## Simulation

The aim of *simulation* is to tune parameters in the two versions of *MART* (with *static* and *dynamic* *d*), then evaluating their performance compared to the *Simple Random Sampling Without Replacement*. For these purposes, it is necessary to introduce *simulation scenarios* and *evaluation criteria*.

### 3.1 Simulation Scenarios

To obtain the simulation scenarios, it is necessary to consider three main factors about the population and the problem size:

* *Type of partitioning*: it represents how test cases are partitioned in test frame. This information is encoded as the test frame’s failure probability, in particular in a *Correct/Failing* value, for which three settings are considered: 0/1, 0.25/0.75 and 0.1/0.9. This probability is necessary to determine the weight associated with each network’s link. In case 0/1, a test frame believed to be *failing* (*correct*) contains only *failing* (*correct*) test cases, hence its failure probability (as proportion of failing test cases) is 1 (0). This refers to a *Perfect Partitioning*, while if the failure test frames are organized in clusters, a *Perfect Clustered Partitioning* is taken into account; the couples 0.25/0.75 and 0.1/0.9 mean that the partitioning is not accurate, because failures are distributed in all test frames. These probabilities generate the *Close to Uniform Partitioning* (for 0.25/0.75) and *Close to Perfect Partitioning* (0.1/0.9), if failure test frames are organized in cluster, the population distribution is *Clustered*.
* *Failing test frame proportion*: his is the proportion of the failing test frames over the total, for which we consider two values, 0.1 and 0.2.
* *Total number of test frames (N):* two order of magnitudes are tried: N = 100, N = 1000.

The combination of first two factors generates 12 different populations, shown in Table 4, after a *completely uniform population distribution* is added, in which the failure probability of each test frame is obtained as a random value between 0 and 1, as a consequence, failures are uniformly distributed between test frames (this represent an ideal case).

Table 1: Populations

|  |  |  |
| --- | --- | --- |
|  | Type of partitioning | Failing test frame proportion |
| 1 | Uniform (Random) | 0.5 |
| 2 | Close to uniform (0.25/0.75) | 0.1 |
| 3 | Close to perfect partitioning (0.1/0.9) | 0.1 |
| 4 | Clustered (0.25/0.75) | 0.1 |
| 5 | Clustered (0.1/0.9) | 0.1 |
| 6 | Perfect | 0.1 |
| 7 | Perfect Clustered (1/0) | 0.1 |
| 8 | Close to uniform (0.25/0.75) | 0.2 |
| 9 | Close to perfect partitioning (0.1/0.9) | 0.2 |
| 10 | Clustered (0.25/0.75) | 0.2 |
| 11 | Clustered (0.1/0.9) | 0.2 |
| 12 | Perfect | 0.2 |
| 13 | Perfect Clustered | 0.2 |

26 different scenarios are obtained adding to the combinations the last factor. Assessment is made at 9 checkpoints: *n1 = 0.1N, n2 = 0.2N, ... 0.9N*.

Described scenarios, according to all features, are generated with *uniform operational profile*, before the execution of respective simulation’s blocks.

The populations from 1 to 7 are defined in ascended order of compatibility with testing strategy (from worst to best case). The same consideration also applies to populations from 8 to 13, but with different failure distribution.

#### 3.1.1 Population generators

Different population distributions are generated using following functions:

* *generatePopulationAndMatrix*: to generate a random set of test frames with random failure probability, random distance factor (between 0 and *maxdistance*); occurrence probabilities can be random or equiprobable.
* *generatePopulationAndMatrixBinary*: to generate a random set of test Frames with failure probability chosen random between two selected value ((0/1), (0.1/0.9), (0.25/0.75)) respecting a failure proportion (proportion of values with the high probability of failure on the total), random distance factor (between 0 and *maxdistance*) and the probability of occurrence can be random or equiprobable;
* *generatePopulationAndMatrixCluster*: to generate a set of test frames with a policy defined below, random distance factor (between 0 and *maxdistance*) and the probability of occurrence can be random or equiprobable.

The steps used to determine the *failure probability* of *N* test frames set with *clustered distribution* are the following:

1. A *t%* of test frames is considered as failing, the *t%-th* part of *N* is called *X*:
2. The lowest failure probability is assigned to all points;
3. Each Cluster of failure points is made by a certain percentage (in this case 10% and 20% are considered) of *X*’s cardinality, called T;
4. points are chosen randomly as centroids;
5. Finally, for each centroid T minimum distance test frames are chosen assigning them the maximum failure probability.

### 3.2 Evaluation Criteria

*Accuracy* and *efficiency* are considered as evaluation criteria estimated as follows. A simulation scenario j is repeated 100 times; denote with r one of such repetitions. At the end of each repetition, the reliability estimates is computed by the technique under assessment as well as the *true* reliability . For simulation, it is known in advance which input t is a failure point (hence, , where is 1 if the input is a failure point and 0 otherwise).

For each scenario j, the sample mean (denoted as *M*), sample variance (*var*) and mean squared error (*MSE*) are computed:

* ;
* ;
* ;

Comparison of estimation accuracy is done by looking at the *MSE*. Comparison of efficiency is done by the *sample variance*. Lastly, the *average* *number of failing points* (*NFP*) is also considered as a *performance* metric.

### 3.3 Empirical correction of Estimator

The simulations showed that for little values of *n*, *z*i at each pass is strongly influenced both by the order in which test frames are selected and by the number of taken failures.

This behavior is a consequence of assumptions defined in formulation, in fact there is an important approximation in case of “failure probability” between [0,1]: *unreliability* is calculated as where is a binary value, because only a test case is taken from each test frame. A consequence is the possibility that a test case taken from a test frame with “high” failure probability (ex.: 0.9) could not fail and that a test case taken from a test frame with “low” failure probability (ex.: 0.1) could fail, causing underestimation or overestimation respectively. This phenomenon is more evident when are considered little values of n.

For the discussed problem it is more dangerous have a Reliability overestimation than an underestimation, thus the idea is to adjust this value for avoid this condition.

The estimate defined in previous chapter is based on the mean value of the estimates calculated with *zi*, influenced by the presence of "outliers”. This presence is very heavy when little values of n are considered, but the mean is still the most representative value of the set.

The simulation shows that testing strategy takes test frames with the highest failure probability in the initial part of the testing, this implies that the probability to obtain an underestimation of *unreliability* (overestimation of *reliability*) increases.

The idea is to calculate the mean of estimated value in a single algorithm execution and use it to consider only:

* values that differing not later than 90% by mean in overestimation;
* values that differing not later than 10% by mean in underestimation.

In other words, a *window with two limit values* is considered: the upper bound is the mean plus 90% of mean, the lower bound is mean minus 10% of mean.

This idea does not influence the *unbiasedness* of the estimators; in fact, this operation consists in calculating the mean choosing values that better representing the examined population.

This adjustment is not functional with big values of *n*, in fact there is a constant increment of MSE (respect the estimator described in the article [2]).

The final solution consists of combination of two explained estimators, the choice depends on *n* value. It is fundamental define what means little or big values of *n*. It is observed that the adjusted estimator has good performance when *n* is included between 0 and 30/40%. This consideration brings to divide *n* values as described in Figure 13, where: for 0-34% *n* is *little* and the *adjusted estimator* is used*;* for 66-100% *n* is *big* and the estimators defined in previous chapter are used.

  
For *n* among 34% and 66% a linear combination of both estimators is used. This linear combination is realized using a “*coefficient of correction*” defined as: and the estimation is calculated as:

Figure 13: Different Estimator usage

.

Table 5 shows the trend of MSE, in particular it converges to 0 when *n* increases. For this purpose, configurations 1, 2 and 10 are considered.

Table 2: Results of the adjusted estimator based on MSE

|  |
| --- |
|  |
| Configuration 1 |
|  |
| Configuration 2 |
|  |
| Configuration 10 |

### 3.4 Sensitivity Analysis

#### 3.4.1 Sensitivity Analysis of *d* in Static Approach

The chosen value of *d* is taken comparing simulation results, which are obtained considering these four values: 0.2, 0.4, 0.6, 0.8. These represent the trust in the *weight based sampling* compared to the *simple random sampling*.

The sensitivity analysis is realized considering only five configurations 1, 2, 8, 9, that represent the limit cases, and 12, as best case example.

The evaluation criteria are *MSE* and *Sample* *Variance*.

Table 3: MSE comparison between d = 0.2, d = 0.4, d = 0.6 and d = 0.8, considering the most significant configurations

|  |
| --- |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf1/Schermata%202018-04-19%20alle%2011.57.19.png |
| Configuration 1 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf2/Schermata%202018-04-19%20alle%2012.01.17.png |
| Configuration 2 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf8/Schermata%202018-04-19%20alle%2012.03.15.png |
| Configuration 8 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf9/Schermata%202018-04-19%20alle%2012.05.14.png |
| Configuration 9 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf12/Schermata%202018-04-19%20alle%2012.07.51.png |
| Configuration 12 |

From results shown in Table 6, the best value of *d* about *MSE* is 0.8, in particular for little values of *n*.

Table 4: Sample Variance comparison between d = 0.2, d = 0.4, d = 0.6 and d = 0.8, considering the most significant configurations

|  |
| --- |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf1/Schermata%202018-04-19%20alle%2011.57.28.png |
| Configuration 1 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf2/Schermata%202018-04-19%20alle%2012.01.26.png |
| Configuration 2 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf8/Schermata%202018-04-19%20alle%2012.03.22.png |
| Configuration 8 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf9/Schermata%202018-04-19%20alle%2012.05.23.png |
| Configuration 9 |
| Simulazione/Simulazioni%20utilizzate/Analisi%20di%20sensitività%20su%20d%20statico/conf12/Schermata%202018-04-19%20alle%2012.07.59.png |
| Configuration 12 |

As shown in Table 7, the *Sample* *Variance* is few influenced by the *d* value in the different configurations, in fact, except for few cases, the variance values have the same order of magnitude. In the configurations 1 and 2 there are better values for *d* equal to 0.2 and 0.4, but in configuration 9 and 12 there are better values for 0.6 and 0.8.

The chosen value of *d* is 0.8, because it has the best trade-off between MSE and Variance.

#### 3.4.2 Sensitivity Analysis in Dynamic approach

In this case the sensitivity analysis is performed for two different values: *d0* (value between 0.5 and 0.9 with step 0.1) and the *shift register size* (3, 4 and 5).

Sensitivity Analysis on *d0*

The chosen value of *d0* is taken comparing simulation results, which are obtained considering five different values between 0.5 and 0.9 with step 0.1. *MSE*, *Sample* *Variance* and Number of Failing Point (*NFP*) are the used evaluation criteria for the sensitivity analysis.

The sensitivity analysis is realized considering only four configurations 1, 2, 8, 9, that represent the limit cases.

As shown in Tables 8, 9 and 10, it is possible to observe three different trends:

* *MSE* has a downward trend with the increment of *d0*;
* *Sample Variance* is more or less growing with the increment of *d0*;
* *NFP* is almost constant for all *d0* values.

Considerations about the limit cases are also interesting, the first subset, in fact there are substantial differences for different values of d0.

Table 5: Sensitivity Analysis on d0 respect to MSE, where Ad1 is the implementation with dynamic d and initial sample unitary, while Ad2 is the implementation with dynamic d and variable initial sample

|  |
| --- |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%201/Schermata%202018-02-02%20alle%2012.07.21.png |
| Configuration 1 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%202/Schermata%202018-02-02%20alle%2012.10.23.png |
| Configuration 2 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%208/Schermata%202018-02-03%20alle%2010.14.57.png |
| Configuration 8 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%209/Schermata%202018-02-03%20alle%2010.16.37.png |
| Configuration 9 |

How described previously the *MSE* have a downward trend with the increment of *d0*. As shown in Table 8 the best values of *d0* are 0.8 and 0.9.

Table 6: Sensitivity Analysis on d0 respect to Sample Variance, where Ad1 is the implementation with dynamic d and initial sample unitary, while Ad2 is the implementation with dynamic d and variable initial sample

|  |
| --- |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%201/Schermata%202018-02-02%20alle%2012.07.35.png |
| Configuration 1 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%202/Schermata%202018-02-02%20alle%2012.10.34.png |
| Configuration 2 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%208/Schermata%202018-02-03%20alle%2010.15.15.png |
| Configuration 8 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%209/Schermata%202018-02-03%20alle%2010.16.46.png |
| Configuration 9 |

In Table 9 it is shown that *Sample Variance* is better for little values of d0, while 0.9 is the worst case.

Table 7: Sensitivity Analysis on d0 respect to NFP, where Ad1 is the implementation with dynamic d and initial sample unitary, while Ad2 is the implementation with dynamic d and variable initial sample

|  |
| --- |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%201/Schermata%202018-02-02%20alle%2012.07.46.png |
| Configuration 1 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%202/Schermata%202018-02-02%20alle%2012.10.45.png |
| Configuration 2 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%208/Schermata%202018-02-03%20alle%2010.15.30.png |
| Configuration 8 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20diversi%20valori%20di%20d0/screen/Configurazione%209/Schermata%202018-02-03%20alle%2010.16.57.png |
| Configuration 9 |

In case of *NFP* performances, shown in Table 10, are more or less the same for each value of *d0*.

Considering all observations, the selected value of *d0* is 0.8, that offers a good trade-off between *MSE* and *Sample Variance*.

Sensitivity Analysis on Shift Register

The analysis is carried out on three values of shift registers’ dimension: 3, 4 and 5.

The values associated with each cell are organized as follow:

|  |  |  |
| --- | --- | --- |
| 0.5 | 0.3 | 0.2 |

|  |  |  |  |
| --- | --- | --- | --- |
| 0.4 | 0.3 | 0.2 | 0.1 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| 0.4 | 0.3 | 0.15 | 0.1 | 0.05 |

As the previous case, simulations are performed for configurations 1, 2, 8 and 9. About *MSE* and *Sample Variance*, the three configuration are few different.

In case of *MSE* the better value of Shift Register’s dimension is 4, because it gives better performances also in the limit cases.

Table 8: Sensitivity Analysis on Shift Register dimension respect to MSE

|  |
| --- |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%201/Schermata%202018-02-05%20alle%2010.18.28.png |
| Configuration 1 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%202/Schermata%202018-02-05%20alle%2010.20.33.png |
| Configuration 2 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%208/Schermata%202018-02-05%20alle%2010.21.39.png |
| Configuration 8 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%209/Schermata%202018-02-05%20alle%2010.23.06.png |
| Configuration 9 |

As shown in Table 11, in these cases performances are better in cases SR = 3 and 4.

As in the case of *MSE*, *Sample Variance* is more or less the same in all cases, except that in the limit cases.

Table 9: Sensitivity Analysis on Shift Register dimension respect to MSE

|  |
| --- |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%201/Schermata%202018-02-05%20alle%2010.18.37.png |
| Configuration 1 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%202/Schermata%202018-02-05%20alle%2010.20.43.png |
| Configuration 2 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%208/Schermata%202018-02-05%20alle%2010.21.52.png |
| Configuration 8 |
| Simulazione/Analisi%20di%20sentività%20versione%20dinamica/Adaptive%20con%20varie%20dimensioni%20SR/screen/Configuration%209/Schermata%202018-02-05%20alle%2010.23.14.png |
| Configuration 9 |

Observing Table 12, the performances of SR=5 are better, while values 4 and 3 show the same performances. It is important to note that in this case there is an uncertainty linked to the third decimal place.

The chosen value is SR=4, because is the best value in case of MSE, with acceptable values for the variance (2nd best value).

### 3.5 Results

Simulation results are based on six different configurations:

1. Static *d* with *n0=1* (1);
2. Static *d* with *n0≥1* (2);
3. Simple Random Sampling (SRS);
4. Dynamic *d* with *n0=1* (Ad1);
5. Dynamic *d* with *n0≥1* (Ad2).

#### 3.5.1 MSE

To evaluate the difference between the three different approaches it is possible observe the histograms in Table 13. The evaluation of MSE is important to make consideration on how the obtained values deviate from the mean.

Table 10: MSE simulation results for each Configuration

|  |
| --- |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%201/Schermata%202018-01-31%20alle%2016.54.15.p |
| Configuration 1 (a) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%202/Schermata%202018-01-31%20alle%2017.00.31.p |
| Configuration 1 (b) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%203/Schermata%202018-01-31%20alle%2017.04.05.p |
| Configuration 3 (c) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%204/Schermata%202018-01-31%20alle%2017.06.32.p |
| Configuration 4 (d) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%205/Schermata%202018-01-31%20alle%2017.13.22.p |
| Configuration 5 (e) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%206/Schermata%202018-01-31%20alle%2017.15.50.p |
| Configuration 6 (f) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%207/Schermata%202018-01-31%20alle%2017.53.21.p |
| Configuration 7 (g) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%208/Schermata%202018-01-31%20alle%2017.56.08.p |
| Configuration 8 (h) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%209/Schermata%202018-01-31%20alle%2017.59.20.p |
| Configuration 9 (i) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2010/Schermata%202018-01-31%20alle%2018.01.41.p |
| Configuration 10 (j) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2011/Schermata%202018-01-31%20alle%2018.04.46.p |
| Configuration 11 (k) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2012/Schermata%202018-01-31%20alle%2018.06.41.p |
| Configuration 12 (l) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2013/Schermata%202018-01-31%20alle%2018.10.06.p |
| Configuration 13 (m) |

#### 3.5.2 Sample Variance

*Sample Variance* is important to evaluate the goodness of estimate; in particular it underline how much estimate is representative of population mean. All results are shown in Table 14.

Table 11: Sample Variance simulation results for each Configuration

|  |
| --- |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%201/Schermata%202018-01-31%20alle%2016.54.49.p |
| Configuration 1 (a) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%202/Schermata%202018-01-31%20alle%2017.01.17.p |
| Configuration 2 (b) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%203/Schermata%202018-01-31%20alle%2017.04.42.p |
| Configuration 3 (c) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%204/Schermata%202018-01-31%20alle%2017.11.22.p |
| Configuration 4 (d) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%205/Schermata%202018-01-31%20alle%2017.13.44.p |
| Configuration 5 (e) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%206/Schermata%202018-01-31%20alle%2017.16.20.p |
| Configuration 6 (f) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%207/Schermata%202018-01-31%20alle%2017.53.52.p |
| Configuration 7 (g) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%208/Schermata%202018-01-31%20alle%2017.56.32.p |
| Configuration 8 (h) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%209/Schermata%202018-01-31%20alle%2017.59.53.p |
| Configuration 9 (i) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2010/Schermata%202018-01-31%20alle%2018.02.06.p |
| Configuration 10 (j) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2011/Schermata%202018-01-31%20alle%2018.05.14.p |
| Configuration 11 (k) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2012/Schermata%202018-01-31%20alle%2018.07.02.p |
| Configuration 12 (l) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2013/Schermata%202018-01-31%20alle%2018.10.39.p |
| Configuration 13 (m) |

#### 3.3.3 Failing Point Number

This quantity explains the trend of different techniques to expose failures, all results are shown in Table 15.

Table 12: NFP simulation results for each Configuration

|  |
| --- |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%201/Schermata%202018-01-31%20alle%2016.54.38.p |
| Configuration 1 (a) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%202/Schermata%202018-01-31%20alle%2017.00.43.p |
| Configuration 2 (b) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%203/Schermata%202018-01-31%20alle%2017.04.17.p |
| Configuration 3 (c) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%204/Schermata%202018-01-31%20alle%2017.11.09.p |
| Configuration 4 (d) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%205/Schermata%202018-01-31%20alle%2017.13.32.p |
| Configuration 5 (e) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%206/Schermata%202018-01-31%20alle%2017.16.03.p |
| Configuration 6 (f) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%207/Schermata%202018-01-31%20alle%2017.53.35.p |
| Configuration 7 (g) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%208/Schermata%202018-01-31%20alle%2017.56.20.p |
| Configuration 8 (h) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%209/Schermata%202018-01-31%20alle%2017.59.42.p |
| Configuration 9 (i) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2010/Schermata%202018-01-31%20alle%2018.01.54.p |
| Configuration 10 (j) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2011/Schermata%202018-01-31%20alle%2018.04.58.p |
| Configuration 11 (k) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2012/Schermata%202018-01-31%20alle%2018.06.50.p |
| Configuration 12 (l) |
| ../../Desktop/UNINA/TESI/screen%20Simulazione/Configurazione%2013/Schermata%202018-01-31%20alle%2018.10.26.p |
| Configuration 13 (m) |

#### 3.3.4 Considerations

The first consideration is about efficiency and accuracy of *SRS* compared with the four different testing strategy versions. For this purpose, configurations 1, 2 and 8, represented in graphs (a), (b) and (h), are considered. In these configurations *MSE* is better in *SRS*, this result depends on the uniform distribution of failures (or close to uniform distribution in case 0.25/0.75), that is an ideal case away from the real world. On the other hand, *SRS* is worse than four versions of test generation algorithm of *MART* about *sample variance*, where the difference is very strong.

For all other configurations our techniques are better than *SRS* both about *sample variance* and except for few isolated point about *MSE*.

At last the *NFP* values are considered, where the four implementations of test generation algorithm are globally better.

Now is useful to verify what is the better between four versions of test generation algorithm of *MART*. The first step is the comparison between the different “initial sample dimensions”. Results show that two techniques with *unitary initial sample* are globally better, both for *MSE* than for *sample variance*.

This consideration brings to the comparison between technique with *static d* and technique with *dynamic d* both with *unitary sample size*, knowing that the differences are very slight. For *MSE*, values are more or less the same, in some configuration i.e. 5 the *dynamic d* is better, while there are other configurations, i.e. 6 in which the *static d* is better. For Variance the *static d* is meanly better.

In conclusion, the *test generation algorithm of MART with static d* is considered better, not only for *Sample Variance*, but also for its simpler formulation than dynamic one.