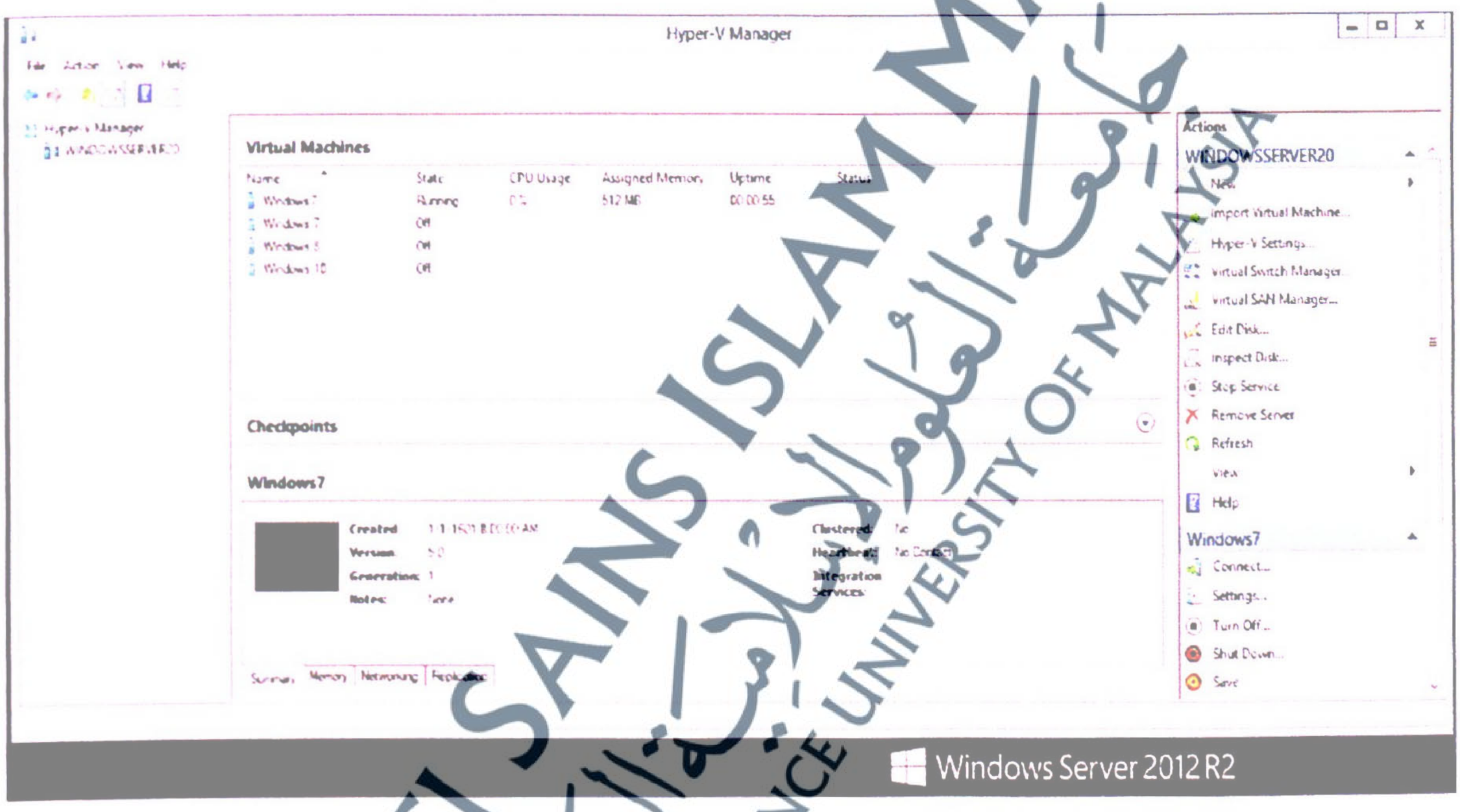
APPENDICES

APPENDIX A

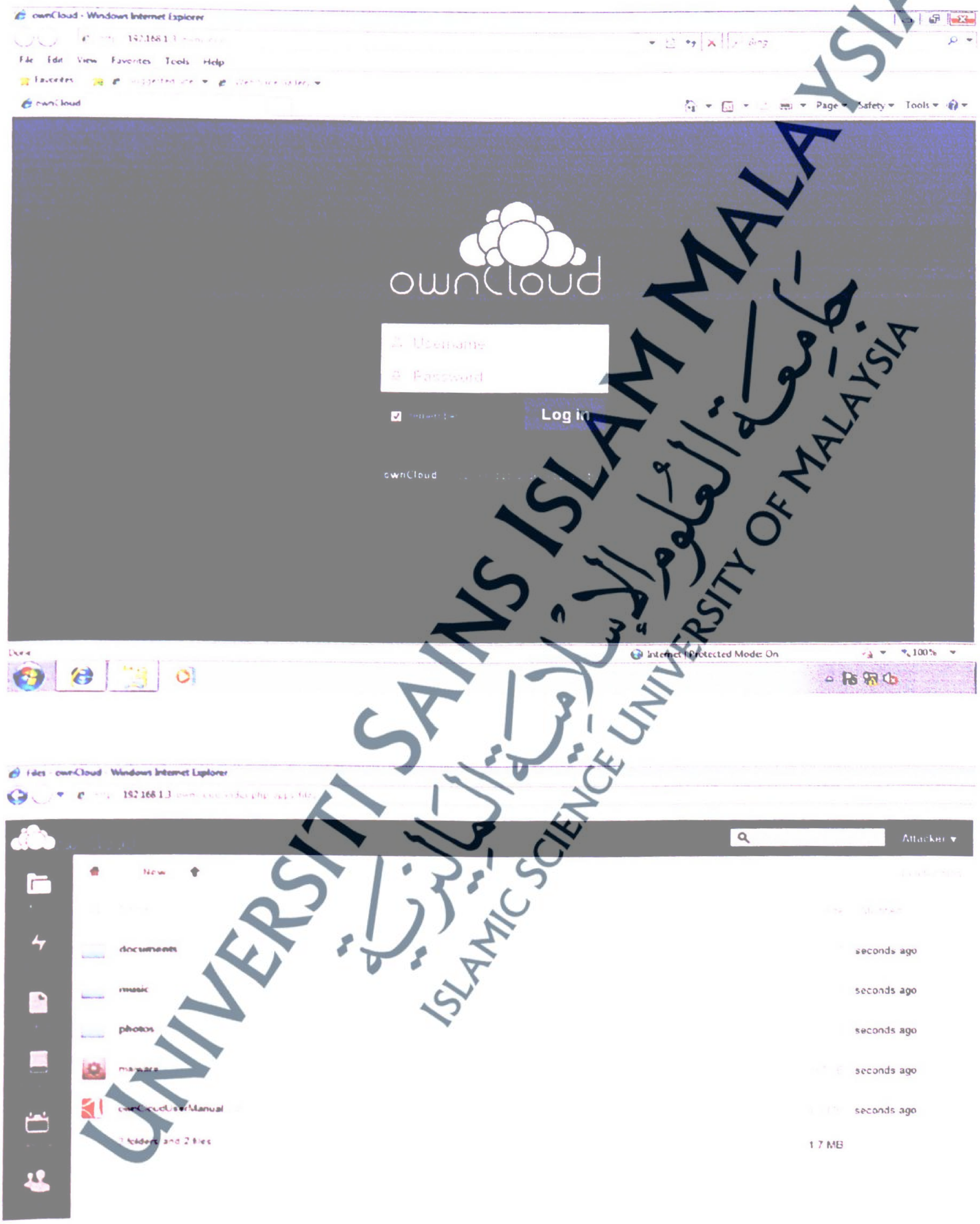
Worm Analysis Process

A.1: Cloud Server



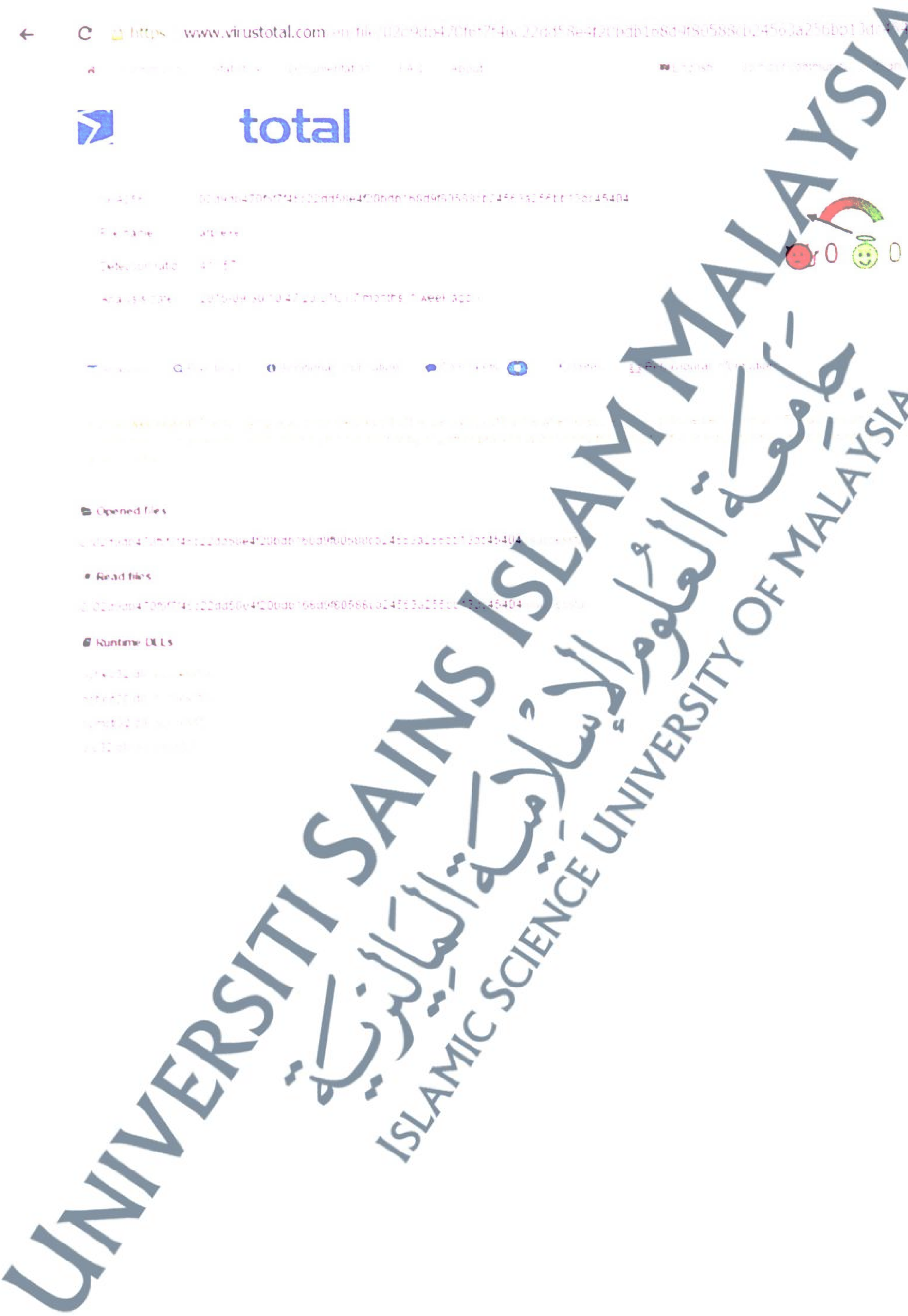
UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

A.2: Own Cloud Platform



UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA
جامعة العلوم الإسلامية الماليزية
YSIA

A.3: Web based Analysis Tool



UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية الماليزية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

APPENDIX B

Chi-Squareand SymmetricMeasure Results

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Infection

Finding 1. Results for relationship between Root privilege and Other users resources.

Rootprivilege * Otherusersresources Crosstabulation

			Otherusersresources		Total
			Yes	No	
Rootprivilege	Yes	Count	0	411	411
		Expected Count	245.1	165.9	411.0
		% within Rootprivilege	0.0%	100.0%	100.0%
		% within Otherusersresources	0.0%	100.0%	40.4%
		% of Total	0.0%	40.4%	40.4%
	No	Count	607	0	607
		Expected Count	361.9	245.1	607.0
		% within Rootprivilege	100.0%	0.0%	100.0%
		% within Otherusersresources	100.0%	0.0%	59.6%
		% of Total	59.6%	0.0%	59.6%
Total		Count	607	411	1018
		Expected Count	607.0	411.0	1018.0
		% within Rootprivilege	59.6%	40.4%	100.0%
		% within Otherusersresources	100.0%	100.0%	100.0%
		% of Total	59.6%	40.4%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	1018.000[a]	1	.000		
Continuity Correction[b]	1013.850	1	.000		
Likelihood Ratio	1373.274	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	1017.000	1	.000		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 165.93.
b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	-1.000-	.000
	Cramer's V	1.000	.000
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Finding 2.Results for relationship between Root privilegeand Hypervisor

Rootprivilege* Hypervisor Crosstabulation

			Hypervisor		Total
			Yes	No	
Rootprivilege	Yes	Count	393	18	411
		Expected Count	339 9	71 1	411 0
		% within Rootprivilege	95 6%	4 4%	100 0%
		% within Hypervisor	46 7%	10 2%	40 4%
		% of Total	38 6%	1 8%	40 4%
	No	Count	449	158	607
		Expected Count	502 1	104 9	607 0
		% within Rootprivilege	74 0%	26 0%	100 0%
		% within Hypervisor	53 3%	89 8%	59 6%
		% of Total	44 1%	15 5%	59 6%
Total		Count	842	176	1018
		Expected Count	842 0	176 0	1018 0
		% within Rootprivilege	82 7%	17 3%	100 0%
		% within Hypervisor	100 0%	100 0%	100 0%
		% of Total	82 7%	17 3%	100 0%

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)	Exact Sig (2-sided)	Exact Sig (1-sided)
Pearson Chi-Square	80 329[a]	1	000		
Continuity Correction[b]	78 822	1	000		
Likelihood Ratio	93 565	1	000		
Fisher's Exact Test				000	000
Linear-by-Linear Association	80 250	1	000		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5 The minimum expected count is 71 06
b Computed only for a 2x2 table

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	281	000
	Cramer's V	281	000
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Activation

Finding 1 .Results for relationship between HummanTriger and Scheduled Process

HummanTriger * ScheduledProcess Crosstabulation

			ScheduledProcess		Total
			Yes	No	
HummanTriger	Yes	Count	456	539	995
		Expected Count	459 4	535 6	995 0
		% within HummanTriger	45 8%	54 2%	100 0%
		% within ScheduledProcess	97 0%	98 4%	97 7%
	No	Count	14	9	23
		Expected Count	10 6	12 4	23 0
		% within HummanTriger	60 9%	39 1%	100 0%
		% within ScheduledProcess	3 0%	1 6%	2 3%
Total		Count	470	548	1018
		Expected Count	470 0	548 0	1018 0
		% within HummanTriger	46 2%	53 8%	100 0%
		% within ScheduledProcess	100 0%	100 0%	100 0%

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)	Exact Sig (2-sided)	Exact Sig (1-sided)
Pearson Chi-Square	2 046[a]	1	153		
Continuity Correction[b]	1 486	1	223		
Likelihood Ratio	2 045	1	153		
Fisher's Exact Test				204	112
Linear-by-Linear Association	2 044	1	153		
N of Valid Cases	1018				

a. 0 cells (0 0%) have expected count less than 5 The minimum expected count is 10 62

b Computed only for a 2x2 table

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	- 045-	153
	Cramer's V	045	153
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Finding 2. Results for relationship between ScheduledProcess and Self Activation

ScheduledProcess * SelfActivation Crosstabulation

			SelfActivation		Total
			Yes	No	
ScheduledProcess	Yes	Count	449	21	470
		Expected Count	423 8	46 2	470 0
		% within ScheduledProcess	95 5%	4 5%	100 0%
		% within SelfActivation	48 9%	21 0%	46 2%
		% of Total	44 1%	2 1%	46 2%
	No	Count	469	79	548
		Expected Count	494 2	53 8	548 0
		% within ScheduledProcess	85 6%	14 4%	100 0%
		% within SelfActivation	51 1%	79 0%	53 8%
		% of Total	46 1%	7 8%	53 8%
Total		Count	918	100	1018
		Expected Count	918 0	100 0	1018 0
		% within ScheduledProcess	90 2%	9 8%	100 0%
		% within SelfActivation	100 0%	100 0%	100 0%
		% of Total	90 2%	9 8%	100 0%

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)	Exact Sig (2-sided)	Exact Sig (1-sided)
Pearson Chi-Square	28 265[a]	1	000		
Continuity Correction[b]	27 153	1	000		
Likelihood Ratio	30 292	1	000		
Fisher's Exact Test				000	000
Linear-by-Linear Association	28 237	1	000		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5 The minimum expected count is 46 17
b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	167	000
	Cramer's V	167	000
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Payload

Finding 1. Results for relationship between Destructiveand Steal Information.

Destructive * StealInformation Crosstabulation

			StealInformation		Total
			Yes	No	
Destructive	Yes	Count	572	15	587
		Expected Count	537.4	49.6	587.0
		% within Destructive	97.4%	2.6%	100.0%
		% within StealInformation	61.4%	17.4%	57.7%
		% of Total	56.2%	1.5%	57.7%
	No	Count	360	71	431
		Expected Count	394.6	36.4	431.0
		% within Destructive	83.5%	16.5%	100.0%
		% within StealInformation	38.6%	82.6%	42.3%
		% of Total	35.4%	7.0%	42.3%
Total		Count	932	86	1018
		Expected Count	932.0	86.0	1018.0
		% within Destructive	91.6%	8.4%	100.0%
		% within StealInformation	100.0%	100.0%	100.0%
		% of Total	91.6%	8.4%	100.0%

Chi-Square Tests

	Value	df	Asymp. Sig. (2-sided)	Exact Sig. (2-sided)	Exact Sig. (1-sided)
Pearson Chi-Square	62.244[a]	1	.000		
Continuity Correction[b]	60.458	1	.000		
Likelihood Ratio	64.263	1	.000		
Fisher's Exact Test				.000	.000
Linear-by-Linear Association	62.183	1	.000		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 36.41.
b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approx. Sig.
Nominal by Nominal	Phi	.247	.000
	Cramer's V	.247	.000
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Finding 2. Results for relationship between Backdoorand Destructive

Backdoor * Destructive Crosstabulation

			Destructive		Total
			Yes	No	
Backdoor	Yes	Count	482	169	651
		Expected Count	375 4	275 6	651 0
		% within Backdoor	74 0%	26 0%	100 0%
		% within Destructive	82 1%	39 2%	63 9%
		% of Total	47 3%	16 6%	63 9%
	No	Count	105	262	367
		Expected Count	211 6	155 4	367 0
		% within Backdoor	28 6%	71 4%	100 0%
		% within Destructive	17 9%	60 8%	36 1%
		% of Total	10 3%	25 7%	36 1%
Total		Count	587	431	1018
		Expected Count	587 0	431 0	1018 0
		% within Backdoor	57 7%	42 3%	100 0%
		% within Destructive	100 0%	100 0%	100 0%
		% of Total	57 7%	42 3%	100 0%

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)	Exact Sig (2-sided)	Exact Sig (1-sided)
Pearson Chi-Square	198 407[a]	1	000		
Continuity Correction[b]	196 550	1	000		
Likelihood Ratio	202 280	1	000		
Fisher's Exact Test				000	000
Linear-by-Linear Association	198 212	1	000		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5 The minimum expected count is 155 38

b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	441	000
	Cramer's V	441	000
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Operating Algorthim

Finding 1. Results for relationship between Stealthand Polymorphic

Stealth * Polymorphic Crosstabulation

			Polymorphic		Total
			Yes	No	
Stealth	Yes	Count	440	545	985
		Expected Count	430 6	554 4	985 0
		% within Stealth	44 7%	55 3%	100 0%
		% within Polymorphic	98 9%	95 1%	96 8%
		% of Total	43 2%	53 5%	96 8%
	No	Count	5	28	33
		Expected Count	14 4	18 6	33 0
		% within Stealth	15 2%	84 8%	100 0%
		% within Polymorphic	1 1%	4 9%	3 2%
		% of Total	0 5%	2 8%	3 2%
Total		Count	445	573	1018
		Expected Count	445 0	573 0	1018 0
		% within Stealth	43 7%	56 3%	100 0%
		% within Polymorphic	100 0%	100 0%	100 0%
		% of Total	43 7%	56 3%	100 0%

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)	Exact Sig (2-sided)	Exact Sig (1-sided)
Pearson Chi-Square	11 308[a]	1	001		
Continuity Correction[b]	10 140	1	001		
Likelihood Ratio	12 753	1	000		
Fisher's Exact Test				001	000
Linear-by-Linear Association	11 297	1	001		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5 The minimum expected count is 14 43
b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	105	001
	Cramer's V	105	001
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

Finding 2. Results for relationship between Stealthand Anti anti- virus.

Stealth * AntiantiVirus Crosstabulation

			AntiantiVirus		Total
			Yes	No	
Stealth	Yes	Count	789	196	985
		Expected Count	787 6	197 4	985 0
		% within Stealth	80 1%	19 9%	100 0%
		% within AntiantiVirus	96 9%	96 1%	96 8%
		% of Total	77 5%	19 3%	96 8%
	No	Count	25	8	33
		Expected Count	26 4	6 6	33 0
		% within Stealth	75 8%	24 2%	100 0%
		% within AntiantiVirus	3 1%	3 9%	3 2%
		% of Total	2 5%	0 8%	3 2%
Total		Count	814	204	1018
		Expected Count	814 0	204 0	1018 0
		% within Stealth	80 0%	20 0%	100 0%
		% within AntiantiVirus	100 0%	100 0%	100 0%
		% of Total	80 0%	20 0%	100 0%

Chi-Square Tests

	Value	df	Asymp Sig (2-sided)	Exact Sig (2-sided)	Exact Sig (1-sided)
Pearson Chi-Square	376[a]	1	540		
Continuity Correction[b]	154	1	695		
Likelihood Ratio	359	1	549		
Fisher's Exact Test				511	335
Linear-by-Linear Association	376	1	540		
N of Valid Cases	1018				

a. 0 cells (0.0%) have expected count less than 5. The minimum expected count is 6 61

b. Computed only for a 2x2 table

Symmetric Measures

		Value	Approx Sig
Nominal by Nominal	Phi	019	540
	Cramer's V	019	540
N of Valid Cases		1018	

UNIVERSITI SAINS ISLAM MALAYSIA
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

APPENDIX C

EGATechniquefor Worm Detection

Results

E.1: Naïve Bayes Results from Weka

```
=== Run information ===

Scheme:weka.classifiers.bayes.NaiveBayes
Relation:     Classification_result
Instances:    1195
Attributes:   6
              Infection
              Activation
              Payload
              Operation algorithm
              Propagation
              worm
Test mode:10-fold cross-validation

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1184               99.0795 %
Incorrectly Classified Instances        11                0.9205 %
Kappa statistic                          0.9634
Mean absolute error                      0.0092
Root mean squared error                  0.0959
Relative absolute error                  3.641  %
Root relative squared error             27.0094 %
Total Number of Instances             1195

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.995    0.034      0.994     0.995     0.995      0.981    worm
                 0.966    0.005      0.972     0.966     0.969      0.981    Benign
Weighted Avg.    0.991    0.03       0.991     0.991     0.991      0.981

=== Confusion Matrix ===

    a    b   <-- classified as
 1013    5 |    a = worm
    6  171 |    b = Benign
```

UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية الماليزية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

E.2: J48 Results from Weka

```
=== Run information ===

Scheme:weka.classifiers.trees.J48 -C 0.25 -M 2
Relation:     Classification_result
Instances:    1195
Attributes:   6
              Infection
              Activation
              Payload
              Operation algorithm
              Propagation
              worm
Test mode:10-fold cross-validation

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1182               98.9121 %
Incorrectly Classified Instances        13                1.0879 %
Kappa statistic                          0.9576
Mean absolute error                      0.0138
Root mean squared error                  0.0939
Relative absolute error                  5.4664 %
Root relative squared error             26.4459 %
Total Number of Instances             1195

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.99      0.017       0.997     0.99      0.994      0.998    worm
                 0.983     0.01        0.946     0.983     0.964      0.998    Benign
Weighted Avg.    0.989     0.016       0.989     0.989     0.989      0.998

=== Confusion Matrix ===

    a     b   <-- classified as
 1008    10 |     a = worm
    3   174 |     b = Benign
```

UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية الماليزية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

E.3: IBK Results from Weka

```
=== Run information ===

Scheme:weka.classifiers.lazy.IBk -K 1 -W 0 -A"weka.core.neighboursearch.LinearNNSearch
-A \"weka.core.EuclideanDistance -R first-last\""

Relation:     Classification_result
Instances:    1195
Attributes:   6
              Infection
              Activation
              Payload
              Operation algorithm
              Propagation
              worm
Test mode:10-fold cross-validation

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1185               99.1632 %
Incorrectly Classified Instances        10                0.8368 %
Kappa statistic                          0.9674
Mean absolute error                      0.0121
Root mean squared error                  0.0753
Relative absolute error                  4.7874 %
Root relative squared error             21.2049 %
Total Number of Instances             1195

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 0.991     0.006       0.999     0.991     0.995      1        worm
                 0.994     0.009       0.951     0.994     0.972      1        Benign
Weighted Avg.    0.992     0.006       0.992     0.992     0.992      1

=== Confusion Matrix ===

    a     b   <-- classified as
 1009     9 |     a = worm
    1   176 |     b = Benign
```

UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية الماليزية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

E.4: OlexGA Results from Weka

```
=== Run information ===

Scheme:weka.classifiers.rules.OlexGA -Y 0 -P 60 -X 0 -R 1.0 -M 0.001 -S 1 -I 500 -G
200 -A 1 -E 0.2 -C 1
Relation:     Classification_result
Instances:    1195
Attributes:   6
              Infection
              Activation
              Payload
              Operation algorithm
              Propagation
              worm
Test mode:10-fold cross-validation

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances         845               70.7113 %
Incorrectly Classified Instances       350               29.2887 %
Kappa statistic                          0.3612
Mean absolute error                      0.2929
Root mean squared error                  0.5412
Relative absolute error                115.8487 %
Root relative squared error            152.3533 %
Total Number of Instances             1195

=== Detailed Accuracy By Class ===

                 TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                   0.656     0          1         0.656     0.792      0.828    worm
                   1         0.344      0.336     1         0.503      0.828    Benign
Weighted Avg.      0.707     0.051      0.902     0.707     0.75       0.828

=== Confusion Matrix ===

   a   b   <-- classified as
 668 350 |   a = worm
   0 177 |   b = Benign
```

UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية الماليزية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA

E.5: EGA Results from Weka

```
=== Run information ===

Scheme:weka.classifiers.rules.EGA
Relation:     Classification_result
Instances:    1195
Attributes:   6
              Infection
              Activation
              Payload
              Operation algorithm
              Propagation
              worm
Test mode:10-fold cross-validation

=== Stratified cross-validation ===
=== Summary ===

Correctly Classified Instances        1192               99.749  %
Incorrectly Classified Instances         3                0.251  %
Kappa statistic                          0.99
Mean absolute error                      0.0025
Root mean squared error                  0.0501
Relative absolute error                  0.993  %
Root relative squared error             14.1052 %
Total Number of Instances             1195

=== Detailed Accuracy By Class ===

               TP Rate   FP Rate   Precision   Recall  F-Measure   ROC Area  Class
                 1         0.017      0.997       1         0.999      0.992    worm
                 0.983     0          1           0.983     0.991      0.992    Benign
Weighted Avg.    0.997     0.014      0.997       0.997     0.997      0.992

=== Confusion Matrix ===

    a    b   <-- classified as
 1018    0 |    a = worm
    3  174 |    b = Benign
```

UNIVERSITI SAINS ISLAM MALAYSIA
جامعة العلوم الإسلامية الماليزية
ISLAMIC SCIENCE UNIVERSITY OF MALAYSIA