NI-KRY - Popis testů v sadě STS od NITS

Tomáš Chvosta (chvostom)

Říjen 2020

Všechny testy byly použity pro soubory o velikosti 5400000 bitů. Pro většinu testů byly tyto soubory rozděleny na 576 podposloupností o velikosti 9375 bitů ($n = 9375$). Pro test číslo 10 byly rozděleny na 12 podposloupností o velikosti 450000 bitů a pro test č.12, č.13 a č.15 na 5 podposloupností o velikosti 1080000 bitů.

Test 1 - Frequency test

Popis testu

Test se zaměřuje na poměr jedniček a nul v celé sekvenci dat. Cílem testu je rozhodnout, zda je počet jedniček a nul přibližně stejný jako je očekáván u náhodné sekvence.

Doporučení

- Každá sekcence by měla mít alespoň 100 bitů.

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů.

Test 2 - Block Frequency test

Popis testu

Obdobný test jako předchozí, ale testují se M-bitové bloky. Sekvence se poté rozdělí na díly $N = n/M$, kde n je celkový počet bitů celé sekvence a M je počet dílů jednotlivých bloků. V každém z bloků musí platit, že počet jedniček je přibližně polovina celkového počtu jeho bitů.

Doporučení

- Každá sekcence by měla mít alespoň 100 bitů.

- $n \geq M \cdot N$
- $M \geq 20$
- $M > 0,01 \cdot n$
- $N < 100$

Zvolené parametry

- $M = 125$

Test 3 - Cumulative Sums test

Popis testu

V tomto testu se vypočítává maximální odchylka od nuly pro kumulativní součet všech bitů, kdy každý průběžný součet má být blízký nule. Bit 1 reprezentuje kladnou jedničku v součtu a bit 0 zápornou jedničku. První součet je tvořen pouze prvním bitem. Pro každý následující bit je hodnota součtu tvořena všemi předcházejícími bity včetně aktuálního bitu. Pro náhodnou sekvenci by se měly součty blížit nule, naopak pro sekvenci, která náhodná není, je odchylka od nuly veliká.

Doporučení

- Každá sekcence by měla mít alespoň 100 bitů.

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů.

Test 4 - Runs test

Popis testu

Cílem testu je celkový počet průběhu, kde průběh znamená nepřerušenou sekvenci stejných znaků. Průběh délky k obsahuje k stejných bitů, které jsou ohraničeny bity opačné hodnoty. Účel testu je určit, zda počet těchto průběhů jedniček a nul s různými délkami je takový, jaký se u náhodných sekvencí očekává. Dochází tedy k rozhodnutí, zda střídání jedniček nebo nul není příliš rychlé, nebo naopak příliš pomalé.

Doporučení

- Každá sekcence by měla mít alespoň 100 bitů.

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů.

Test 5 - Longest Run of Ones test

Popis testu

Význam testů spočívá v kontrole nejdelšího průběhu jedniček v M-bitovém bloku. Účelem je porovnat, zda je délka nejdelšího průběhu jedniček v testované sekvenci stejná, jako délka očekávaná v náhodné sekvenci. Zajímavostí je, že pokud existuje nepravidelnost v délce nejdelšího průběhu jedniček, bude zároveň i nepravidelnost v této délce pro nuly. Proto nám stačí testovat pouze jedničky.

Doporučení

- Každá sekcence by měla mít minimum bitů podle následující tabulky:

minimum n	M
128	8
6272	128
75000	10^4

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů.

Test 6 - Rank test

Popis testu

Test se zaměřuje na lineární závislost mezi podřetězci pevné délky pro celou sekvenci. Toho se dosahuje testováním hodnosti (rank) matice, tj. číslo představující nezávislý počet řádků nebo sloupců matice.

Doporučení

- $n \geq 38 \cdot M \cdot Q$

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů.

Test 7 - Discrete Fourier Transform test

Popis testu

Test pracuje se špičkovými hodnotami (maximální hodnoty amplitudy) v diskrétní Furierově transformaci pro danou sekvenci. Zjišťuje periodické funkce, tj. opakující se vzory nacházející se blízko sebe, které znamenají odchylku od předpokládané náhodnosti. Test počítá počet špičkových hodnot a porovnává zda počet špiček pod prahovou hodnotou je více nebo méně než 95% ze všech špiček.

Doporučení

- Každá sekcence by měla mít alespoň 1000 bitů.

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů.

Test 8 - Nonperiodic Template Matchings test

Popis testu

Smyslem testu je detekce velkého počtu výskytu vzoru m-bitové délky. Pod pojmem vzor je myšlen řetězec bitů, který je právě obsažen v pracovním okně. Okno má velikost m-bitů. Pokud se řetězec v sekvenci nevyskytuje, okno se posouvá o jeden bit. Pokud se řetězec vyskytne, okno se vymaže a posouvá se na řetězec začínající prvním bitem po nalezeném řetězci.

Doporučení

- $m = 9$ nebo $m = 10$

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů (ponecháno $m = 9$).

Test 9 - Overlapping Template Matchings test

Popis testu

Test pracuje na podobném principu jako předchozí test, avšak pokud je hledaný vzor nalezen, celé okno se posouvá jen o jeden bit, místo posunu za řetězec.

Doporučení

- $m = 9$ nebo $m = 10$

Zvolené parametry

Nejsou potřeba žádná nastavení parametrů (ponecháno $m = 9$).

Test 10 - Universal Statistical test

Popis testu

Test kontroluje počet bitů mezi shodnými vzory (opatření vztahující se k délce komprimované sekvence). Účelem je zjistit, zda může být sekvence komprimovaná bez toho, aby ztratila informace. Pokud je možno dosáhnout velké komprimace, značí to porušení vlastností náhodnosti.

Doporučení

- Test vyžaduje dlouhé sekvence
- $n \geq (Q + K) \cdot L$
- $6 \geq L \geq 16$
- $Q = 10 \cdot 2^L$
- $K = \lfloor n/L \rfloor - Q \approx 1000 \cdot 2^L$
- pro $L = 6$ by mělo být $n \geq 387840$

Zvolené parametry

- Změna délky všech podposloupností na 450000. Vznikne tedy 12 podposloupností.

Test 11 - Approximate Entropy test

Popis testu

Zaměření testu je stejné jako u Serial Test. Test porovnává, zda je frekvence překrývajících se bloků (po sobě jdoucích velikostí m a $m+1$) stejná jako očekávaná frekvence pro náhodnou sekvenci.

Doporučení

- $m < \lfloor log_2(n) \rfloor - 5$

Zvolené parametry

- $m < \lfloor log_2(9375) \rfloor - 5$
- $m < 8$
- $m = 5$ (aby nebylo potřeba doplňovat bloky nulami)

Test 12 - Random Excursions test

Popis testu

Počítá počet cyklů, které mají přesně K návštěv v kumulativním součtu pro náhodný průchod. Kumulativní součet náhodným průchodem je odvozen z částečných součtů poté, co je (0, 1) posloupnost převedena do příslušné (-1, + 1) sekvence. Cyklus náhodného průchodu se skládá ze sekvence kroků náhodně vybrané délky, které začínají a vrací se k počátku. Smyslem testu je zjistit, zda se počet návštěv určitého stavu liší, od očekávaných hodnot pro náhodné sekvence. Tento test je ve skutečnosti sérií osmi testů (a závěrů), jeden test a závěr pro každý ze stavů: -4, -3, -2, -1 a +1, +2, +3, +4.

Doporučení

- Každá sekcence by měla mít alespoň 1000000 bitů.

Zvolené parametry

- Změna délky všech podposloupností na 1080000. Vznikne tedy 5 podposloupností.

Test 13 - Random Excursions Variant test

Popis testu

Zaměřuje se na počet návštěv určitého stavu pro kumulativní součet náhodného průchodu. Účelem je určit odchylky od očekávaných počtu návštěv, pro různé stavy v během průchodu. Test je sérií osmnácti testů (a závěrů), jeden test a závěr pro každý ze stavů: -9, -8, ..., -1 a +1, +2, ..., +9.

Doporučení

- Každá sekcence by měla mít alespoň 1000000 bitů.

Zvolené parametry

- Změna délky všech podposloupností na 1080000. Vznikne tedy 5 podposloupností.

Test 14 - Serial test

Popis testu

Testuje frekvenci všech možných překrývajících se m-bitových vzorů v celé sekvenci. Účelem je rozhodnout, zda počet výskytů 2^m m-bitových překrývajících se vzorů je přibližně stejný, jako u náhodné sekvence. Jednotnost náhodných sekvencí je v tom, že každý m-bitový vzor má stejnou pravděpodobnost výskytu jako všechny ostatní. Pro $m = 1$ je test ekvivalentní s Frequency Test.

Doporučení

- $m < \lfloor log_2(n) \rfloor - 2$

Zvolené parametry

- $m < \lfloor log_2(9375) \rfloor - 2$
- $m < 11$
- $m = 5$ (aby nebylo potřeba doplňovat bloky nulami)

Test 15 - Linear Complexity test

Popis testu

Test je zaměřen na délku posuvného registru s lineární zpětnou vazbou (LFSR - linear feedback shift register). Smyslem je rozhodnout, zda je testovaná sekvence dostatečně složitá na to, aby mohla být považována za náhodnou. Náhodné sekvence jsou charakterizovány dlouhými LFSR, kdežto LFSR s krátkou délkou vyvrací jejich náhodnost.

Doporučení

- $n \geq 10^6$
- $500 \leq M \leq 5000$
- $N \geq 200$

Zvolené parametry

- Změna délky všech podposloupností na 1080000. Vznikne tedy 5 podposloupností.
- $M = 1000$