ParanoYa

Lóránt Boráros, Filip Budáč, Martin Cehelský, Silvia Holecová

Október 2019

1 Analýza

- **Súčasný stav aplikácie**

  Aplikácia s názvom ParanoYa, slúži na štatistické testovanie pseudonáhodných postupností. V aplikácii sú implementované rôzne sady testov ako napríklad NIST, FIPS, Diehard. Pomocou aplikácie je tiež možné vyhodnocovať jednotlivé testované postupnosti a to na základe dvoch metodík. Výstup aplikácie sa spracúva pomocou Microsoft Excel dokumentu, pomocou ktorého je možné vyhodnotiť dosiahnuté výsledky. Aplikácia bola vytvorená pomocou frameworku Qt a jednotlivé sady testov sú implementované v jazyku C.

- **Metodiky použité pri vyhodnocovaní testov** porovnanie s odporucanymi/ interval odporucanych (pouzivnie pre ucely kryptografie na predmete) diehard test sa nedaju pouzit 10 MB

- **Použité sady testov**

  Testovanie je process vykonávania jedného alebo viacerých testovacích prípadov na základe stanovených podmienok, počas ktorého porovnávame aktuálne a očakávané správanie. V aplikácií sú implementované rôzne sady testov, medzi ktorými sú NIST, FIPS a Diehard.

  1. **NIST**

     NIST je štatistický balík testov, ktorý slúži na testovanie náhodnosti lubovoľne dlhých binárnych postupností vytvorených pomocou generátora náhodných alebo pseudonáhodných postupností. Tento balík pozostáva z 15 nasledujúcich testov:

     (a) **The Frequency (Monobit) Test**
     Cieľom tohto testu je zistiť, či pomer núl a jednotiek v danej postupnosti zodpovedá očakávanemu pomeru pre náhodnú postupnosť. Počet jednotiek a núl v postupnosti by mal byť približne rovnaký, čo tiež skúma daný test.

     (b) **Frequency Test within a Block**

Tento test posudzuje pomer núl a jednotiek v M-bitových blokoch. Cieľom testu je určiť, či je ich frekvencia M-bitový blok približne M/2.

(c) **The Runs Test**
V tomto teste ide o celkový počet núl a jednotiek runov v celej postupnosti, kde run predstavuje nepretržitú sekvenciu rovnakých bitov.Run o dĺžke k znamená, že pozostáva z k identických bitov a je ohraničený pred a za s bitom, ktorý má opačnú hodnotu. Cieľom tohto testu je zistiť, či počet runov jednotiek a núl rozličnej dĺžky je taký ako sa očakáva pre náhodné postupnosti.Tento test sa používa najmä na posúdenie či je menlivosť medzi takýmito substringmi je príliš pomalá alebo príliš rýchla.

(d) **Tests for the Longest-Run-of-Ones in a Block**
Tento test sa zameriava na najdlhší run jednotiek v rámci M-bitových blokov.Jeho cieľom je zistiť, či dĺžka najdlhšieho runu jednotiek v testovanej postupnosti je konzistentná s dĺžkou najdlhšieho runu jednotiek, ktorá je očakávaná v náhodných postupnostiach. Nepravidelnosť v očakávanej dĺžke najdlhšieho runu jednotiek znamená, že existuje nepravidelnosť v očakávanej dĺžke najdlhšieho runu núl. Dlhé runy núl sa nevyhodnocujú separátne z dôvodu obáv zo štatistickej nezávislosti medzi testami.

(e) **The Binary Matrix Rank Test**
Test je zameraný na nesúvislé poradie podmatíc v celej postupnosti.Účel tohto testu je skontrolovať lineárnu závislosť u pevnej dĺžky podreťazcov pôvodnej postupnosti.

(f) **The Discrete Fourier Transform (Spectral) Test**
Ťažiskom tohto testu sú výšky vrcholov vo Fourierovej transformácii. Cieľom tohto testu je odhaliť periodické funkcie(napríklad opakujúce sa vzory, ktoré sa nachádzajú blízko seba) v testovanej postupnosti, ktoré by naznačovali odchýlku od predpokladu náhodnosti.

(g) **The Non-overlapping Template Matching Test**
Náhodná číselná sekvencia je rozdelená do nezávislých podreťazcov s dĺžkou M a počet výskytov šablóny B, kttorá predstavuje m-bitový run jednotiek v každom z podreťazcov.IfP-hodnota chi-kvadrátovej štatistiky je menšia ako úroveň signifikantnosti, test dospieva k záveru, že testovaná sekvencia sa javí ako náhodná. V opačnom prípade sa pri teste dospelo k záveru, že opakovaný test sa zdá byť náhodný. Priepustnosť je definovaná pomerom sekvencií, ktoré prešli testom.

(h) **The Overlapping Template Matching Test**
Tento test zisťuje počet výskytov vo vopred špecifikovaných cieľových reťazcoch. Test používa m-bitové okno na hľadanie špecifického m-bitového vzoru. Pokiaľ vzor nie je nájdený, okno sa posúva o jednu bitovú pozíciu.Ak sa hľadaný vzor nájde, okno sa posúva

len jeden bit pred obnovením hľadania.

(i) **Maurer's "Universal Statistical" Test**
Účelom tohto testu je odhaliť, či postupnosť môže byť výrazne komprimovaná bez straty informácií alebo nie. Príliš komprimovaná postupnosť sa považuje za nenáhodnú.

(j) **The Linear Complexity Test**
Cieľom tohto testu je určiť, či je sekvencia dostatočne zložitá na to, aby sa považovala za náhodnú.

(k) **The Serial Test**
Účelom tohto testu je zistiť, či je počet výskytov prekrývajúcich sa vzorov m-bitov približne rovnaký, ako by sa očakávalo v prípade náhodnej sekvencie.

(l) **The Approximate Entropy Test**
Test sa zameriava na frekvenciu všetkých možných prekrývaní m-bitových vzorov v celej sekvencii. Účelom tohto testu je porovnať frekvenciu prekrývajúcich sa blokov dvoch po sebe nasledujúcich alebo susediacich dĺžok (m a m+1) s očakávaným výsledkom pre náhodnú postupnosť.

(m) **The Cumulative Sums (Cusums) Test**
Tento test sa zameriava na maximálnu odchýlku (od nuly) náhodnej prechadzky(walk?????) definovanej kumulatívnym súčtom upravených (-1, +1) číslic v postupnosti. Cieľom testu je určiť, či kumulatívny súčet čiastkových sekvencií vyskytujúcich sa v testovanej sekvencii je príliš veľký alebo príliš malý relatívny k očakávanemu správaniu tejto kumulatívnej sumy pre náhodné sekvencie. Táto kumulatívna suma, môže byť považovaná za náhodnu walk. Pre náhodnú sekvenciu by mala byť odchýlka náhodnej walk blízko nuly. Pre určité typy náhodných sekvencií budú odchýlky tejto náhodnej walk väčšie od nuly.

(n) **The Random Excursions Test**
Test je zameraný na počet cyklov, ktoré majú presne K výskytov v kumulatívnom súčte náhodných krokov. Kumulatívny súčet môžeme zistiť ak sú čiastkové súčty (0, 1) sekvencie upravené na (-1, +1). Náhodná odchýlka náhodných krokov pozostáva zo sekvencie n krokov jednotkovej dĺžky. Cieľom testu je zistiť, či počet výskytov stavu s náhodnými krokmi prekračuje to, čo sa očakáva od náhodnej sekvencie.

(o) **The Random Excursions Variant Test**
Tento test skúma koľkokrát sa konkrétny stav vyskytne v kumulatívnom súčte náhodných krokov. Cieľom je odhaliť odchýlky od očakávaného počtu výskytov rôznych stavov v náhodných krokoch.

Tieto testy sa zaoberajú rôznymi typmi náhodností, ktoré by mohli vzniknúť v postupnosti. Niektoré z testov by bolo možné rozložiť na

rôzne subtesty. Poradie spustenia jednotlivých testov je lubovoľné, ale odporúča sa, aby bol Frequency test spustený ako prvý, pretože pokiaľ tento test zlyhá, pravdepodobnosť zlyhania ďalších testov je veľmi vysoká.

2. **FIPS** nist sp-822,fips 140-2 Test Federal Information Processing Standard predstavuje americký vládny bezpečnostný štandard, ktorý sa používa na shcválenie použitých kryptografických modulov. FIPS poskytuje rôzne druhy zabezpečenia a to na základe definovanej úrovne bezpečnosti. Takéto úrovne poznáme štyri:
   (a) **Úroveň 1** - najnižšia bezpečnostná úroveň, ktorá nevyžaduje špecifické mechanizmy fyzickej bezpečnosti, ale vyžaduje použitie aspoň jedného schváleného bezpečnostného algoritmu alebo funkcie
   (b) **Úroveň 2** - táto úroveň vyžaduje riadenie prístupu na základe rolí, vyžaduje tiež fyzické zabezpečenie
   (c) **Úroveň 3** - v tejto úrovni je poskytnutá autentifikácia založená na identite a fyzické zabezpečenie, mala by obsahovať mechanizmus na detekciu útoku a pokiaľ dôjde k hacknutiu systému, systém by mal byť schopný vymazať kritické bezpečnostné parametre
   (d) **Úroveň 4** - ide o najvyššiu úroveň zabezpečenia, okrem vyššie spomenutých náležitostí na systém sa v nej sprísňujú požiadavky fyzického zabezpečenia, je výhodná najmä pre prácu vo fyzicky nechránenom prostredí

   Validácia FIPS zahŕňa intenzívne testovanie na zistenie konkrétnych nedostatkov a slabín. Na to, aby systém spĺňal validáciu na základe FIPS, je potrebné, aby obsahoval kryptografické algoritmy a hashovacie funkcie.K trom najznámejším príkladom patrí AES, Triple DES a HMAC SHA-1

3. **Diehard**

   Diehard tests sú štatistické testy, ktoré slúžia na zhodnotenie miery kvality generátora náhodných čísel. Diehardová batéria testov pozostáva z rôznych, nezávislých štatistických testov. Výsledky týchto testov sa označujú ako p-hodnoty. Medzi Diehardové testy patria:
   (a) **The Birthday spacings test**
   V tomto teste sa najprv vyberá m narodenín v roku s n dňami, následne sa uvádza zoznam medzier medzi narodeninami. A nakoniec sa posudzuje asymptoticky Poissonove rozdelenie hodnoty j. Hodnota j predstavuje počet hodnôt, ktoré sa nachádzajú v uvedenom zozname medzier a pokiaľ sa v tomto zozname nachádza viackrát, tak j je asymptoticky Poissonovo rozdelené s priemerom $m^3/(4n)$. n musí byť dostatočne veľké, na porovnanie výsledkov s Poissonovým rozdelením.
   (b) **Overlapping permutations**

Tento test sleduje postupnosť jedného milióna 32-bitových náhodných celých čísel. Každá sada piatich po sebe idúcich celých čísel môže byť v jednom zo 120 stavov pre 5! možných usporiadaní piatich čísel.

(c) **Ranks of matrices**
Tento test sa vykonáva výberom určitého počtu bitov z určitého počtu náhodných čísel, aby sa vytvorila matica nad [0,1] a následne sa určí poradie matice. Počet radov by mal sledovať určité rozdelenie.!!!!!!!!!!!!!!

(d) **Monkey test**
Tiež nazývaný ako test bitových tokov. Tento test má svoj názov z nekonečnej "monkey theorem". Najlepšie sa dosiahne spracovaním sekvencií určitého počtu bitov ako slov" a spočítaním prekrývajúcich sa slov v pare. Počet slov", ktoré sa neobjavia, by mal nasledovať po známej distribúcii.

(e) **Count the 1s**
The test is done through counting the 1 bits in each of either successive or chosen bytes and converting the counts to "letters", and counting the occurrences of five-letter "words".

(f) **Parking lot test**
Randomly place unit circles in a 100 x 100 square. If the circle overlaps an existing one, try again. After 12,000 tries, the number of successfully "packed" circles should follow a certain normal distribution.

(g) **Minimum distance test**
Randomly place 8000 points in a 10,000 x 10,000 square and then find the minimum distance between the pairs. The square of this distance should be exponentially distributed with a certain mean.

(h) **Random spheres test**
Randomly choose 4000 points in a cube of edge 1000. Center a sphere on each point, whose radius is the minimum distance to another point. The smallest sphere s volume should be exponentially distributed with certain mean.

(i) **The squeeze test**
Multiply 231 by float random integers on [0,1) until you reach 1. Repeat this 100,000 times. The number of floats needed to reach 1 should follow a certain distribution.

(j) **Overlapping sums test**
Generate a long sequence of random floats on [0,1). Add sequence of 100 consecutive floats. The sums should be normally distributed with characteristic mean and sigma.

(k) **Runs test**
Generate a long sequence of random floats on [0,1). Count ascending and descending runs. The counts should follow a certain distribution.

(l) **The craps test**
Play 200,000 games of craps, counting the wins and the number of throws per game. Each count should follow a certain distribution.

1.1 Spôsob riešenia

Základ aplikácie, teda sady testov, sú rezprezentované ako knižnice v programovacom jazyku C. Používateľské rozhranie je vo frameworku Qt a výstup aplikácie sa v súčasnosti realizuje zapísaním do súboru typu *.xls*, ktorý používateľ môže čítať v tabuľkových editoroch. Návrh riešenia spočíva vo viacerých krokoch:

1. **Aktualizácia projektu pre súčasné vývojové prostredia**

   Pre vytvorenie používateľského rozhrania sme sa rozhodli použiť jazyk Python. Pomocou knižnice CTypes sme zabezpečili prepojenie s knižnicami v jazyku C. Z knižníc sme vytvorili Shared Object File "*.so*". Tento Shared Object nám umožní dynamicky prepájať knižnicu s rôznymi programmi.

1.2 Funkcionálne požiadavky

- **Štatistické testovanie pseudonáhodných postupností** - používateľ si bude schopný štatisticky otestovať rôzne pseudonáhodné postupnosti pomocou implementovaných sád testov
- **Nastavovanie jednotlivých testov** - používateľovi bude umožnené nastavovať a upravovať jednotlivé testy, podľa jeho kritérií
- **Vyhodnocovanie testovaných postupností** - po testovaní vybraných postupností, si bude môcť používateľ prezrieť vyhodnotenie testov na základe zvolenej metodiky

2 Existujúce programy

1. **Ent** je konzolová aplikácia, ktorá slúži na vyhodnocovanie pseudnáhodných postupností pre štatistické vzorkovacie aplikácie, šifrovacie a kompresné algoritmy. Ent vykonáva nasledujúce testy:

   - **Entropia**
   - **Chi-square Test**
   - **Arithmetic mean**
   - **Monte Carlo Value of Pi**
   - **Serial Correlational Coefficient**

   Ent poskyutje viacerých možností na modifikovanie spracovania údajov ako aj formát výstupu:

- **-b**
  Vstupné dáta s spracované ako bity namiesto bajtov.
- **-c**
  Na štandarný výstup vytlačí tabuľku vyskytnutých znakov, ktoré sú zobrazené s ich desatinnými bajtovými hodnotami spolu s prislúchajúcim znakom zo sady ISO 8859-1 Latin-1.
- **-f**
  Veľké písmená sú zmenené na malé pred vykonaním testov.
- **-t**
  Výstup je vo forme tzv. *terse mode*, v ktorom výstupy sú oddelené čiarkou(CSV formát).
- **-u**
  Výpis informácií

2. **Dieharder** je vylepšená verzia programu *Diehard battery of tests* s vyčisteným zdrojovým kódom, implementované v jazyku C. Najdôležitejšie vylepšenia sú jeho rýchlosť a vďaka architektúry rozšíriteľnosť sady testov. Je open source projektom, program je voľne dostupný a stiahnuteľný na jeho web-stránke. Dieharder umožňuje otestovanie generátorov priamo, vstupom môže byť aj neobmedzený tok náhodných čísel.

3. **Practically random** je knižnica implementované v programovacom jazyku C++ a slúži na testovanie generátorov náhodných postupností-*RNG*

4. **TestU01** je knižnica implementované v programovacom jazyku ANSI C. Obsahuje funkcie na empirické štatistické testovanie generátorov náhodných postupností. Aplikácia

3 Implementácia

3.1 UML Diagramy

3.1.1 Use Case Diagram

```
#==============================================================================#
#          dieharder version 3.29.4beta Copyright 2003 Robert G. Brown         #
#==============================================================================#
Installed dieharder tests:
 Test Number                          Test Name                Test Reliability
==============================================================================
  -d 0                            Diehard Birthdays Test              Good
  -d 1                               Diehard OPERM5 Test            Suspect
  -d 2                     Diehard 32x32 Binary Rank Test              Good
  -d 3                       Diehard 6x8 Binary Rank Test              Good
  -d 4                            Diehard Bitstream Test               Good
  -d 5                                     Diehard OPSO                Good
  -d 6                                  Diehard OQSO Test              Good
  -d 7                                   Diehard DNA Test              Good
  -d 8                 Diehard Count the 1s (stream) Test              Good
  -d 9                   Diehard Count the 1s Test (byte)              Good
  -d 10                         Diehard Parking Lot Test               Good
  -d 11          Diehard Minimum Distance (2d Circle) Test             Good
  -d 12          Diehard 3d Sphere (Minimum Distance) Test             Good
  -d 13                               Diehard Squeeze Test             Good
  -d 14                                  Diehard Sums Test         Do Not Use
  -d 15                                  Diehard Runs Test              Good
  -d 16                                Diehard Craps Test               Good
  -d 17                      Marsaglia and Tsang GCD Test               Good
  -d 100                                  STS Monobit Test              Good
  -d 101                                     STS Runs Test              Good
  -d 102                    STS Serial Test (Generalized)               Good
  -d 200                        RGB Bit Distribution Test               Good
  -d 201            RGB Generalized Minimum Distance Test               Good
  -d 202                           RGB Permutations Test                Good
  -d 203                             RGB Lagged Sum Test                Good
  -d 204                 RGB Kolmogorov-Smirnov Test Test               Good
```

Figure 1: Sada testov Dieharder

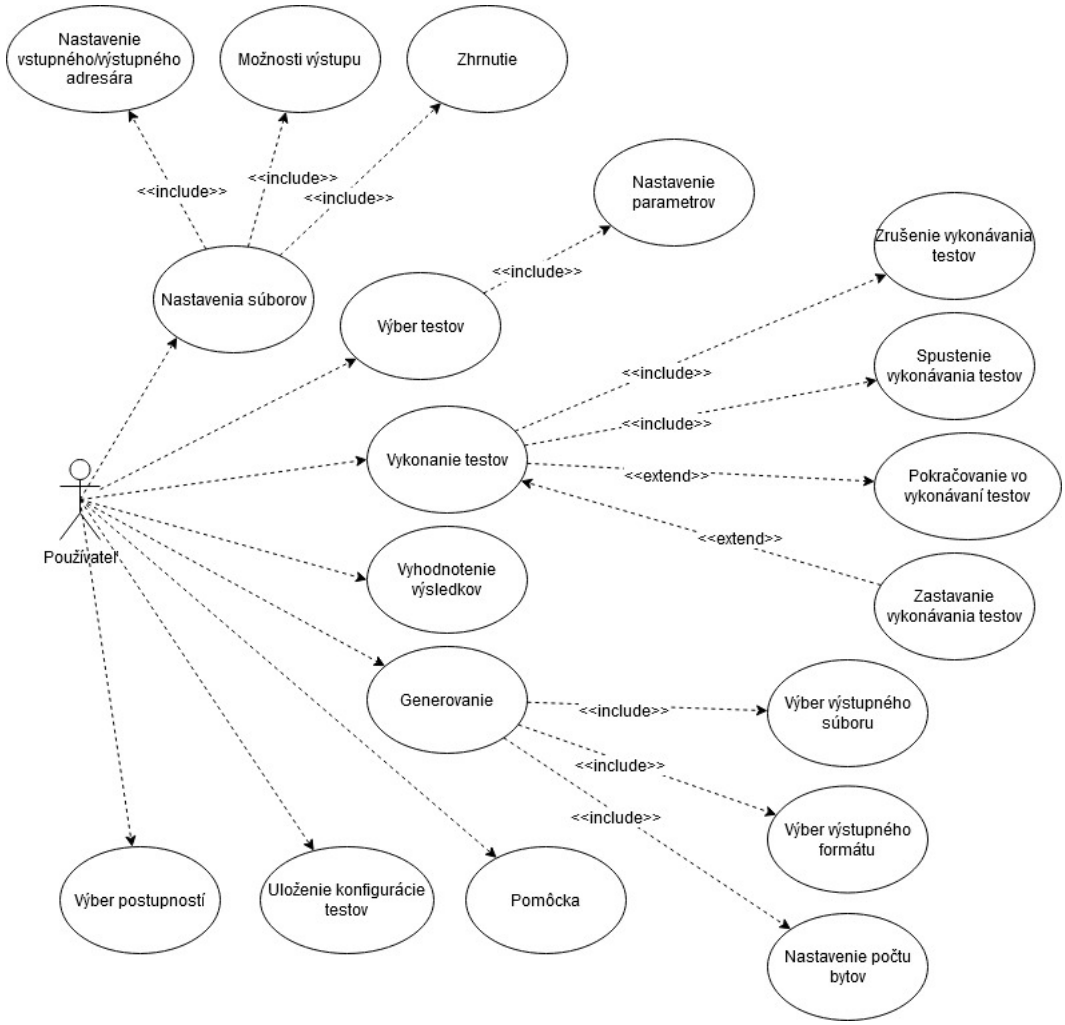

Figure 2: paranoYa - Use case diagram

3.1.2 Sequence Diagrams

Používateľ interaguje s grafickým rozhraním aplikácie. V záložke *Tasks* v navigačnej lište aplikácie, zvolí *Run*. Spustí sa testovanie pseudonáhodných postupností. Spusteniu predchádza nahranie postupnosti, zvolenie metodiky.

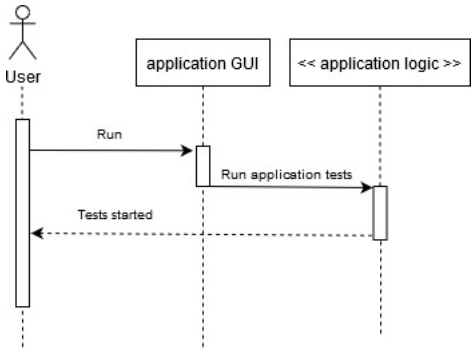

Figure 3: paranoYa - Sequence diagram (Run)

Používateľ interaguje s grafickým rozhraním aplikácie. V záložke *Tasks* v navigačnej lište aplikácie, zvolí *Continue*. Testovanie pseudonáhodných postupností pokračuje. Pokračovaniu testovania predchádza spustenie *Run* a pozastavenie testovania *Pause*.

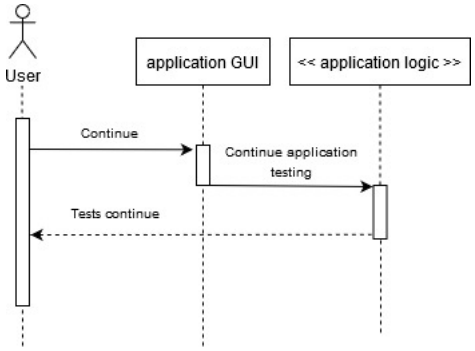

Figure 4: paranoYa - Sequence diagram (Continue)

Používateľ interaguje s grafickým rozhraním aplikácie. V záložke *Tasks* v navigačnej lište aplikácie, zvolí *Pause*. Preruší sa testovanie pseudonáhodných postupností. Prerušeniu predchádza spustenie testovania *Run*.

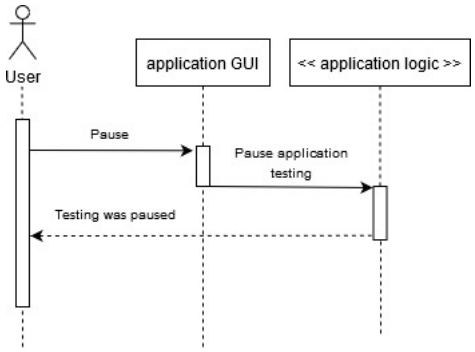

Figure 5: paranoYa - Sequence diagram (Pause)

Používateľ interaguje s grafickým rozhraním aplikácie. V záložke *Tasks* v navigačnej lište aplikácie, zvolí *Cancel*. Zastaví sa testovanie pseudonáhodných postupností. Zastaveniu predchádza spustenie testovania *Run*.

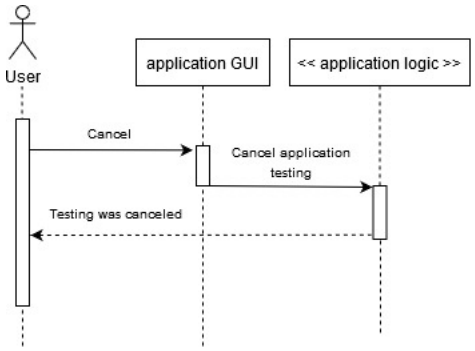

Figure 6: paranoYa - Sequence diagram (Cancel)

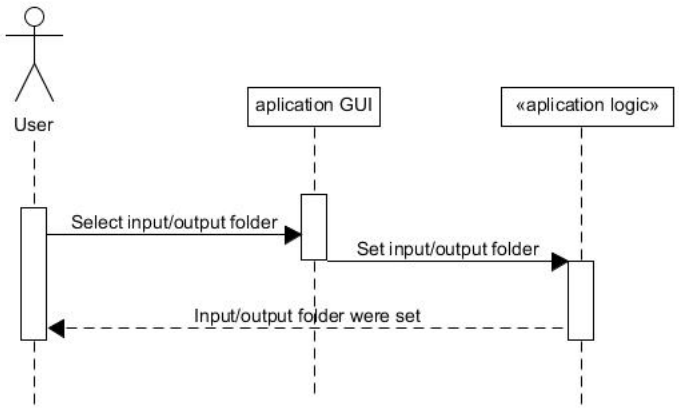

Figure 7: paranoYa - Sequence diagram (Set input/output folder)

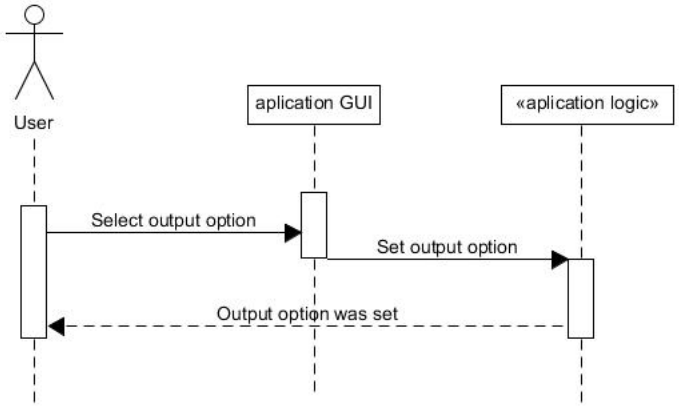

Figure 8: paranoYa - Sequence diagram (Select output option)

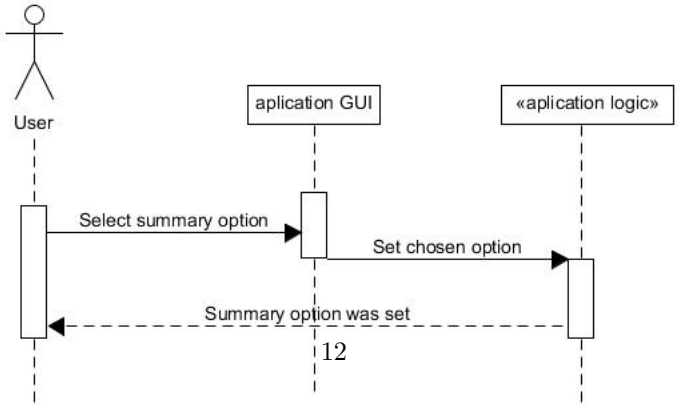

Figure 9: paranoYa - Sequence diagram (Select summary option)

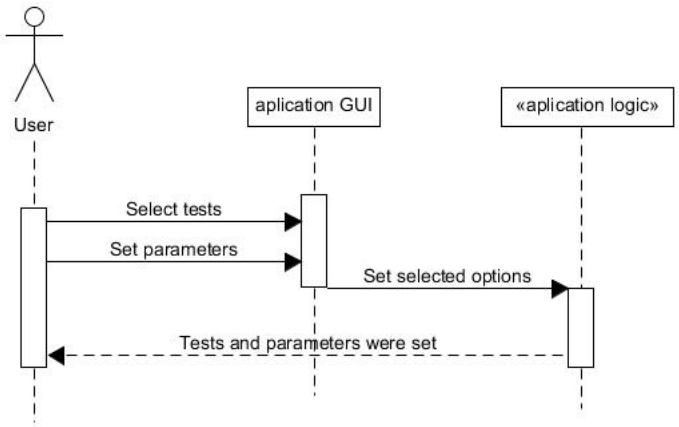

Figure 10: paranoYa - Sequence diagram (Set parameters for selected tests)

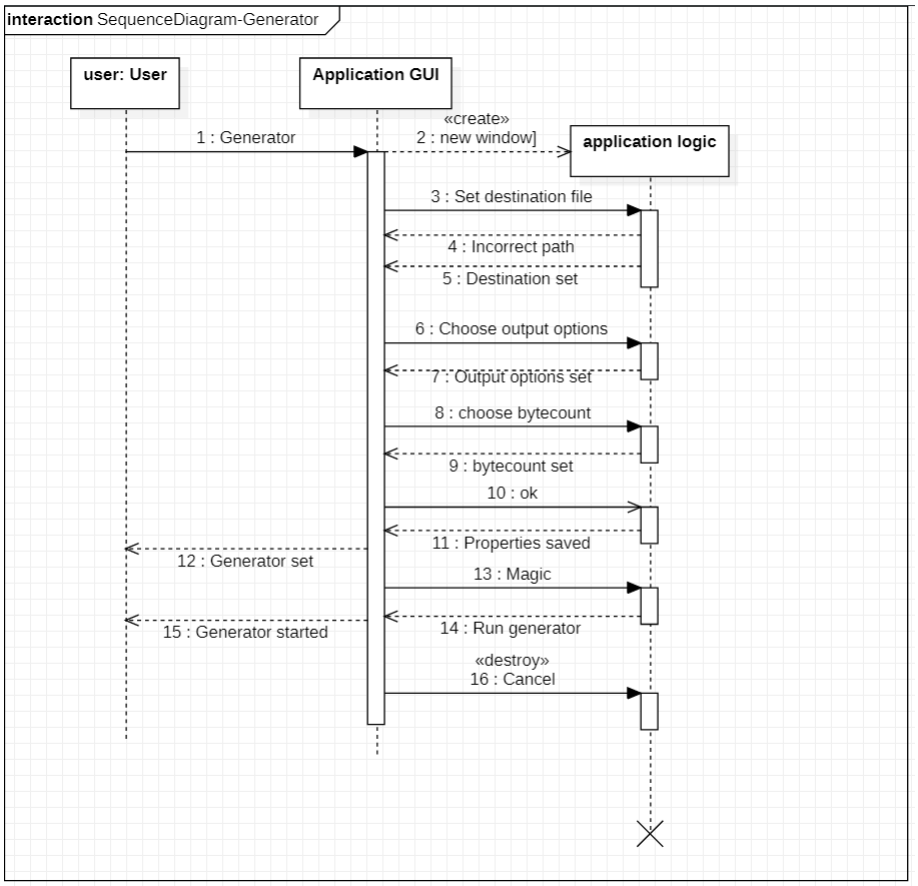

Figure 11: paranoYa - Sequence diagram (Generator)

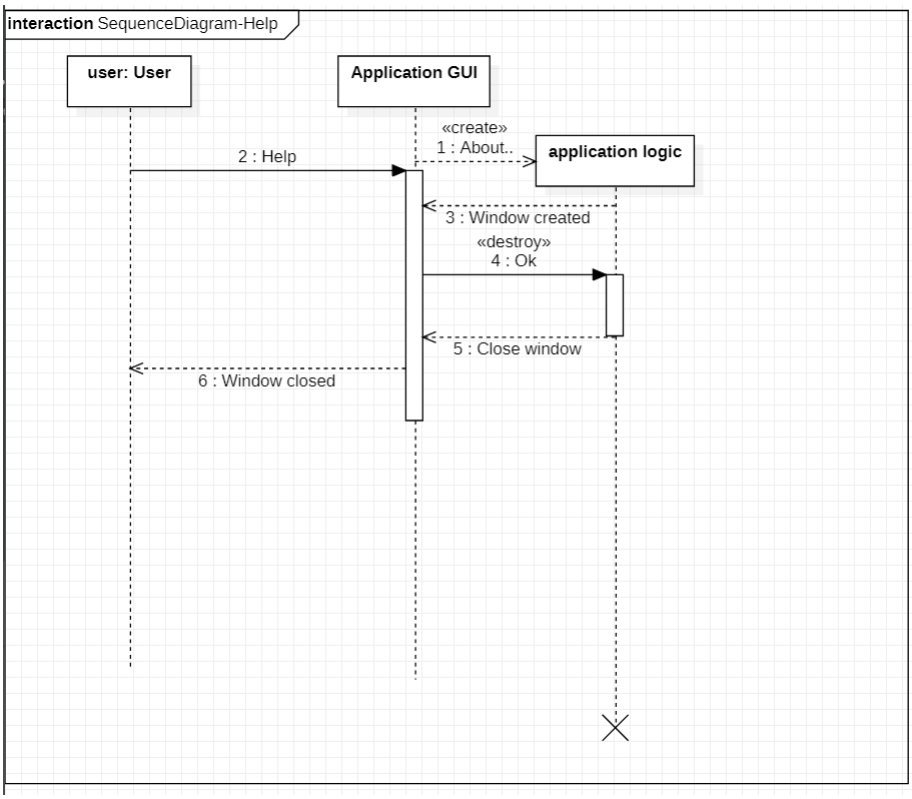

Figure 12: paranoYa - Sequence diagram (Help)

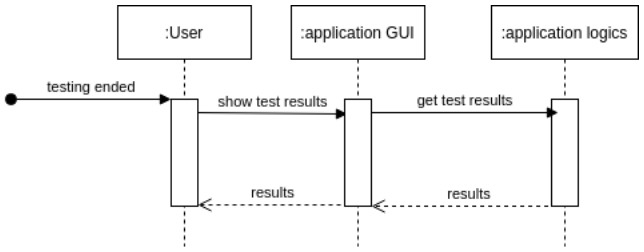

Figure 13: paranoYa - Sequence diagram (Show test results)

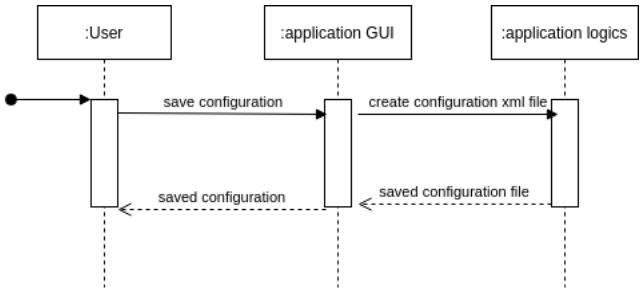

Figure 14: paranoYa - Sequence diagram (Save configuration)

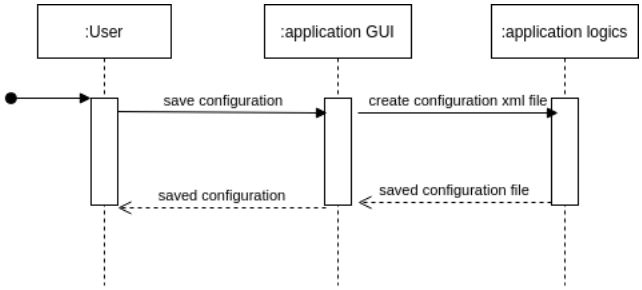

Figure 15: paranoYa - Sequence diagram (Load sequence)

3.1.3 Activity Diagrams

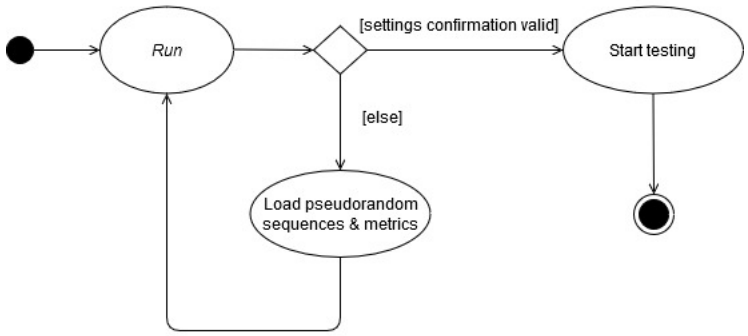

Figure 16: paranoYa - Activity diagram (Run)

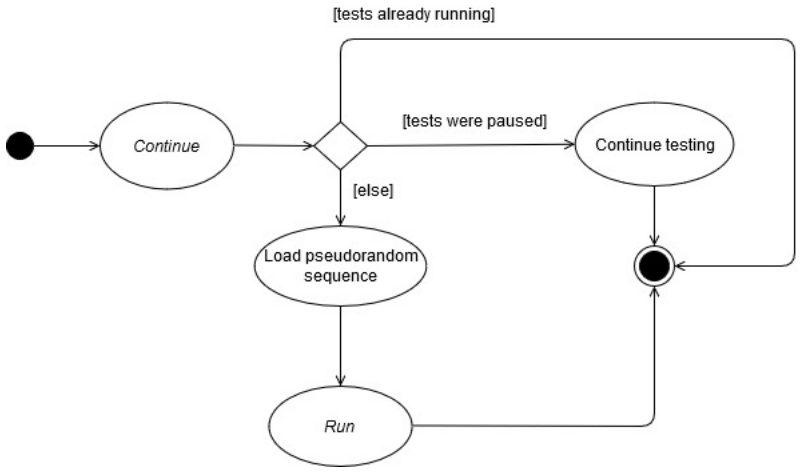

Figure 17: paranoYa - Activity diagram (Continue)

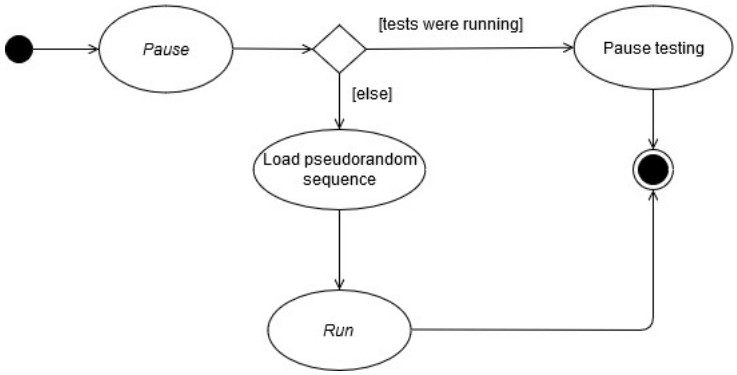

Figure 18: paranoYa - Activity diagram (Pause)

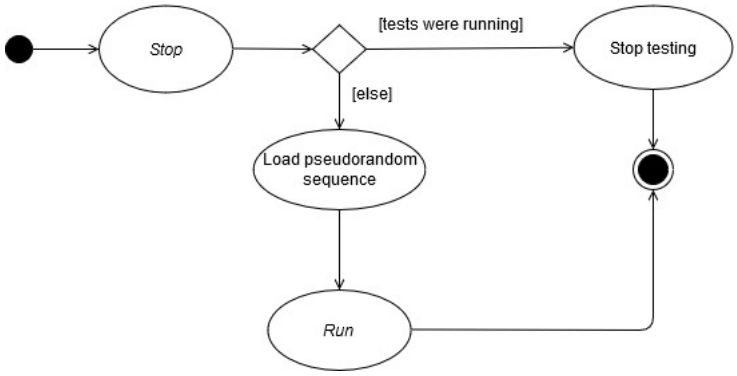

Figure 19: paranoYa - Activity diagram (Stop)

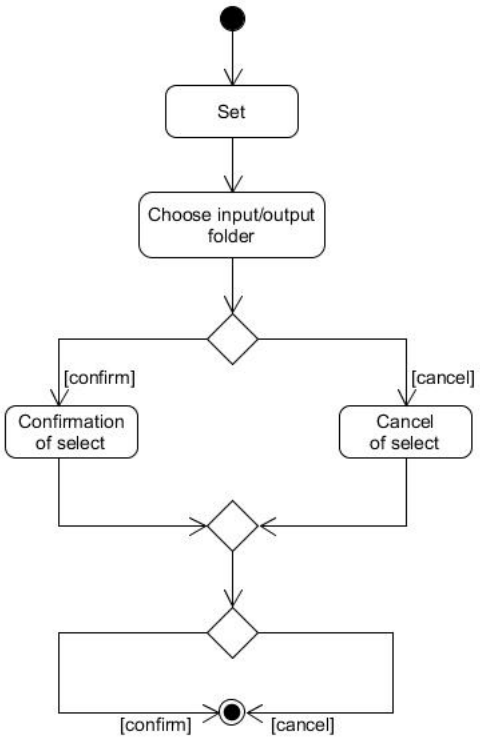

Figure 20: paranoYa - Activity diagram (Set input/output folder)

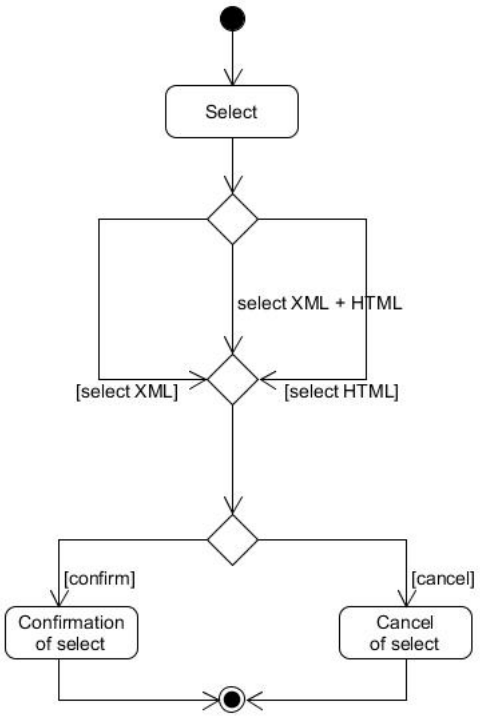

Figure 21: paranoYa - Activity diagram (Select output option)

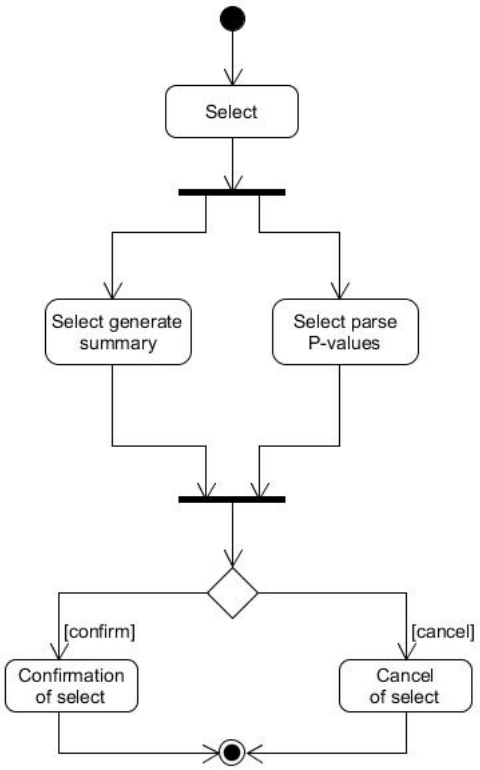

Figure 22: paranoYa - Activity diagram (Select summary option)

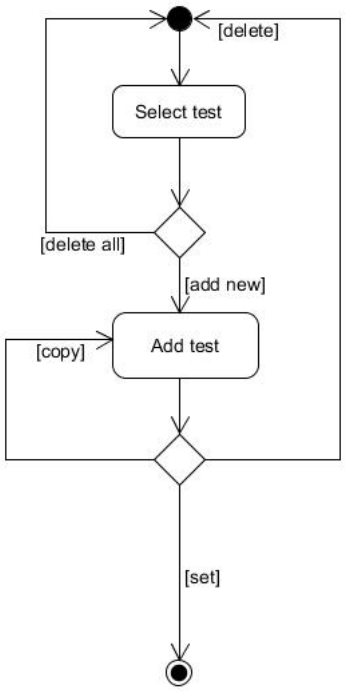

Figure 23: paranoYa - Activity diagram (Set parameters for selected tests)

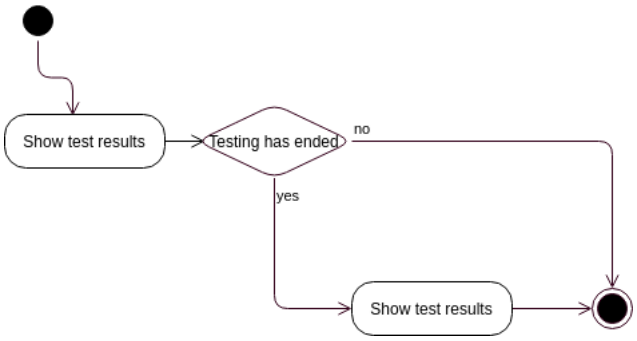

Figure 24: paranoYa - Activity diagram (Show test results)

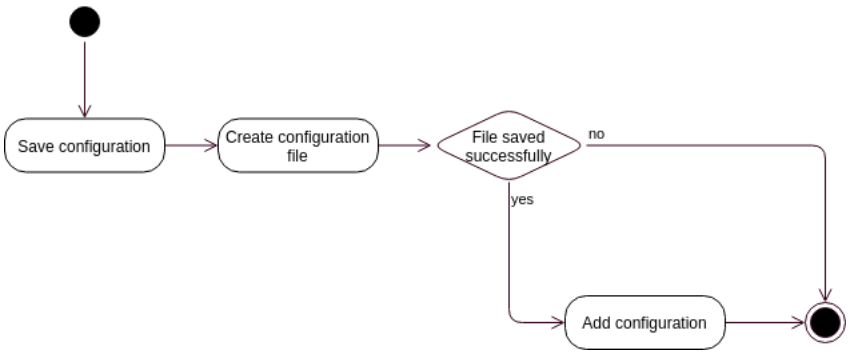

Figure 25: paranoYa - Activity diagram (Save configuration)

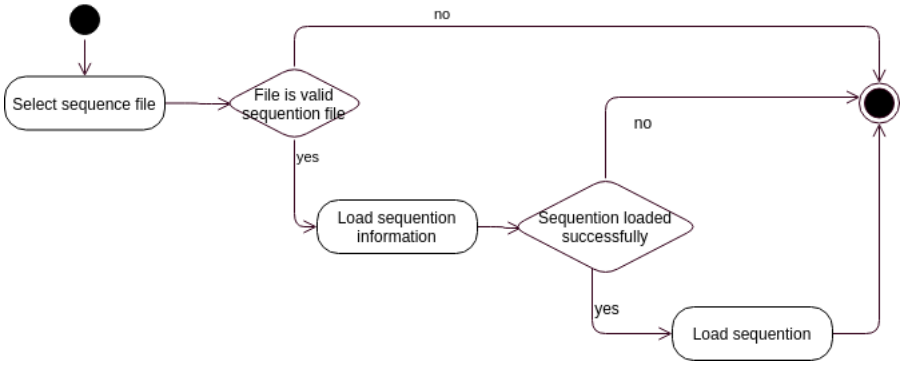

Figure 26: paranoYa - Activity diagram (Load sequence)