

Project Completion Report

Development of a Defence System against Polymorphic and Metamorphic Internet Worms for Enterprise Networks

Submitted

by

Prof. D. K. Saikia, Chief Investigator,

Dr. N. Sarma, Co-investigator,

Mr. S. S. Satapathy, Co-investigator



DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

TEZPUR UNIVERSITY

Napaam, Tezpur - 784028, Assam

September, 2011

COMPLETION REPORT

PART - 1

1. Title of the project : *Development of a Defence System against Polymorphic and Metamorphic Internet Worms for Enterprise Networks.*
2. Implementing Organisation : *Department of Computer Sc. & Engineering, Tezpur University.*
3. DIT Sanction No. and Date : *12(3)/08-ESD dt. 16th October, 2008.*
- 4 (a) Total Budget Outlay : *Original: Rs. 47.61 lakh Revised, if any*
- (b) Duration of project : *2 years.*
- (c) Date of completion and reasons for : *30th Sept, 2011.*
delay, if any : *Period of the project was extended by 11 months to allow for fine tuning and testing.*
5. Total funds spent under various approved : *Enclosed*
budgetary Heads/actual expenditure. Reasons
for deviation, if any (as per enclosed Table 1)
6. Details of equipment/assets acquired out of : *Enclosed*
DIT funds with the name of equipment,
sources of supply, total cost/whether Indian
or imported (as per enclosed Table 2 and 2A)
7. Details of manpower associated with the : *Enclosed*
project (as per enclosed Table 3)
8. Details of yearwise audited statement of : *Enclosed*
accounts and utilization certificates submitted
to DIT (as per G.F.R.19 & 19A)

**TABLE : 1 HEADWISE BREAK-UP OF EXPENDITURE
(Rs. In Lakhs)**

Sl. No	Head	Approved Budget Outlay (Rs.)	Amount Released (Rs.)	Expenditure Incurred up to end of the FY (2009-10) (Rs.)	Expenditure Incurred during the FY 2010-11	Expenditure Incurred during the FY 2011-12	Total Expenditure as on 30 th Sept, 2011	Balance (Rs.)	Remarks
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	
1.	Capital Equipment (including software)	13,30,000/-	13,30,000/-	13,26,248/-	3,500/-	-	13,29,748/-	252/-	
2.	Consumable Items / components	3,50,000/-	3,50,000/-	36,760/-	1,095/-	3,12,145/-	3,50,000/-	Nil	
3.	Duty on Imports	Nil	Nil	Nil	Nil	Nil	Nil	Nil	
4.	Manpower	11,28,000/-	11,28,000/-	6,17,547/-	4,23,363/-	42,500/-	10,83,410/-	44,590/-	
5.	Travel	5,00,000/-	3,83,346/-	48,801/-	31,605/-	2,90,738/-	3,71,144/-	12,202/-	
6.	Contingencies	5,00,000/-	5,00,000/-	37,458/-	51,404/-	4,11,138/-	5,00,000/-	Nil	
7.	Overheads	9,53,000/-	9,22,837/-	5,16,704/-	1,27,742/-	2,64,129/-	9,08,575/-	14,262/-	
8.	Other expenditure debitable to this project	Nil	Nil	Nil	Nil	Nil	Nil	Nil	
	Total -	47,61,000/-	46,14,183/-	25,83,518/-	6,38,709/-	13,20,650/-	45,42,877/-	71,306/-	

**TABLE 2 : CAPITAL EQUIPMENT PROCURED FOR THE PROJECT
(Rs. In Lakhs)**

Sl. No	Description	Manufacturer/ Supplier	Brief Specifications	Purchase Order No.&Date	Date of Receipt	Total Cost (Rs.)	Duty Paid, If any	Conditions G-Good B-Bad#
1.	Workstation - 2 nos.	HP/ BMG Informatics, Ghy	HP XW-6600 Intel Core 2 Quad	TU/11-55/Pur/CSE/2008/ 7142 Dtd. 03-03-09	17-06-09	3,35,000/-	VAT 4% (HW) 12.36% (S/W)	Good
2.	PC – 7 nos.	HCL Infosystems	HCL Infinity BL 1295	TU/11-55/Pur/CSE/2008/ 7143 Dtd. 03-03-09	19-06-09	2,32,316/-	VAT 4% (HW) 12.36% (S/W)	Good
3.	Router – 1 no.	CISCO/ HCL Infosys	CISCO 2821 with security bundle			1,86,703/-	Good	
4.	L2 Switches - 2 nos.	CISCO/ Wipro	CISCO Catalyst 3560	TU/11-55/Pur/CSE/2008/ 7144 Dtd. 03-03-09	19-06-09	3,80,671/-	4% VAT 14,641/-	Good
5.	Laptop PC - 1 no.	DELL/ Cyber Space	Vosto 1510	TU/11-55/Pur/CSE/2008/ 7145 Dtd. 03-03-09	08-04-09	1,74,928/-	4% VAT 6,728/-	Good
6.	Tablet PC - 1 no.	HP/ Cyber Space	HP 2730-P					Good
7.	Laser Printer - 1 no.	Samsung/ Cyber Space	ML2851 ND	TU/11-55/Pur/CSE/2008/ 252 Dtd. 13-04-09	27-05-09	16,630/-	4% VAT 639/-	Good
8.	Red Hat Linux				Included in Workstations & PCs			Good
9.	Windows Vista							Good
10	IDA Pro	Hex-Rays SA Belgium	-	TU/Fin/Project/60-89/08/202 dt. 18-09-2009	19-10-09	21,413/-	-	Good

**TABLE 2.A SALE/TRANSFER OF CAPITAL GOODS
(WITH PRIOR PERMISSION OF DIT)**

NONE

TABLE 3 : MANPOWER ASSOCIATED WITH THE PROJECT :

Sl. No.	Name	Designation	Qualification	% of time devoted to this project	Salary Drawn From the Project Funds (Y/N)	Date of Joining	Date of Leaving	Total Average Emoluments (Monthly) (Rs.)
1.	Prof. D. K. Saikia	Professor, PI	Ph. D.	25%	No	-	-	Nil
2.	Dr. N. Sarma	Assoc. Prof.	Ph. D.	25%	No	-	-	Nil
3.	Mr. S. Satapathy	Asst. Prof.	M. Tech	15%	No	-	-	Nil
4.	Mr. Rinku Buragohain	Research Associate	B. Tech	100%	Yes	Jan 2009	31 st Dec 2009	Rs. 12,000/-
5.	Mr. W. Lamjing Meitei	Research Associate	MCA	100%	Yes	Jan 2009	1 st Mar 2011	Rs. 12,000/-
6.	Mr. Amarjyoti Pathak	Student Assistant	M. Tech Student	-	Yes	Dec 2008	15 th June 2010	Rs. 5000/-
7.	Mr. Amitabha Nath	Student Assistant	M. Tech Student	-	Yes			Rs. 5000/-
8.	Mr. Monjit Sonar	Student Assistant	M. Tech Student	-	Yes			Rs. 5000/-
9.	Mr. Nitin Gupta	Student Assistant	B. Tech Student	-	Yes			Rs. 2500/-
10.	Md. Mustafizur Rahman	Student Assistant	B. Tech Student	-	Yes		31 st May 2010	Rs. 2500/-
11.	Mr. Bronjon Gogoi	Student Assistant	MCA Student	-	Yes	22 nd Jan, 2010	31 st May 2011	Rs. 2500/-
12.	Mr. Amit Dhar	Student Assistant	MCA Student	-	Yes		1 st Jan, 2011	Rs. 2500/-
13.	Mr. Srinu Bavera	Student Assistant	M. Tech Student	-	Yes	16 th June, 2010	30 th June 2011	Rs. 5000/-
14.	Mr. Pankaj Agarwala	Student Assistant	B. Tech Student	-	Yes			Rs. 2500/-
15.	Mr. Pankaj Goswami	Student Assistant	B. Tech Student	-	Yes		31 st May 2011	Rs. 2500/-
16.	Mr. Bhadreswar Choudhury	Student Assistant	B. Tech Student	-	Yes			Rs. 2500/-
17.	Mr. Rishi Koushik Sarmah	Student Assistant	B. Tech Student	-	Yes			Rs. 2500/-
18.	Mr. Nitin Gupta	Research Associate	B. Tech	100%	Yes	2 nd Aug 2010	1 st Dec 2010	Rs. 12,000/-

FORM G.F.R. 19

(SEE GOVERNMENT OF INDIA'S DECISION 7(B) UNDER RULE 148(3))

Assets Acquired wholly or substantially out of Government Grants

Register maintained by grantee institution

Block Account maintained by Sanctioning Authorities

**Name of the Authority: Department of Information Technology (DIT),
Ministry of Information & Communication Technology, GoI**

1.	Name of Grantee Institution	<i>Tezpur Universty, Tezpur 784 028 Assam</i>
2.	Name & Date of sanction	<i>12(3)/08-ESD dt. 16th October, 2008</i>
3.	Amount of the sanctioned grant	<i>Rs. 47.61 Lakh</i>
4.	Brief purpose of the grant	<i>Development of a Defence System against Polymorphic and Metamorphic Internet Worms for Enterprise Networks</i>
5.	Whether any condition regarding the right of ownership of Govt. in the property or other assets acquired out of the grant was incorporated in the grant-in-aid sanction	<i>Yes</i>
6.	Particulars of assets actually credited or acquired	<i>Details in Table 2</i>
7.	Value of the assets as on 30 th June 2011.	<i>Rs. 13,29,748/-</i>
8.	Purpose for which utilized at present	<i>Student training, faculty and student research.</i>
9.	Encumbered or not	<i>No.</i>
10.	Reasons if encumbered	<i>NA</i>
11.	Disposed of or not	<i>Not Disposed</i>
12.	Reasons & authority, if any for disposal	<i>NA</i>
13.	Amount realised on disposal	<i>NA</i>
14.	Remarks	

Form GFR 19 - A

Form of Utilization Certificate

Sl. No.	Letter No.	Amount Released	Certified that out of Rs. 47.61 lakh of Grants -in-aid sanctioned during the year 2008-09 (and Rs. 46,14,183/- released) in favour of <i>Tezpur University</i> under the Ministry / Department Letter No. given in the margin and interest <i>nil*</i> , a sum of Rs. 45,42,877/- has been utilized for the purpose for which it was sanctioned and that the balance of Rs. 71,306/- remaining unutilized at the end of the year has been surrendered to Government (vide DD no. 457018 dated. 30/11/2011) will be adjusted towards the grants-in-aid payable during the next year.....
1.	12(3)/08-ESD dt. 16.10.2008	Rs. 29,93,000/-	
2.	12(3)/08-ESD (pt-II) dt. 14.12.2010	Rs. 16,21,183/-	

TOTAL- Rs. 46,14,183/-

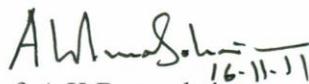
- 2 Certified that I have satisfied myself that the conditions on which the grant-in-aid was sanctioned have been duly fulfilled/are being fulfilled and that I have exercised the following checks to see that the money was actually utilised for the purpose for which it was sanctioned.

Kinds of checks exercised.

1. Procurement of equipment as per Tezpur University rules.
2. Standard procedures and the Tezpur University rules have been followed in recruiting the project personnel.


Dr. N Sarma
Co-Investigator


R R Borah
Finance Officer
Finance Officer
TEZPUR UNIVERSITY


Prof. A K Buragohain
Registrar
Registrar
Tezpur University

*As the grants-in-aid received for different sponsored research projects are maintained in a single account in the bank to minimize administrative costs, it is not possible to determine the interest accrued for the individual projects. However, the interest earned in the account is utilized for the maintenance of the equipment acquired for the projects after the completion of the projects.

PART – II

1. Project work and achievements:

a. Executive Summary:

The objective of the project was to develop a defence system against polymorphic and metamorphic internet worms. Based on the studies and the deliberations in the PRSG it was decided to build the defence system based on Vulnerability Signature of the vulnerable applications. It was decided to have the vulnerability signatures at protocol level so that these signatures can be deployed in the edge routers to filter out the exploit packets. As the process of deployment of protocol level signatures in a packet filter is a known one the job at hand boiled down to developing a Vulnerability Signature Generator(VSG) for a known vulnerability of an application.

During the course of the project the following activities have been carried out:

- 1. Studies have been carried out on the following:*
 - a. Different worms and their exploits,*
 - b. Different type of vulnerabilities and details of some specific vulnerabilities exploited by known worms,*
 - c. Different defense strategies proposed against polymorphic and metamorphic worms, namely, Signature Based and Anomaly Based.*
 - d. Use of Honeypots for detection and capturing of worms.*
 - e. Tools for packet capturing*
 - f. Mechanisms for packet filtering*
 - g. Tools for worm attack creation*
- 2. Installed and carried out studies on the following:*
 - a. Tools for packet capturing*
 - b. Tools for packet filtering*
 - c. Tools for worm attack creation*
 - d. Honeypots*
- 3. Based on the studies it was decided to adopt the vulnerability based approach for the defense against polymorphic and metamorphic worms because of its strengths in terms of its ruggedness and efficiency in preventing exploitation of vulnerabilities.*
- 4. Packet filters with known protocol level vulnerability signatures have been implemented and tested against simulated attacks by polymorphic worms to*

confirm the effectiveness of the protocol based vulnerability signature approach.

- 5. Proposed schemes for generation of vulnerability signatures available in the literature have been studied and their weaknesses identified.*
- 6. A scheme for identifying the vulnerability point in a vulnerable program with a buffer overflow vulnerability has been developed. With the scheme, given an example exploit sample the vulnerability point can always be detected.*
- 7. A new efficient scheme for generation of protocol level vulnerability signature generator has been developed with a modular structure as depicted in the DFD in Figure 1. The different modules of the vulnerability signature generator have been developed in the IDA Pro environment and these have been integrated and tested.*
- 8. A workshop on “Malware Trends and Defence” was organized during 20-21 June, 2011 in which several stalwarts in the field delivered talks and demonstrations. Around 80 participants from academia as well as industry took part in the workshop and benefitted from it.*

The Vulnerability Signature Generator(VSG):

It was necessary to develop the VSG based on analysis of binary executable code of the vulnerable application as the source of an application is often not readily available and also that the source may not give the correct picture of the runtime state due to inaccuracies introduced by the compilers. The complexity involved in binary analysis is however well known. To minimize the computational complexity and to achieve complete coverage of all the possible execution paths in the application program it was resolved to use static analysis, in the form of analysis of the CFG of the vulnerable program, wherever possible. Difficulties arise in static analysis due to pointer aliasing, use of indirect jumps, and the lack of types and other higher-level abstractions in binaries. To resolve these difficulties it was necessary to resort to dynamic analysis. Therefore, in the VSG developed, we use a combination of static and dynamic analysis of the binary executable code. We use the static code and the CFG of the vulnerable program, wherever possible, to keep the cost low and use execution traces to resolve the ambiguities where necessary.

The signature generation process in the VSG developed involves the following steps:

- 1. Disassemble the binary executable of the vulnerable program.*
- 2. Compute the CFG for the vulnerable program.*

3. *Compute the Pruned CFG to have only those branches that lead to the vulnerability point.*
4. *Remove the initial nodes in the Pruned CFG so as to start with the node that has the code for reading in the input message. This is done to eliminate the terms in the predicate that are not dependent on the input but on the connection establishment process.*
5. *Compute the decision nodes in the pruned CFG.*
6. *Merge the loops and branches that do not contain any decision node. This is done to reduce complexity in the predicate computation process.*
7. *Compute the Vulnerability Point Reachability Predicate(VPRP) in terms of the decision nodes (Preliminary VPRP) through a depth first search of the pruned CFG considering the VP as the root node and joining the decision nodes appropriately.*
8. *Take a sample input message to execute the vulnerable program.*
 - a. *Trace the execution to identify the message receive buffer.*
 - b. *Identify the protocol variable locations based on protocol specification and receive buffer access.*
 - c. *Perform taint analysis to determine the mapping of protocol variables to path variables.*
 - d. *Map the Path Variable locations to Temporary Variable Names/ Stack Frame Offset Address.*
9. *Compute the branch conditions in the decision nodes in terms of the temporary variable names/ stack frame offset addresses and replace the decision nodes in the VPRP with these conditions.*
10. *Substitute the temporary variable names/ stack frame offset addresses in the branch conditions with protocol variable names.*
11. *Take the conjunction of the VPRP and the Vulnerability exploit condition to obtain the Vulnerability Signature.*

In the above, the Step 8 involves a dynamic process as it needs tracing the execution of the vulnerable application program to resolve the identity of the path variables in the program. All the remaining steps are carried out on the static code or it's CFG.

The strength of this scheme lies in the following:

1. *The predicate computation is done through a depth first search of the CFG of the code and therefore all the possible execution paths of the program get covered.*

2. *For the path coverage tracing of execution is not involved and therefore the choice of input message is not critical. The sample input message is required only to locate the input buffer.*
 3. *As most of the process is done statically and signature generated is at the protocol level, the computation complexity is low.*
 4. *Generated signature being at protocol level, its deployment in the packet filters is simple.*
- b. Details covering targets, achievements in quantitative term and reasons for variations, if any on the following:
- i. Scope of the project :
 - *Study the existing schemes for defence against polymorphic and metamorphic internet worms.*
 - *Develop the architecture for a vulnerability signature generator that generates signature for an application with a known vulnerability.*
 - *Design and implement the vulnerability signature generator.*
 - ii. Systems/ Sub-systems with specifications or feasibility report on futuristic studies

The modular structure of the Signature Generator is depicted in the DFD in Figure 1. The different modules are-

Module 1.0: Disassembler:

The disassembler is used to obtain the assembly level code from the executable binary code of the application for further processing. The IDA Pro disassembler is used for this purpose.

Module 2.0: CFG Generator:

The CFG generator generates the CFG of the application from the assembly code. The CFG generator of IDA Pro is used for this purpose.

Module 3.0: CFG Pruning & Decision Node Computation:

The CFG pruning module removes the branches in the CFG of the vulnerable program that do not lead to the Vulnerability Point(VP). During this process the identification of the nodes in the CFG where decisions are made on the path towards the VP is also carried out. A program in C has been developed for this purpose. The pruning and decision node identification are done by tracing the CFG backwards starting at the VP. The algorithmic details are provided in enclosed Technical Report.

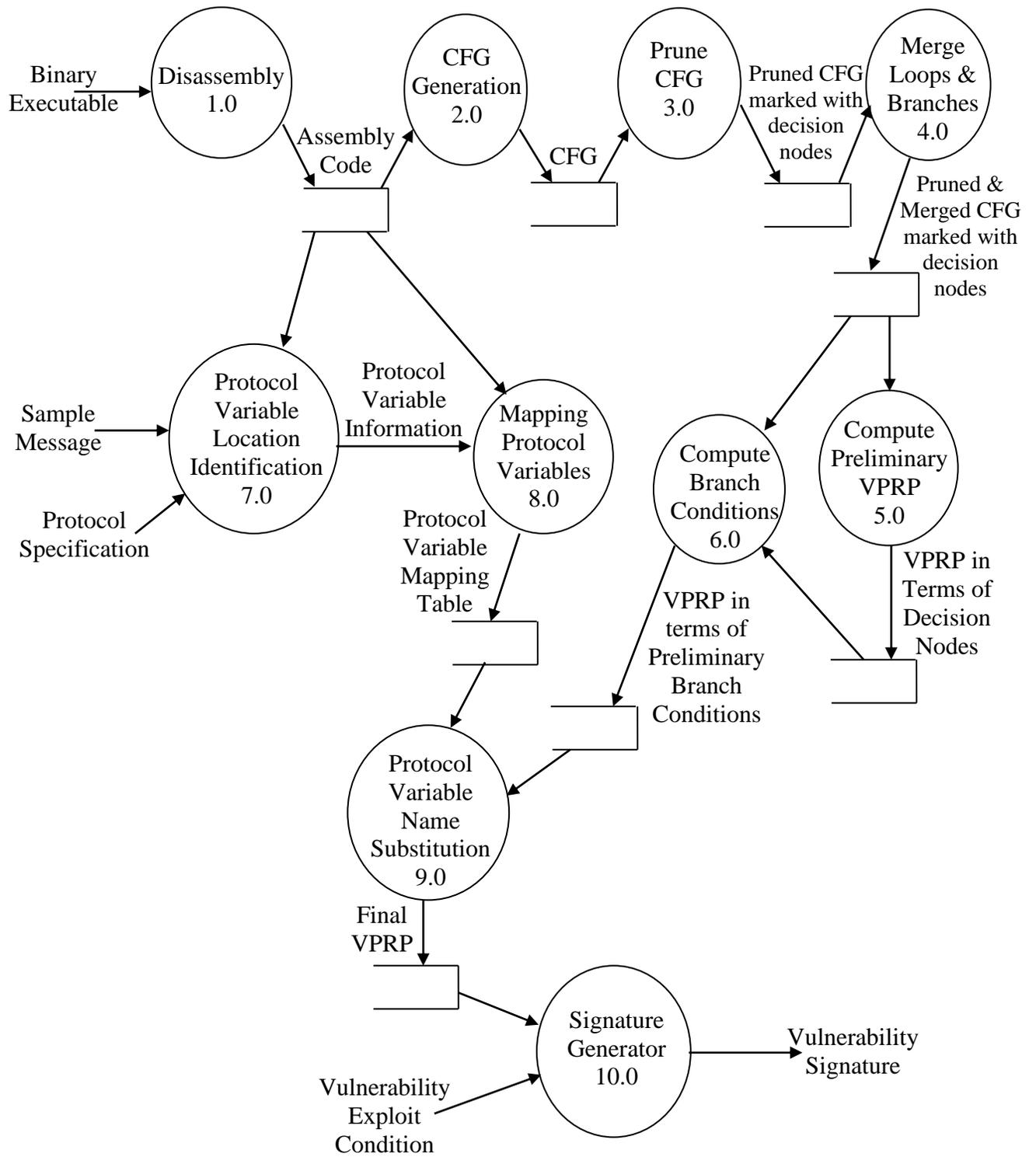


Figure 1: DFD for the Vulnerability Signature Generator

Module 4.0: Loop & Branch Merging:

This module merges the loops and branches that do not contain any decision nodes. This is done as the loops & branches not containing any decision node do not have any effect on the vulnerability point reachability predicate(VPRP). Merging of these loops & branches reduces the complexity of depth first search in the predicate computation module. A C program module has been developed to implement the algorithms below for the purpose. Please refer to Technical Report for algorithmic details.

Module 5.0: Preliminary VPRP Computation:

This module computes the Preliminary Vulnerability Point Reachability-path Predicate (VPRP) by processing the Pruned & Merged CFG of with marked decision nodes. In this predicate the branch conditions in the decision nodes remain to be computed. Instead these are represented with the decision nodes. The algorithm used for this is as follows. A C program module has been developed for this purpose. Please refer to Technical Report for algorithmic details.

Module 6.0/ Branch Condition Computation:

A C program module has been developed to extract the branch condition in each of the decision nodes in terms of temporary variable name, the stack frame offset, or the global address, as the case may be, used by each of the protocol variables in the static code and replace the decision nodes with these branch conditions in preliminary VPRP to obtain the ad hoc VPRP. The program utilizes the information already produced by the IDA Pro disassembler during the disassembly process for this purpose.

The module uses a mapping for the different branch instructions of the Intel x86 processors to the corresponding condition operator. This mapping is used to derive the operator in the branch conditions.

The details are provided in enclosed Technical Report.

Module 7.0/ Protocol Variable Location Identification:

This module first identifies the location of the receive buffer in memory at runtime where the input message is held before the individual fields are copied to the respective protocol variables in the program.

After identification of the Receive Buffer the module identifies the location of the protocol variables in memory at runtime. This is done by noting the

location in memory to which the concerned field in the receive buffer is copied.

An IDA Pro Plug in module has been developed for the purpose. Please refer to enclosed Technical Report for algorithmic details.

Module 8.0/ Protocol Variable Mapping:

From the execution trace this module produces a mapping table that shows the correspondence between the protocol variable with the temporary variable name, the stack frame offset, or the global address, as the case may be, used by each of the protocol variables in the static code. This is done with an IDA Pro Plug in.

Module 9.0/ Protocol Variable Name Substitution:

This module substitutes the temporary variable name, the stack frame offset, or the global address, as the case may be, in the ad hoc VPRP produced by Module 6.0 with the appropriate protocol variable names using the mapping table produced by Module 8.0 to produce the final VPRP. An IDA Pro plug-in has been developed for this purpose.

Module 10.0/ Signature Generation:

The job of this final module, the Signature Generator is trivial. It simply joins the VPRP produced by Module 9.0 and the Vulnerability Exploit Condition that is part of the Vulnerability Specification to produce the Vulnerability Signature. An IDA Pro plug-in does this job.

- iii. Research papers/Technical Reports brought out : None
- iv. Manpower trained:
 - 1. Fourteen M. Tech/ B. Tech/ MCA students participated in the project work and acquired valuable experience of R&D work in the field of network security.
 - 2. A two-day workshop on malware issues was organized during 20-21 June, 2011 in which several stalwarts in the field delivered talks and demonstrations. Around 80 participants took part in the workshop and benefitted from it.
- v. Anticipated know-how transfer to industry : NA

- vi. Technology/Know-how developed (Hardware, software & other details, if any); know-how document available or not : *NA*
- vii. No. of industries shown interest for know-how utilization/commercialization : *NA*
- viii. No. of users/interested for taking prototype/finished product : *NA*
- ix. No. of industries/users interested in applying the know-how developed for enhanced productivity : *NA*

2. Additional information

i) Details of patents registered, if any : *None*

ii) Technological spin offs, seeding of a major activity and how the project has helped in enhancing the technological base/capabilities in the country : *NA*

iii) Future areas for work :

a. *Further work needs to be carried out on the vulnerability signature generator to enhance its capability for the following:*

- *to handle vulnerable applications that use multiple processes or multiple threads;*
- *to be capable of handling MS Windows programs with embedded DLLs;*
- *to take care of vulnerable applications that have path variables different from the protocol variables;*
- *to generate signature for vulnerable applications compiled with compilers other than VC++, Dev C++ or gcc;*
- *to handle functions as part of a branch condition.*

b. *To make the vulnerability signature work for any application there is need to develop a program tracer for dynamic analysis that can trace programs with multiple threads and with multiple processes.*

**TABLE : 1 HEADWISE BREAK-UP OF EXPENDITURE
(Rs. In Lakhs)**

Sl. No	Head	Approved Budget Outlay (Rs.)	Amount Released (Rs.)	Expenditure Incurred up to end of the FY (2009-10) (Rs.)	Expenditure Incurred during the FY 2010-11	Expenditure Incurred during the FY 2011-12	Total Expenditure as on 30 th Sept, 2011	Balance (Rs.)	Remarks
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	
1.	Capital Equipment (including software)	13,30,000/-	13,30,000/-	13,26,248/-	3,500/-	-	13,29,748/-	252/-	
2.	Consumable Items / components	3,50,000/-	3,50,000/-	36,760/-	1,095/-	3,12,145/-	3,50,000/-	Nil	
3.	Duty on Imports	Nil	Nil	Nil	Nil	Nil	Nil	Nil	
4.	Manpower	11,28,000/-	11,28,000/-	6,17,547/-	4,23,363/-	42,500/-	10,83,410/-	44,590/-	
5.	Travel	5,00,000/-	3,83,346/-	48,801/-	31,605/-	2,90,738/-	3,71,144/-	12,202/-	
6.	Contingencies	5,00,000/-	5,00,000/-	37,458/-	51,404/-	4,11,138/-	5,00,000/-	Nil	
7.	Overheads	9,53,000/-	9,22,837/-	5,16,704/-	1,27,742/-	2,64,129/-	9,08,575/-	14,262/-	
8.	Other expenditure debitable to this project	Nil	Nil	Nil	Nil	Nil	Nil	Nil	
	Total -	47,61,000/-	46,14,183/-	25,83,518/-	6,38,709/-	13,20,650/-	45,42,877/-	71,306/-	

(N. Sarma)

15/11/11
Finance Officer
TEZPUR UNIVERSITY

Ahmedul Haque
Registrar
16.11.11
Tezpur University