building process, usage, MWE). Results show that documentation is often missing or very poor (*e.g.*, Lotrack), which makes the rebuild process very complex and the first analysis of a MWE.

We finally excluded Choi *et al.*[6] as their tool works on the sources of Android applications, and Poeplau *et al.*[27] that focus on Android hardening. As a summary, in the end we have 20 tools to compare. Some specificities should be noted. The IC3 tool will be duplicated in our experiments because two versions are available: the original version of the authors and a fork used by other tools like IccTa. For Androguard, the default task consists of unpacking the bytecode, the resources, and the Manifest. Cross-references are also built between methods and classes. Because such a task is relatively simple to perform, we decided to duplicate this tool and ask to Androguard to decompile an APK and create a control flow graph of the code using its decompiler: DAD. We refer to this variant of usage as `androguard_dad`. For Thresher and Lotrack, because these tools cannot be built, we excluded them from experiments.

Finally, starting with 26 tools of Table 2, with the two variations of IC3 and Androguard, we have in total 22 static analysis tools to evaluate in which two tools cannot be built and will be considered as always failing.



## 4.2.2 Source Code Selection and Building Process

Tool	Origin		Alive Forks		Last Commit Date	Authors Reached	Environment Language – OS
	Stars	Alive	Nb	Usable			
Adagio [10]	74	✓	0	✗	2022-11-17	✓	Python – U20.04
Amandroid [33]	161	✗	2	✗	2021-11-10	✓	Scala – U22.04
Anadroid [19]	10	✗	0	✗	2014-06-18	✗	Scala/Java/Python – U22.04
Androguard [7]	4430	✓	3	✗	2023-02-01	✗	Python – Python 3.11 slim
Apparecium [31]	0	✗	1	✗	2014-11-07	✗	Python – U22.04
BlueSeal [30]	0	✗	0	✗	2018-07-04	✓	Java – U14.04
DIALDroid [5]	16	✗	1	✗	2018-04-17	✗	Java – U18.04
DidFail [15]	4	✗			2015-06-17	✓	Java/Python – U12.04
DroidSafe [12]	92	✗	3	✗	2017-04-17	✓	Java/Python – U14.04
Flowdroid [2]	868	✓	1	✗	2023-05-07	✓	Java – U22.04
Gator [29, 36]					2019-09-09	✓	Java/Python – U22.04
IC3 [23]	32	✗	3	✓	2022-12-06	✗	Java – U12.04 / 22.04
IccTA [16]	83	✗	0	✗	2016-02-21	✓	Java – U22.04
Lotrack [20]	5	✗	2	✗	2017-05-11	✓	Java – ?
MalloDroid [9]	64	✗	10	✗	2013-12-30	✗	Python – U16.04
PerfChecker [21]		✗			–	✓	Java – U14.04
Redexer [14]	153	✗	0	✗	2021-05-20	✓	Ocaml/Ruby – U22.04
SAAF [13]	35	✗	5	✗	2015-09-01	✓	Java – U14.04
Thresher [4]	31	✗	1	✗	2014-10-25	✓	Java – U14.04
Wognsen <i>et al.</i> [34]				✗	2022-06-27	✗	Python/Prolog – U22.04

✓: yes, ✗: no, UX.04: Ubuntu X.04

Table 3: Selected tools, forks, selected commits and running environment

In a second step, we explored the best sources to be selected among the possible forks of a tool. We reported some indicators about the explored forks and our decision about the selected one in Table 3. For each source code repository called “Origin”, we reported in Table 3 the number of GitHub stars attributed by users and we mentioned if the project is still alive (✓ in column Alive when a commit exist in the last two years). Then, we analyzed the fork tree of the project. We searched recursively if any forked repository contains a more recent commit than the last one of the branch mentioned in the documentation of the original repository. If such a commit is found (number of such commits are reported in column Alive Forks Nb), we manually looked at the reasons behind this commit and considered if we should prefer this more up-to-date repository instead of the original one (column “Alive Forks Usable”). As reported in Table 3, we excluded all forks, except IC3 for which we selected the fork JordanSamhi/ic3, because they always contain experimental code with no guarantee of stability. For example, a fork of Aparecium contains a port for Windows 7 which does not suggest an improvement of the stability of the tool. For IC3, the fork seems promising: it has been updated to be usable on a recent operating system (Ubuntu 22.04 instead of Ubuntu 12.04 for the original version)

and is used as a dependency by IccTa. We decided to keep these two versions of the tool (IC3 and IC3\_fork) to compare their results.

Then, we self-allocated a maximum of four days for each tool to successfully read and follow the documentation, compile the tool and obtain the expected result when executing an analysis of a MWE. We sent an email to the authors of each tool to confirm that we used the more suitable version of the code, that the command line we used to analyze an application is the most suitable one and, in some cases, requested some help to solve issues in the building process. We reported in Table 3 the authors that answered our request and confirmed our decisions.

From this building phase, several observations can be made. Using a recent operating system, it is almost impossible in a reasonable amount of time to rebuild a tool released years ago. Too many dependencies, even for Java based programs, trigger compilation or execution problems. Thus, if the documentation mentions a specific operating system, we use a Docker image of this OS. Most of the time, tools require additional external components to be fully functional. It could be resources such as the android.jar file for each version of the SDK, a database, additional libraries or tools. Depending of the quality of the documentation, setting up those components can take hours to days. This is why we automatized in a Dockerfile the setup of the environment in which the tool is built and run<sup>1</sup>

### 4.2.3 Runtime Conditions

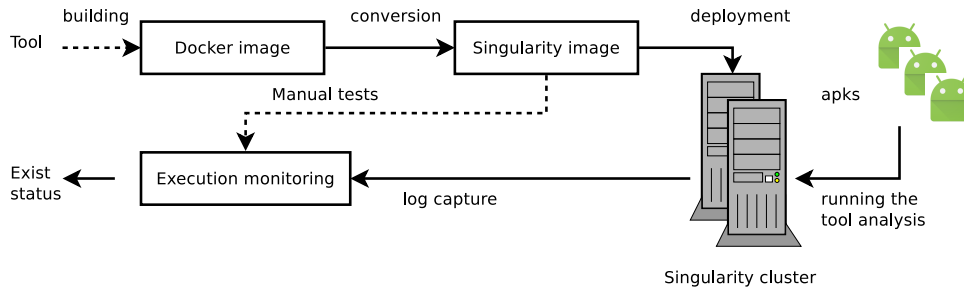


Figure 2: Methodology overview

As shown in Figure 2, before benchmarking the tools, we built and installed them in a Docker containers for facilitating any reuse of other researchers. We converted them into Singularity containers because we had access to such a cluster and because this technology is often used by the HPC community for ensuring the reproducibility of experiments. We performed manual tests using these Singularity images to check:

1. To guarantee reproducibility we published the results, datasets, Dockerfiles and containers: <https://github.com/histausse/rasta>, <https://zenodo.org/records/10144014>, <https://zenodo.org/records/10980349> and on Docker Hub as `histausse/rasta-<toolname>:icsr2024`

- the location where the tool is writing on the disk. For the best performances, we expect the tools to write on a mount point backed by an SSD. Some tools may write data at unexpected locations which required small patches from us.
- the amount of memory allocated to the tool. We checked that the tool could run a MWE with a 64 GB limit of RAM.
- the network connection opened by the tool, if any. We expect the tool not to perform any network operation such as the download of Android SDKs. Thus, we prepared the required files and cached them in the images during the building phase. In a few cases, we patched the tool to disable the download of resources.

A campaign of tests consists in executing the 20 selected tools on all APKs of a dataset. The constraints applied on the clusters are:

- No network connection is authorized in order to limit any execution of malicious software.
- The allocated RAM for a task is `ramlimit`.
- The allocated maximum time is 1 hour.
- The allocated object space / stack space is 64 GB / 16 GB if the tool is a Java based program.

For the disk files, we use a mount point that is stored on a SSD disk, with no particular limit of size. Note that, because the allocation of 64 GB could be insufficient for some tool, we evaluated the results of the tools on 20% of our dataset (described later in Section 4.2.4) with 128 GB of RAM and 64 GB of RAM and checked that the results were similar. With this confirmation, we continued our evaluations with 64 GB of RAM only.

#### 4.2.4 Dataset

We built a dataset named **Rasta** to cover all dates between 2010 to 2023. This dataset is a random extract of Androzoo[1], for which we balanced applications between years and size. For each year and inter-decile range of size in Androzoo, 500 applications have been extracted with an arbitrary proportion of 7% of malware. This ratio has been chosen because it is the ratio of goodware/malware that we observed when performing a raw extract of Androzoo. For checking the maliciousness of an Android application we rely on the VirusTotal detection indicators. If more than 5 antiviruses have flagged the application as malicious, we consider it as a malware. If no antivirus has reported the application as malicious, we consider it as a goodware. Applications in between are dropped.

For computing the release date of an application, we contacted the authors of Androzoo to compute the minimum date between the submission to Androzoo and the first upload to VirusTotal. Such a computation is more reliable than using the DEX date that is often obfuscated when packaging the application.

## 4.3 Experiments

### 4.3.1 RQ1: Re-Usability Evaluation

**TODO 12** ▶ *alt text for figure rasta-exit / rasta-exit-drebin* ◀

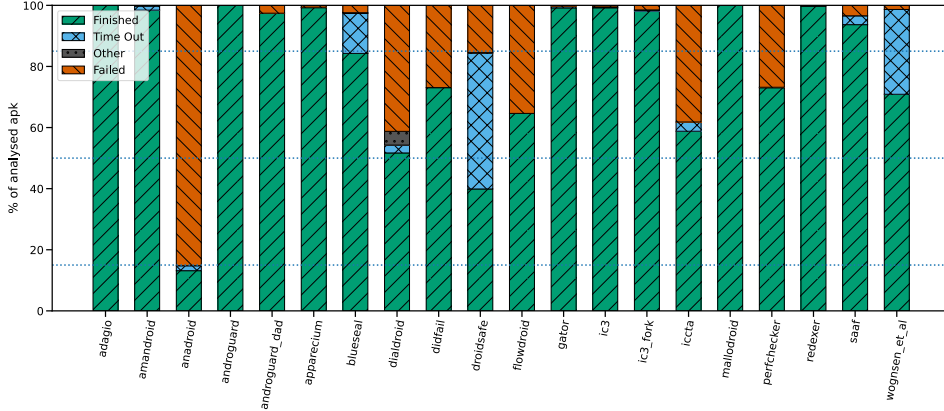


Figure 3: Exit status for the Drebin dataset

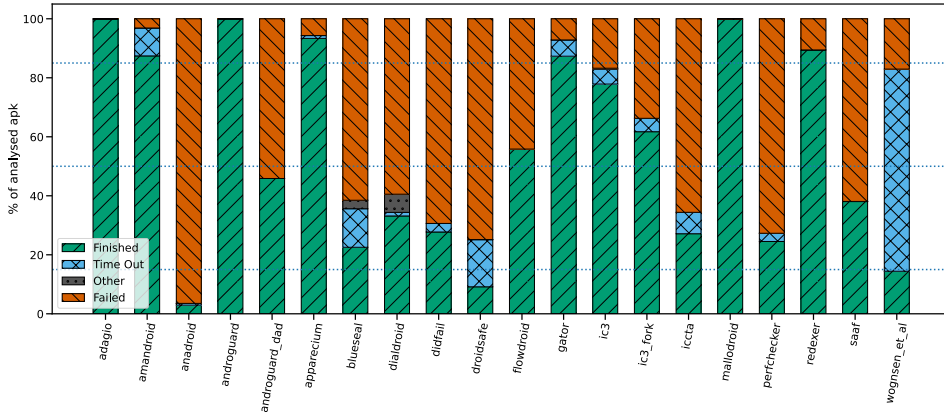


Figure 4: Exit status for the Rasta dataset

FiguresFigure 3 andFigure 4 compare the Drebin and Rasta datasets. They represent the success/failure rate (green/orange) of the tools. We distinguished failure to compute a result from timeout (blue) and crashes of our evaluation framework (in grey, probably due to out of memory kills of the container itself). Because it may be caused by a bug in our own analysis stack, exit status represented in grey (Other) are considered as unknown errors and not as failure of the tool. **TODO 13** ▶ *We discuss further errors for which we have information in the logs in Section.* ◀

Results on the Drebin datasets shows that 11 tools have a high success rate (greater than 85%). The other tools have poor results. The worst, excluding Lotrack and Tresher, is Anadroid with a ratio under 20% of success.

On the Rasta dataset, we observe a global increase of the number of failed status: 12 tools (54.55 %) have a finishing rate below 50%. The tools that have bad results with Drebin are of course bad result on Rasta. Three tools (androguard\_dad, blueséal, saaf) that were performing well (higher than 85%) on Drebin surprisingly fall below the bar of 50% of failure. 7 tools keep a high success rate: Adagio, Amandroid, Androguard, Apparecium, Gator, Mallodroid, Redexer. Regarding IC3, the fork with a simpler build process and support for modern OS has a lower success rate than the original tool.

Two tools should be discussed in particular. Androguard has a high success rate which is not surprising: it used by a lot of tools, including for analyzing application uploaded to the Androzoo repository. Nevertheless, when using Androguard decompiler (DAD) to decompile an APK, it fails more than 50% of the time. This example shows that even a tool that is frequently used can still run into critical failures. Concerning Flowdroid, our results show a very low timeout rate (0.06 %) which was unexpected: in our exchanges, Flowdroid's author were expecting a higher rate of timeout and fewer crashes.

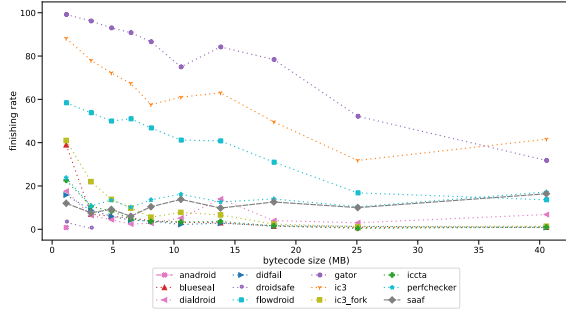
As a summary, the final ratio of successful analysis for the tools that we could run is 54.9 %. When including the two defective tools, this ratio drops to 49.9 %.

**RQ1 answer:** On a recent dataset we consider that result unusable of the tools are unusable. For the tools that we could run, result ratio of analysis are finishing successfully.%(those with less than 50% of successful execution and including the two tools that we were unable to build).

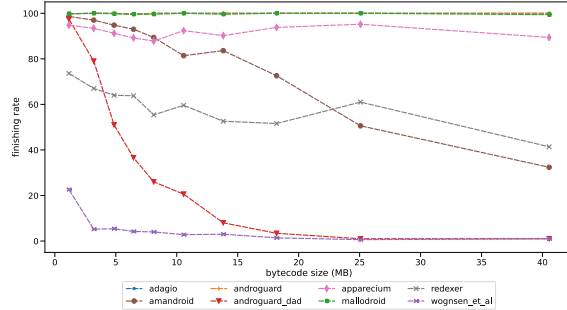
#### 4.3.2 RQ2: Size, SDK and Date Influence

To measure the influence of the date, SDK version and size of applications, we fixed one parameter while varying an other. For the sake of clarity, we separated Java based / non Java based tools.

**TODO 14** ► *Alt text for fig rasta-decorelation-size* ◀



Subfigure 6: Java based tools

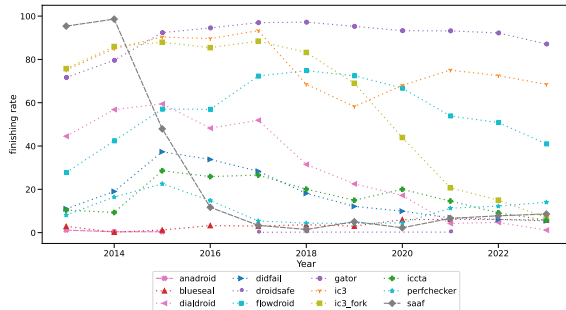


Subfigure 7: Non Java based tools

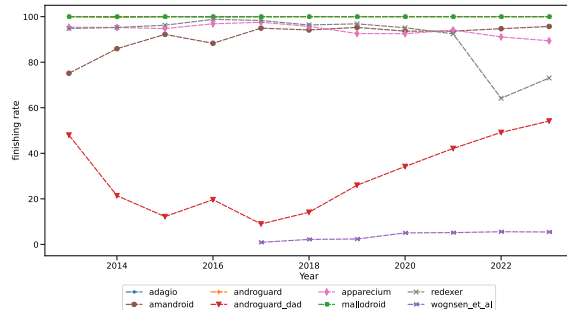
Figure 5: Finishing rate by bytecode size for APK detected in 2022

548 *Fixed application year. (5000 APKs)* We selected the year 2022 which has a good amount of  
 549 representatives for each decile of size in our application dataset. Subfigure 6} (resp. Subfigure 7)  
 550 shows the finishing rate of the tools in function of the size of the bytecode for Java based tools  
 551 (resp. non Java based tools) analyzing applications of 2022. We can observe that all Java based  
 552 tools have a finishing rate decreasing over years. 50% of non Java based tools have the same  
 553 behavior.

554 **TODO 15** ► *Alt text for fig rasta-decorelation-size* ◀



Subfigure 9: Java based tools

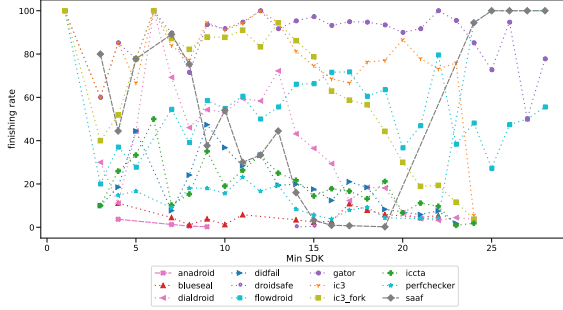


Subfigure 10: Non Java based tools

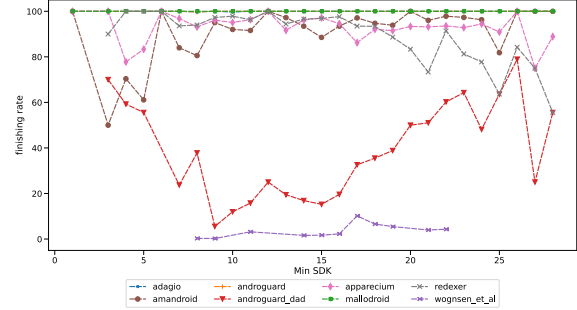
Figure 8: Finishing rate by discovery year with a bytecode size  $\in [4.08, 5.2]$  MB

555 *Fixed application bytecode size. (6252 APKs)* We selected the sixth decile (between 4.08 and  
 556 5.20 MB), which is well represented in a wide number of years. Subfigure 9 (resp. Subfigure 10)  
 557 represents the finishing rate depending of the year at a fixed bytecode size. We observe that 9  
 558 tools over 12 have a finishing rate dropping below 20% for Java based tools, which is not the  
 559 case for non Java based tools.

560 **TODO 16** ► *Alt text for fig rasta-decorelation-min-sdk* ◀



Subfigure 12: Java based tools



Subfigure 13: Non Java based tools

Figure 11: Finishing rate by min SDK with a bytecode size  $\in [4.08, 5.2]$  MB

We performed similar experiments by varying the min SDK and target SDK versions, still with a fixed bytecode size between 4.08 and 5.2 MB, as shown in Subfigure 12 and Subfigure 13. We found that contrary to the target SDK, the min SDK version has an impact on the finishing rate of Java based tools: 8 tools over 12 are below 50% after SDK 16. It is not surprising, as the min SDK is highly correlated to the year.

**RQ2 answer:** The success rate varies based on the size of bytecode and SDK version. The date is also correlated with the success rate for Java based tools only.

### 4.3.3 RQ3: Malware vs Goodware

Decile	Average DEX size (MB)		Finishing Rate: FR		Ratio Size Good/Mal	Ratio FR Good/Mal
	Good	Mal	Good	Mal		
1	0.13	0.11	0.85	0.82	1.17	1.04
2	0.54	0.55	0.74	0.72	0.97	1.03
3	1.37	1.25	0.63	0.66	1.09	0.97
4	2.41	2.34	0.57	0.62	1.03	0.92
5	3.56	3.55	0.53	0.59	1	0.9
6	4.61	4.56	0.5	0.61	1.01	0.82
7	5.87	5.91	0.47	0.57	0.99	0.83
8	7.64	7.63	0.43	0.56	1	0.76
9	11.39	11.26	0.39	0.58	1.01	0.67
10	24.24	21.36	0.33	0.46	1.13	0.73

Table 4: DEX size and Finishing Rate (FR) per decile

We compared the finishing rate of malware and goodware applications for evaluated tools. Because, the size of applications impacts this finishing rate, it is interesting to compare the success rate for each decile of bytecode size. Table 4 reports the bytecode size and the finishing rate of goodware and malware in each decile of size. We also computed the ratio of the bytecode size and finishing rate for the two populations. We observe that the ratio for the finishing rate

574 decreases from 1.04 to 0.73, while the ratio of the bytecode size is around 1. We conclude from  
575 this table that analyzing malware triggers less errors than for goodware.

576 **RQ3 answer:** Analyzing malware applications triggers less errors for static analysis tools  
577 than analyzing goodware for comparable bytecode size.



## CONTRIBUTION 2

---

580 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 581 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequale doleamus animo,  
 582 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 583 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 584 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
 585 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
 586 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
 587 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
 588 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
 589 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
 590 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis  
 591 mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc  
 592 sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita  
 593 prorsus existimo, neque eum Torquatium, qui hoc primum cognomen invenerit, aut torquem  
 594 illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt  
 595 vitae sine metu degendae praesidia firmissima. – Filium morte multavit. – Si sine causa, nollem  
 596 me ab eo delectari, quod ista Platonis, Aristoteli, Theophrasti orationis ornamenta neglexerit.  
 597 Nam illud quidem physici, credere aliquid esse minimum, quod profecto numquam putavisset,  
 598 si a Polyaeo, familiari suo, geometrica discere maluisset quam illum etiam ipsum dedocere.  
 599 Sol Democrito magnus videtur, quippe homini erudito in geometriaque perfecto, huic pedalis  
 600 fortasse; tantum enim esse omnino in nostris poetis aut inertissimae segnitiae est aut fastidii  
 601 delicatissimi. Mihi quidem videtur, inermis ac nudus est. Tollit definitiones, nihil de dividendo ac  
 602 partiendo docet, non quo ignorare vos arbitrer, sed ut ratione et via procedat oratio. Quaerimus  
 603 igitur, quid sit extremum et ultimum bonorum, quod omnium philosophorum sententia tale  
 604 debet esse, ut eius magnitudinem celeritas, diuturnitas allevatio consoletur. Ad ea cum  
 605 accedit, ut neque divinum numen horreat nec praeteritas voluptates effluere patiatur earumque  
 606 assidua recordatione laetetur, quid est, quod huc possit, quod melius sit, migrare de vita. His  
 607 rebus instructus semper est in voluptate esse aut in armatum hostem impetum fecisse aut in  
 608 poetis evolvendis, ut ego et Triarius te hortatore facimus, consumeret, in quibus hoc primum  
 609 est in quo admirer, cur in gravissimis rebus non delectet eos sermo patrius, cum idem fabellas  
 610 Latinas ad verbum e Graecis expressas non inviti legant. Quis enim tam inimicus paene nomini

---

611 Romano est, qui Ennii Medeam aut Antiopam Pacuvii spernat aut reiciat, quod se isdem  
612 Euripidis fabulis delectari dicat, Latinas litteras oderit? Synephebos ego, inquit, potius Caecilii  
613 aut Andriam Terentii quam utramque Menandri legam? A quibus tantum dissentio, ut, cum  
614 Sophocles vel optime scripserit Electram, tamen male conversam Atilii mihi legendam putem,  
615 de quo Lucilius: 'ferreum scriptorem', verum, opinor, scriptorem tamen, ut legendus sit. Rudem  
616 enim esse omnino in nostris poetis aut inertissimae segnitiae est aut in dolore. Omnis autem  
617 privatione doloris putat Epicurus.

## CONTRIBUTION N

---

620 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 621 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aeque doleamus animo,  
 622 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 623 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 624 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
 625 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
 626 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
 627 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
 628 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
 629 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
 630 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis  
 631 mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc  
 632 sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita  
 633 prorsus existimo, neque eum Torquatum, qui hoc primum cognomen invenerit, aut torquem  
 634 illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt  
 635 vitae sine metu degendae praesidia firmissima. – Filium morte multavit. – Si sine causa, nollem  
 636 me ab eo delectari, quod ista Platonis, Aristoteli, Theophrasti orationis ornamenta neglexerit.  
 637 Nam illud quidem physici, credere aliquid esse minimum, quod profecto numquam putavisset,  
 638 si a Polyaeo, familiari suo, geometrica discere maluisset quam illum etiam ipsum dedocere.  
 639 Sol Democrito magnus videtur, quippe homini erudito in geometriaque perfecto, huic pedalis  
 640 fortasse; tantum enim esse omnino in nostris poetis aut inertissimae segnitiae est aut fastidii  
 641 delicatissimi. Mihi quidem videtur, inermis ac nudus est. Tollit definitiones, nihil de dividendo ac  
 642 partiendo docet, non quo ignorare vos arbitrer, sed ut ratione et via procedat oratio. Quaerimus  
 643 igitur, quid sit extremum et ultimum bonorum, quod omnium philosophorum sententia tale  
 644 debet esse, ut eius magnitudinem celeritas, diuturnitas allevatio consoletur. Ad ea cum  
 645 accedit, ut neque divinum numen horreat nec praeteritas voluptates effluere patiatur earumque  
 646 assidua recordatione laetetur, quid est, quod huc possit, quod melius sit, migrare de vita. His  
 647 rebus instructus semper est in voluptate esse aut in armatum hostem impetum fecisse aut in  
 648 poetis evolvendis, ut ego et Triarius te hortatore facimus, consumeret, in quibus hoc primum  
 649 est in quo admirer, cur in gravissimis rebus non delectet eos sermo patrius, cum idem fabellas  
 650 Latinas ad verbum e Graecis expressas non inviti legant. Quis enim tam inimicus paene nomini

---

651 Romano est, qui Ennii Medeam aut Antiopam Pacuvii spernat aut reiciat, quod se isdem  
652 Euripidis fabulis delectari dicat, Latinas litteras oderit? Synephebos ego, inquit, potius Caecilii  
653 aut Andriam Terentii quam utramque Menandri legam? A quibus tantum dissentio, ut, cum  
654 Sophocles vel optime scripserit Electram, tamen male conversam Atilii mihi legendam putem,  
655 de quo Lucilius: 'ferreum scriptorem', verum, opinor, scriptorem tamen, ut legendus sit. Rudem  
656 enim esse omnino in nostris poetis aut inertissimae segnitiae est aut in dolore. Omnis autem  
657 privatione doloris putat Epicurus.

## CONCLUSION

---

TODO 17 ► *Conclude* ◀

661 Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt  
 662 ut labore et dolore magnam aliquam quaerat voluptatem. Ut enim aequale doleamus animo,  
 663 cum corpore dolemus, fieri tamen permagna accessio potest, si aliquod aeternum et infinitum  
 664 impendere malum nobis opinemur. Quod idem licet transferre in voluptatem, ut postea variari  
 665 voluptas distinguere possit, augeri amplificarique non possit. At etiam Athenis, ut e patre  
 666 audiebam facete et urbane Stoicos irridente, statua est in quo a nobis philosophia defensa  
 667 et collaudata est, cum id, quod maxime placeat, facere possimus, omnis voluptas assumenda  
 668 est, omnis dolor repellendus. Temporibus autem quibusdam et aut officiis debitis aut rerum  
 669 necessitatibus saepe eveniet, ut et voluptates repudiandae sint et molestiae non recusandae.  
 670 Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis,  
 671 saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis  
 672 mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc  
 673 sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita  
 674 prorsus existimo, neque eum Torquatum, qui hoc primus cognomen invenerit, aut torquem  
 675 illum hosti detraxisse, ut aliquam ex eo est consecutus? – Laudem et caritatem, quae sunt  
 676 vitae sine metu degendae praesidia firmissima. – Filium morte multavit. – Si sine causa, nollem  
 677 me ab eo delectari, quod ista Platonis, Aristoteli, Theophrasti orationis ornamenta neglexerit.  
 678 Nam illud quidem physici, credere aliquid esse minimum, quod profecto numquam putavisset,  
 679 si a Polyaeno, familiari suo, geometrica discere maluisset quam illum etiam ipsum dedocere.  
 680 Sol Democrito magnus videtur, quippe homini erudito in geometriaque perfecto, huic pedalis  
 681 fortasse; tantum enim esse omnino in nostris poetis aut inertissimae segnitiae est aut fastidii  
 682 delicatissimi. Mihi quidem videtur, inermis ac nudus est. Tollit definitiones, nihil de dividendo ac  
 683 partiendo docet, non quo ignorare vos arbitrer, sed ut ratione et via procedat oratio. Quaerimus  
 684 igitur, quid sit extremum et ultimum bonorum, quod omnium philosophorum sententia tale  
 685 debet esse, ut eius magnitudinem celeritas, diuturnitatem allevatio consoletur. Ad ea cum  
 686 accedit, ut neque divinum numen horreat nec praeteritas voluptates effluere patiatur earumque  
 687 assidua recordatione laetetur, quid est, quod huc possit, quod melius sit, migrare de vita. His  
 688 rebus instructus semper est in voluptate esse aut in armatum hostem impetum fecisse aut in  
 689 poetis evolvendis, ut ego et Triarius te hortatore facimus, consumeret, in quibus hoc primum

---

690 est in quo admirer, cur in gravissimis rebus non delectet eos sermo patrius, cum idem fabellas  
691 Latinas ad verbum e Graecis expressas non inviti legant. Quis enim tam inimicus paene nomini  
692 Romano est, qui Ennii Medeam aut Antiopam Pacuvii spernat aut reiciat, quod se isdem  
693 Euripidis fabulis delectari dicat, Latinas litteras oderit? Synephebos ego, inquit, potius Caecilii  
694 aut Andriam Terentii quam utramque Menandri legam? A quibus tantum dissentio, ut, cum  
695 Sophocles vel optime scripserit Electram, tamen male conversam Atilii mihi legendam putem,  
696 de quo Lucilius: 'ferreum scriptorem', verum, opinor, scriptorem tamen, ut legendus sit. Rudem  
697 enim esse omnino in nostris poetis aut inertissimae segnitiae est aut in dolore. Omnis autem  
698 privatione doloris putat Epicurus.

- 
- [1] Kevin Allix, Tegawendé F. Bissyandé, Jacques Klein, and Yves Le Traon. 2016. AndroZoo: Collecting Millions of Android Apps for the Research Community. In *13th Working Conference on Mining Software Repositories (MSR)*, May 2016. 468–471.
- [2] Steven Arzt, Siegfried Rasthofer, Christian Fritz, Eric Bodden, Alexandre Bartel, Jacques Klein, Yves Le Traon, Damien Octeau, and Patrick McDaniel. 2014. FlowDroid: Precise Context, Flow, Field, Object-sensitive and Lifecycle-aware Taint Analysis for Android Apps. In *ACM SIGPLAN Conference on Programming Language Design and Implementation*, June 05, 2014. ACM Press, Edinburgh, UK, 259–269. <https://doi.org/10.1145/2666356.2594299>
- [3] Tanzirul Azim and Iulian Neamtiu. 2013. Targeted and Depth-First Exploration for Systematic Testing of Android Apps. In *Proceedings of the 2013 ACM SIGPLAN International Conference on Object Oriented Programming Systems Languages & Applications, OOPSLA 2013, Part of SPLASH 2013, Indianapolis, IN, USA, October 26-31, 2013*, 2013. ACM, 641–660. <https://doi.org/10.1145/2509136.2509549>
- [4] Sam Blackshear, Bor-Yuh Evan Chang, and Manu Sridharan. 2013. Thresher: Precise Refutations for Heap Reachability. *ACM SIGPLAN Notices* 48, 6 (June 2013), 275–286. <https://doi.org/10.1145/2499370.2462186>
- [5] Amiangshu Bosu, Fang Liu, Danfeng (Daphne) Yao, and Gang Wang. 2017. Collusive Data Leak and More: Large-scale Threat Analysis of Inter-app Communications. In *Proceedings of the 2017 ACM on Asia Conference on Computer and Communications Security*, April 02, 2017. ACM, Abu Dhabi United Arab Emirates, 71–85. <https://doi.org/10.1145/3052973.3053004>
- [6] Kwanghoon Choi and Byeong-Mo Chang. 2014. A Type and Effect System for Activation Flow of Components in Android Programs. *Information Processing Letters* 114, 11 (2014), 620–627. <https://doi.org/10.1016/j.ipl.2014.05.011>
- [7] Anthony Desnos and Geoffroy Gueguen. 2011. Android: From Reversing to Decompilation. *Black Hat Abu Dhabi* (2011). Retrieved from [https://media.blackhat.com/bh-ad-11/Desnos/bh-ad-11-DesnosGueguen-Andriod-Reversing\\_to\\_Decompile\\_WP.pdf](https://media.blackhat.com/bh-ad-11/Desnos/bh-ad-11-DesnosGueguen-Andriod-Reversing_to_Decompile_WP.pdf)
- [8] William Enck, Peter Gilbert, Byung-Gon Chun, Landon P. Cox, Jaeyeon Jung, Patrick McDaniel, and Anmol N. Sheth. 2010. TaintDroid: An Information-Flow Tracking System for Realtime Privacy Monitoring on Smartphones. In *9th USENIX Symposium on Operat-*

---

ing Systems Design and Implementation, October 2010. USENIX Association, Vancouver, BC, Canada, 393–407.

[9] Sascha Fahl, Marian Harbach, Thomas Muders, Lars Baumgärtner, Bernd Freisleben, and Matthew Smith. 2012. Why Eve and Mallory Love Android: An Analysis of Android SSL (in)Security. In *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, October 16, 2012. ACM, Raleigh North Carolina USA, 50–61. <https://doi.org/10.1145/2382196.2382205>

[10] Hugo Gascon, Fabian Yamaguchi, Daniel Arp, and Konrad Rieck. 2013. Structural Detection of Android Malware Using Embedded Call Graphs. In *Proceedings of the 2013 ACM Workshop on Artificial Intelligence and Security*, November 04, 2013. ACM, Berlin Germany, 45–54. <https://doi.org/10.1145/2517312.2517315>

[11] Dimitris Geneiatakis, Igor Nai Fovino, Ioannis Kounelis, and Pasquale Stirparo. 2015. A Permission Verification Approach for Android Mobile Applications. *Computers & Security* 49, (March 2015), 192–205. <https://doi.org/10.1016/j.cose.2014.10.005>

[12] Michael I. Gordon, Deokhwan Kim, Jeff H. Perkins, Limei Gilham, Nguyen Nguyen, and Martin C. Rinard. 2015. Information Flow Analysis of Android Applications in DroidSafe. In *22nd Annual Network and Distributed System Security Symposium, NDSS 2015, San Diego, California, USA, February 8-11, 2015*, 2015. The Internet Society.

[13] Johannes Hoffmann, Martin Ussath, Thorsten Holz, and Michael Spreitzenbarth. 2013. Slicing Droids: Program Slicing for Smali Code. In *Proceedings of the 28th Annual ACM Symposium on Applied Computing (SAC '13)*, March 18, 2013. Association for Computing Machinery, New York, NY, USA, 1844–1851. <https://doi.org/10.1145/2480362.2480706>

[14] Jinseong Jeon, Kristopher K. Micinski, Jeffrey A. Vaughan, Ari Fogel, Nikhilesh Reddy, Jeffrey S. Foster, and Todd Millstein. 2012. Dr. Android and Mr. Hide: Fine-Grained Permissions in Android Applications. In *Proceedings of the Second ACM Workshop on Security and Privacy in Smartphones and Mobile Devices*, October 19, 2012. ACM, Raleigh North Carolina USA, 3–14. <https://doi.org/10.1145/2381934.2381938>

[15] William Klieber, Lori Flynn, Amar Bhosale, Limin Jia, and Lujo Bauer. 2014. Android Taint Flow Analysis for App Sets. In *Proceedings of the 3rd ACM SIGPLAN International Workshop on the State of the Art in Java Program Analysis*, June 12, 2014. ACM, Edinburgh United Kingdom, 1–6. <https://doi.org/10.1145/2614628.2614633>

[16] Li Li, Alexandre Bartel, Tegawende F. Bissyande, Jacques Klein, Yves Le Traon, Steven Arzt, Siegfried Rasthofer, Eric Bodden, Damien Outeau, and Patrick McDaniel. 2015. IccTA: Detecting Inter-Component Privacy Leaks in Android Apps. In *2015 IEEE/ACM*



- 37th IEEE International Conference on Software Engineering, May 2015. IEEE, Florence, Italy, 280–291. <https://doi.org/10.1109/ICSE.2015.48>
- [17] Li Li, Alexandre Bartel, Tegawendé F. Bissyandé, Jacques Klein, and Yves Le Traon. 2015. ApkCombiner: Combining Multiple Android Apps to Support Inter-App Analysis. In *ICT Systems Security and Privacy Protection*, 2015. Springer International Publishing, Cham, 513–527. [https://doi.org/10.1007/978-3-319-18467-8\\_34](https://doi.org/10.1007/978-3-319-18467-8_34)
- [18] Li Li, Tegawendé F. Bissyandé, Mike Papadakis, Siegfried Rasthofer, Alexandre Bartel, Damien Ochteau, Jacques Klein, and Yves Le Traon. 2017. Static Analysis of Android Apps: A Systematic Literature Review. *Information and Software Technology* 88, (2017), 67–95. <https://doi.org/10.1016/j.infsof.2017.04.001>
- [19] Shuying Liang, Andrew W. Keep, Matthew Might, Steven Lyde, Thomas Gilray, Petey Aldous, and David Van Horn. 2013. Sound and Precise Malware Analysis for Android via Pushdown Reachability and Entry-Point Saturation. In *Proceedings of the Third ACM Workshop on Security and Privacy in Smartphones & Mobile Devices (SPSM '13)*, November 08, 2013. Association for Computing Machinery, New York, NY, USA, 21–32. <https://doi.org/10.1145/2516760.2516769>
- [20] Max Lillack, Christian Kästner, and Eric Bodden. 2014. Tracking Load-Time Configuration Options. In *Proceedings of the 29th ACM/IEEE International Conference on Automated Software Engineering (ASE '14)*, September 15, 2014. Association for Computing Machinery, New York, NY, USA, 445–456. <https://doi.org/10.1145/2642937.2643001>
- [21] Yepang Liu, Chang Xu, and Shing-Chi Cheung. 2014. Characterizing and Detecting Performance Bugs for Smartphone Applications. In *Proceedings of the 36th International Conference on Software Engineering*, May 31, 2014. ACM, Hyderabad India, 1013–1024. <https://doi.org/10.1145/2568225.2568229>
- [22] Linghui Luo, Felix Pauck, Goran Piskachev, Manuel Benz, Ivan Pashchenko, Martin Mory, Eric Bodden, Ben Hermann, and Fabio Massacci. 2022. TaintBench: Automatic Real-World Malware Benchmarking of Android Taint Analyses. *Empirical Software Engineering* 27, 1 (January 2022), 16. <https://doi.org/10.1007/s10664-021-10013-5>
- [23] Damien Ochteau, Daniel Luchaup, Matthew Dering, Somesh Jha, and Patrick McDaniel. 2015. Composite Constant Propagation: Application to Android Inter-Component Communication Analysis. In *2015 IEEE/ACM 37th IEEE International Conference on Software Engineering*, May 2015. IEEE, Florence, Italy, 77–88. <https://doi.org/10.1109/ICSE.2015.30>
- [24] Damien Ochteau, Patrick McDaniel, Somesh Jha, Alexandre Bartel, Eric Bodden, Jacques Klein, and Yves Le Traon. 2013. Effective Inter-Component communication mapping in

---

800        android: An essential step towards holistic security analysis. In *22nd USENIX Security*  
801        *Symposium (USENIX Security 13)*, 2013. 543–558.

802 [25] Felix Pauck, Eric Bodden, and Heike Wehrheim. 2018. Do Android Taint Analysis Tools  
803        Keep Their Promises?. In *Proceedings of the 2018 26th ACM Joint Meeting on European*  
804        *Software Engineering Conference and Symposium on the Foundations of Software Engi-*  
805        *neering*, October 26, 2018. ACM, Lake Buena Vista FL USA, 331–341. [https://doi.org/](https://doi.org/10.1145/3236024.3236029)  
806        10.1145/3236024.3236029

807 [26] Feargus Pendlebury, Fabio Pierazzi, Roberto Jordaney, Johannes Kinder, and Lorenzo  
808        Cavallaro. 2018. TESSERACT: Eliminating Experimental Bias in Malware Classification  
809        across Space and Time. (2018).

810 [27] Sebastian Poeplau, Yanick Fratantonio, Antonio Bianchi, Christopher Kruegel, and Gio-  
811        vanni Vigna. 2014. Execute This! Analyzing Unsafe and Malicious Dynamic Code Loading  
812        in Android Applications. In *21st Annual Network and Distributed System Security Sympo-*  
813        *sium, NDSS 2014, San Diego, California, USA, February 23-26, 2014*, 2014. The Internet  
814        Society.

815 [28] Bradley Reaves, Jasmine Bowers, Sigmund Albert Gorski III, Olabode Anise, Rahul  
816        Bobhate, Raymond Cho, Hiranava Das, Sharique Hussain, Hamza Karachiwala, Nolen  
817        Scaife, Byron Wright, Kevin Butler, William Enck, and Patrick Traynor. 2016. \*droid:  
818        Assessment and Evaluation of Android Application Analysis Tools. *ACM Computing*  
819        *Surveys* 49, 3 (October 2016), 1–30. <https://doi.org/10.1145/2996358>

820 [29] Atanas Rountev and Dacong Yan. 2014. Static Reference Analysis for GUI Objects in  
821        Android Software. In *Proceedings of Annual IEEE/ACM International Symposium on*  
822        *Code Generation and Optimization*, February 15, 2014. ACM, Orlando FL USA, 143–153.  
823        <https://doi.org/10.1145/2544137.2544159>

824 [30] Feng Shen, Namita Vishnubhotla, Chirag Todarka, Mohit Arora, Babu Dhandapani, Eric  
825        John Lehner, Steven Y. Ko, and Lukasz Ziarek. 2014. Information Flows as a Permission  
826        Mechanism. In *Proceedings of the 29th ACM/IEEE International Conference on Automated*  
827        *Software Engineering*, September 15, 2014. ACM, Vasteras Sweden, 515–526. [https://doi.](https://doi.org/10.1145/2642937.2643018)  
828        [org/10.1145/2642937.2643018](https://doi.org/10.1145/2642937.2643018)

829 [31] Dennis Titze and Julian Schutte. 2015. Apparecium: Revealing Data Flows in Android  
830        Applications. In *2015 IEEE 29th International Conference on Advanced Information*  
831        *Networking and Applications*, March 2015. IEEE, Gwangju, South Korea, 579–586. [https://](https://doi.org/10.1109/AINA.2015.239)  
832        [doi.org/10.1109/AINA.2015.239](https://doi.org/10.1109/AINA.2015.239)

833 [32] Timothy Vidas, Jiaqi Tan, Jay Nahata, Chaur Lih Tan, Nicolas Christin, and Patrick  
834        Tague. 2014. A5: Automated Analysis of Adversarial Android Applications. In *Proceedings*

- 835        *of the 4th ACM Workshop on Security and Privacy in Smartphones & Mobile Devices*,  
836        November 07, 2014. ACM, Scottsdale Arizona USA, 39–50. [https://doi.org/10.1145/](https://doi.org/10.1145/2666620.2666630)  
837        2666620.2666630
- 838        [33] Fengguo Wei, Sankardas Roy, Xinming Ou, and Robby. 2014. Amandroid: A Precise and  
839        General Inter-component Data Flow Analysis Framework for Security Vetting of Android  
840        Apps. In *ACM SIGSAC Conference on Computer and Communications Security*, Novem-  
841        ber 2014. ACM, Scottsdale Arizona USA, 1329–1341. [https://doi.org/10.1145/2660267.](https://doi.org/10.1145/2660267.2660357)  
842        2660357
- 843        [34] Erik Ramsgaard Wognsen, Henrik Søndberg Karlsen, Mads Chr. Olesen, and René Rydhof  
844        Hansen. 2014. Formalisation and Analysis of Dalvik Bytecode. *Science of Computer*  
845        *Programming* 92, (October 2014), 25–55. <https://doi.org/10.1016/j.scico.2013.11.037>
- 846        [35] Mingyuan Xia, Lu Gong, Yuanhao Lyu, Zhengwei Qi, and Xue Liu. 2015. Effective Real-  
847        Time Android Application Auditing. In *2015 IEEE Symposium on Security and Privacy*,  
848        May 2015. IEEE, San Jose, CA, 899–914. <https://doi.org/10.1109/SP.2015.60>
- 849        [36] Shengqian Yang, Dacong Yan, Haowei Wu, Yan Wang, and Atanas Rountev. 2015. Static  
850        Control-Flow Analysis of User-Driven Callbacks in Android Applications. In *2015 IEEE/*  
851        *ACM 37th IEEE International Conference on Software Engineering*, May 2015. IEEE,  
852        Florence, Italy, 89–99. <https://doi.org/10.1109/ICSE.2015.31>
- 853        [37] Yury Zhauniarovich, Maqsood Ahmad, Olga Gadyatskaya, Bruno Crispo, and Fabio Mas-  
854        sacci. 2015. StaDynA: Addressing the Problem of Dynamic Code Updates in the Security  
855        Analysis of Android Applications. In *Proceedings of the 5th ACM Conference on Data and*  
856        *Application Security and Privacy*, March 02, 2015. ACM, San Antonio Texas USA, 37–48.  
857        <https://doi.org/10.1145/2699026.2699105>







858 Titre : TODO 18 ► *Find a title* ◀

859 Mots clés : Android, Analyse de Maliciels

860 Résumé : Lorem ipsum dolor sit amet, con-  
861 sectetur adipiscing elit, sed do eiusmod tem-  
862 por incididunt ut labore et dolore magnam  
863 aliquam quaerat voluptatem. Ut enim aequa-  
864 doleamus animo, cum corpore dolemus, fieri  
865 tamen permagna accessio potest, si aliquod  
866 aeternum et infinitum impendere malum nobis  
867 opinemur. Quod idem licet transferre in volup-  
868 tatem, ut postea variari voluptas distinguere  
869 possit, augeri amplificarique non possit. At  
870 etiam Athenis, ut e patre audiebam facete et  
871 urbane Stoicos irridente, statua est in quo a  
872 nobis philosophia defensa et collaudata est,  
873 cum id, quod maxime placeat, facere pos-

simus, omnis voluptas assumenda est, omnis  
dolor repellendus. Temporibus autem quibus-  
dam et aut officiis debitis aut rerum necessitatibus saepe eveniet, ut et voluptates repudian-  
dae sint et molestiae non recusandae. Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis, saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus existimo, neque.

875 Title : TODO 19 ► *Find a title* ◀

876 Keywords: Android, Malware Analysis, TODO 20 ► *More Keywords* ◀

877 Abstract: Lorem ipsum dolor sit amet, con-  
878 sectetur adipiscing elit, sed do eiusmod tem-  
879 por incididunt ut labore et dolore magnam  
880 aliquam quaerat voluptatem. Ut enim aequa-  
881 doleamus animo, cum corpore dolemus, fieri  
882 tamen permagna accessio potest, si aliquod  
883 aeternum et infinitum impendere malum nobis  
884 opinemur. Quod idem licet transferre in volup-  
885 tatem, ut postea variari voluptas distinguere  
886 possit, augeri amplificarique non possit. At  
887 etiam Athenis, ut e patre audiebam facete et  
888 urbane Stoicos irridente, statua est in quo a  
889 nobis philosophia defensa et collaudata est,  
890 cum id, quod maxime placeat, facere pos-

simus, omnis voluptas assumenda est, omnis  
dolor repellendus. Temporibus autem quibus-  
dam et aut officiis debitis aut rerum necessitatibus saepe eveniet, ut et voluptates repudian-  
dae sint et molestiae non recusandae. Itaque earum rerum defuturum, quas natura non depravata desiderat. Et quem ad me accedis, saluto: 'chaere,' inquam, 'Tite!' lictores, turma omnis chorusque: 'chaere, Tite!' hinc hostis mi Albucius, hinc inimicus. Sed iure Mucius. Ego autem mirari satis non queo unde hoc sit tam insolens domesticarum rerum fastidium. Non est omnino hic docendi locus; sed ita prorsus existimo, neque.