Strategy and Synergy for Security

CYBSEC4AI

Task Force Report

Society for Electronic Transactions and Security (SETS)
Under Office of the Principal Scientific Adviser to the Government of India

MGR Knowledge City, CIT Campus, Taramani, Chennai – 600 113

July 2020

This page is left blank intentionally

Acknowledgments

We express our sincere gratitude to Prof. K. VijayRaghavan, President, SETS and Principal Scientific Adviser (PSA) to the Government of India for initiating this important study on "CybSec4AI" and guiding us in arriving at the recommendations.

The valuable contributions of "Task Force on CybSec4AI" with Prof. B. Ravindran, Department of CSE and Head, Robert Bosch Centre for Data Science and Artificial Intelligence, IIT Madras as Chairman of Task Force and its members Dr. P.V. Ananda Mohan, Former Technology Adviser, C-DAC, Bangalore; Dr. Sachin Lodha, TCS, Pune; Prof. V.S. Subrahmanian, Dartmouth College, USA; Prof. Chester Rebeiro, Department of CSE, IIT Madras; Prof. Sandeep Shukla, Department of CSE, IIT Kanpur; Dr. Prateek Jain, Microsoft Research & Adjunct Professor, IIT Kanpur; and Dr. N. Subramanian, Senior Director, C-DAC, Pune are gratefully acknowledged. Special thanks for their valuable insights and contributions to this report. Our thanks to Dr. Praveen Kumar Khosla, Executive Director and Dr. Sanjeev Kumar, Joint Director, C-DAC, Mohali; Dr. Prem Laxman Das, Senior Scientist and Mr Kunal Abhishek, Scientist from SETS for providing the inputs to the report.

We are thankful to academic institutions, R&D labs and industries for spending their valuable time and providing necessary inputs to our questionnaire on CyberSec4AI, which has helped us in compiling the national scenario in CybSec4AI.

Our sincere thanks to all the officials of SETS teams for providing excellent support in organizing the meetings of the Task Force and compiling this report. Specially, the services rendered by Ms. N. Eswari Devi, Project Associate, SETS in editing this document is duly acknowledged.

Several technical reports on "AI & Cyber Security" which are available in public domain have been referred in this report and contributions of the anonymous authors/groups/institutions are also gratefully acknowledged.

Dr. N. Sarat Chandra Babu
Executive Director, SETS

R. Pitchiah,
Technology Adviser, SETS
Convener, Task Force on CybSec4AI

के. विजयराघवन
भारत सरकार के प्रमुख वैज्ञानिक सलाहकार
K. VijayRaghavan
Principal Scientific Adviser to the Govt. of India

सत्यमेव जयते

विज्ञान भवन एनेक्सी
मौलाना आजाद मार्ग, नई दिल्ली - 110011
Vigyan Bhawan Annexe
Maulana Azad Road, New Delhi - 110011
Tel. : +91-11-23022112
Fax: +91-11-23022113
E-mail : vijayraghavan@gov.in
office-psa@nic.in
Website : www.psa.gov.in

Message

It is envisaged that Artificial Intelligence will enable the enhanced design of cyber security solutions and products. Equally important are the new cyber security tools which can be used to safeguard AI systems. A brainstorming session on Artificial Intelligence and Cyber Security (AI&CS) was held at SETS Chennai and as an outcome of this, a task force comprising of experts from India and abroad, was constituted for bringing out a report on the next steps and way forward at National level in this area. I am very glad to see that the task force committee, headed by Prof. B. Ravindran, IIT-Madras has brought-out a well-researched task force report.

The task force report summarizes the importance of AI & CS field to our National needs. It has brought out work carried out at national and international level, important R&D areas, mechanisms for promoting research work, establishing a distributed high quality node, creation of data sets and testbeds, skills development etc. In my view this report would be useful as an excellent reference point for taking-up activities in the area of Cyber security for AI (CybSec4AI), including the AI for Cyber Security at National level. Development of good quality manpower in this area would also be very vital for the success of this initiative. It is also pertinent to observe that this report on Cyber Security for AI is complementing the efforts of National AI plan, which is getting unfolded.

It is expected that a Project Advisory Group (PAG) will be formed to arrive at the implementation plan of the recommendations of the task force in three key areas: R & D, Training and Skills Development & Collaborations. It is also expected that as a first step, creation of High-Quality Node (HQN) in a distributed model involving multiple stakeholders, will be implemented. The detailed mission implementation plan would be driven by the PAG.

I do hope that this program will identify high-impact R&D projects in the intersection of AI and Cyber Security area that address India – specific challenges, leading to development of systems, solutions and products.

K. VijayRaghavan

(K. VijayRaghavan)

Dated : 10th November, 2020

Preface

Artificial Intelligence (AI) and Machine Learning (ML) are rapidly advancing technologies and many potential applications, ranging from machine translation to medical image analysis are being built and demonstrated. The impact of AI/ML is increasingly being felt in almost every sector of economy. In addition, AI/ML technologies can contribute to improving the cyber defence capability to protect against the ever-growing cyber threats. But the ubiquitous usage of AI/ ML empowered applications also increases many cyber security challenges.

As consumption of products and services built around AI/ML increases, specialized actions must be undertaken to safeguard not only the customers and their data, but also to protect the AI enabled systems and algorithms from abuse, trolling, and extraction.

AI is a dual-use technology. AI systems can be designed and used in ways that support both civilian and military ends. Dual-use technologies are particularly hard to regulate given competing desires to both encourage benefits and prevent harms. Research intended to make AI systems more resilient against attack also highlights their vulnerabilities to cybercrimes. While AI is enabling more automated defence systems, it is simultaneously enabling more sophisticated attacks.

There have been world-wide efforts to address the cyber security challenges in the domain of AI/ML by research institutions, academia and industry as well as by Government bodies. Within our country, there have been efforts by MeitY, NITI Aayog and others to bring out the potential of AI/ML in the context of Digital India.

A brainstorming session on AI-CS was organized by SETS on 17th November, 2018 under the chairmanship of Prof. K. VijayRaghavan, Principal Scientific Adviser to the Government of India, wherein experts from academia and R & D institutions participated. This has led to the formulation of "Task Force on CybSec4AI" with a vision to create self- reliance in developing cyber security tools using AI/ML technologies, and building secured AI/ML systems and applications.

This draft report on CybSec4AI identifies both the R&D initiatives to be undertaken in the cross-discipline of AI and cyber security and specialized skill building in this domain.

B. Ravindran

Prof. B. Ravindran
Department of CSE and Head Robert Bosch Centre on Data Science and AI
Indian Institute of Technology Madras
Chairman of the Task Force on CybSec4AI

Table of Contents

Executive Summary

This report is an outcome of the deliberations of the Task Force on “CybSec4AI”, which has been set up by Society for Electronic Transactions and Security (SETS), under the office of Principal Scientific Adviser to the Government of India. It aims at building self-reliance in the development of AI/ML based tools and systems for cyber security as well as building secured AI/ML systems and solutions.

The report has been prepared with an emphasis on three major themes namely:

(i) CyberSec4AI infrastructure development such as hardware, tools and datasets

(ii) Training and outreach programs targeting education and reskilling

(iii) Identification of R&D projects leading to systems, solutions and products

The chapter on “AI for Cyber Security” describes the need for AI based cyber security systems in the light of emerging smart adversaries with capability to target training data, test data and modelling parameters. The malware detection and classification using ML/Deep Learning (DL) algorithms and possibility of AI based Side Channel Analysis (SCA) are also mentioned.

The need to protect AI/ML based systems including critical infrastructure is mentioned with possible types of attacks on confidentiality, integrity, availability and replication. The classification of AI attacks based on attacker capabilities and handling of AI attacks are described in the chapter on “Cyber Security for AI”. The importance of adversarial attacks and a summary of adversarial attacks on Imaging, Natural Language Processing, Computing infrastructure and Cyber Physical Systems including IOT are mentioned.

A summary of similar studies carried out by international institutions such as University of California, MIT Lincoln Laboratory, and CMU is presented. Similar studies carried out at National level by NITI Aayog, MeitY, and Ministry of Commerce & Industry are also included in the survey.

The need for creating the necessary computing infrastructure and required datasets has been brought-out in the above mentioned reports. The trends and approaches towards dataset creation, survey of global and national datasets, strategy to create datasets for R&D in cyber security including collaborative approach are mentioned in the chapter on “AI-CS Infrastructure Development”.

The areas of R&D currently being undertaken in this domain based on the response received to the “Questionnaire on CybSec4AI”, which was circulated to IITs, ISI Calcutta, IIITs, NITs, and industry, are described in the chapter on “R&D Infrastructure in India”. A summary of the research papers published in IEEE, ACM and other journals and conferences during the last few years are included in the report.

Appropriate models for accelerating R & D in CybSec4AI, research areas for international collaboration and IPR sharing model for collaborative research involving R&D Labs, industry and international institutions are suggested.

A representative list of start-ups engaged in cyber security product development using ML and the support required for their growth are mentioned.

Possible areas of R & D to be undertaken by Indian institutions and a list of institutions and major areas are identified for international collaboration is provided. A list of such international institutions is provided as reference for the researchers/institutions planning to undertake research in this domain.

Building a ML or AI based system requires significant cost as well as talented engineers and experts both in ML and in cyber security. It is proposed to train the trainers from engineering institutions and universities in the models similar to National Mission on Education through ICT of MHRD and Information Security Education & Awareness programme of MeitY.

Intensive training courses on CybSec4AI are suggested to be offered to working professionals, students and researchers to jumpstart the mission. International experts and faculty members may be considered, initially, in addition to the faculty members from Indian institutions.

Around 200 PhD/M.S fellowships for a period of five years, with funding support from the mission is proposed to meet the growing needs of high end researchers in this domain. Adequate funding support to present research papers in international conferences/workshops is also recommended.

Establishment of a High Quality Nodal centre (HQN) involving R & D institutions, premier academic institutions and industry is proposed to implement the above development strategy in the cross-discipline of cyber security and artificial intelligence to meet the growing needs of manpower, infrastructure and indigenous development of systems, solutions and products.

Finally, the task force report enlists the key recommendations suggested towards national level R&D program, infrastructure building including repository of datasets, national level test bed, High Quality Node establishment, skilling & training, international collaborations, AI-CS and CS-AI standards, and launching an India portal for CyberSec4AI. The taskforce estimated that this programme needs a budget support of Rs.635 Crores, spread over a period of 5 years, towards all the components recommended in the report.

1 Context and Background

> "AI is a core, transformative way by which we are rethinking how we are doing everything"
>
> -Sundar Pichai

1.1 Introduction

The use of Artificial Intelligence (AI) and its applications is growing at a rapid rate, in this ever transforming and all-pervasive digital world we live in. AI is slowly getting embedded in the fabric of our lives today and we hardly notice it.

AI is changing the game for cyber security, analysing massive quantities of risk data, to speed-up response times and augment the capabilities of under-resourced security operations. The use of AI, Machine Learning (ML), Deep Learning (DL) technologies are enabling analysts to respond to threats with greater confidence and speed.

Realizing the importance of use of AI to bolster cyber security, the enterprises, government, defence etc., have initiated the use of AI to automate the complex processes for detecting attacks and reacting to breaches. This results in offering more protection against sophisticated hackers.

In the context of cyber security, AI is able to perceive its own environment well enough that it can independently identify threats and take the appropriate action, all without the need for human intervention. AI is particularly powerful from an incident-response perspective because it is adept at recognizing patterns and anomalies far better than any human agent ever could.

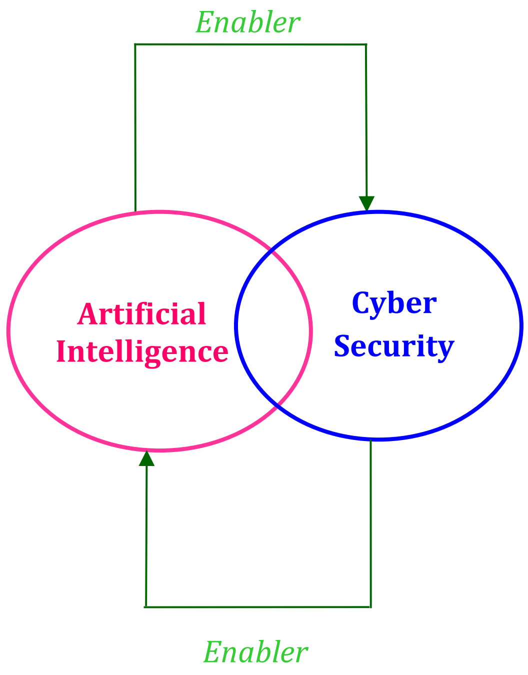
Simply put, as the amount of data continues to grow and the global threat landscape continues to advance, both in number and sophistication of attackers, organizations can no longer rely on antiquated tools and human analysts.

Automated cyber security incident response powered by AI and ML will enable strategic organizations, businesses to stay a step ahead of the threats. But the challenge is non-availability of enough data and enough expertise to take benefit of ML, DL algorithms to mechanize human intelligence.

The protection of AI systems, their data, and their communications is critical for users' safety and privacy, as well as for protecting the eco-system. AI and Security are proving to go hand in hand in enabling delivery of positive outcomes of the implemented technologies (Figure 1.1). Cyber security in the context of AI is looked at as two different implementations:

i. AI solutions are getting widely deployed today and Security must be an important consideration as those solutions evolve. **Cyber security enables better AI** by enhancing the integrity (like by using adversarial ML techniques) and producing accurate results, maintaining privacy of sensitive data of users (secure, privacy preserving ML) & preventing misuse of AI technology. The focus has to be on the protection of systems that compute, process data and algorithms used.

ii. On the other hand, **Artificial Intelligence enables cyber security applications** and plays a major role in the cyber security domain. The integration of AI into the security systems such as intrusion detection & prevention systems, next generation firewalls, Distributed Denial of Service (DDOS) mitigation systems can address the detection of zero day attacks and could take remedial actions.



AI and cyber security enable and complement each other to make systems to work better and more safely and efficiently.
AI enables new cyber security capabilities whereas cyber security enables a better AI and also prevents misuse of AI.
Intersection of AI & CS in Figure 1.1 shows how cyber(in)security will impact the development of AI and how the rise of AI will alter the security landscape.

Figure 1.1: AI and Cyber Security

This application of AI for cyber security and cyber security for AI and the possible benefits made different countries to launch national level initiatives, including India. The following sections explain some international and national efforts in this direction.

1.2 International Efforts

We briefly present the summary of various international whitepapers/reports in the following paragraphs:

One of the observations was that different dimensions of the intersection of AI/ML with cyber security (as shown in Figure1.2) namely, legal and policy issues, human factors (technical and human trust), data, hardware, software, algorithms, mathematically verified trust and operationalization intended for industry, academia, government and standardization bodies need to be considered in arriving at a framework *[IEEE Confluence 2017] [Donegan 2019]*.

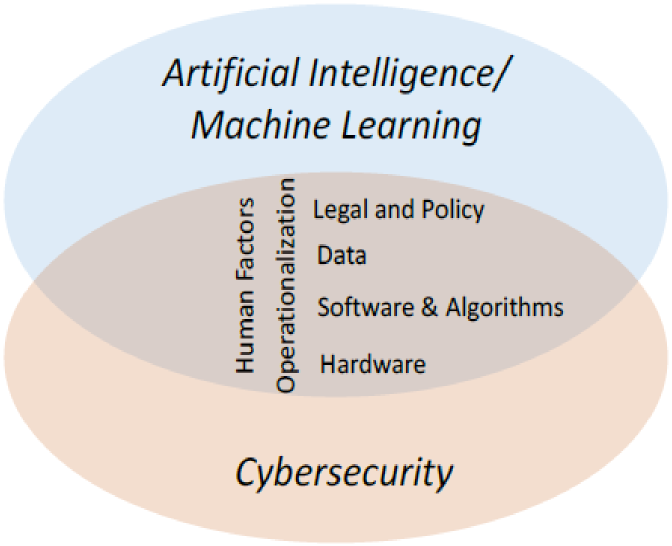


Figure 1.2: The Six dimensions of intersection of AI/ML and cyber security
[IEEE Confluence 2017]

In *[Newman 2019]*, the authors suggest including political and economic domain threats and opportunities that need to be considered for arriving at an AI security map. They also suggest global coordination and cooperation by all countries by considering gaps in different nations' current AI policy approaches.

In *[Martinez 2019]*, the authors suggest looking at areas of adversarial AI which is very critical for Department of Defense (DoD), IC and Homeland Security. But controlled introduction is needed to mitigate the very real new risks that AI creates.

In *[Loaiza 2019]*, it was suggested that automated vulnerability testing and intrusion defence, automated patching with AI using correlations of the data will be a boon to ensure cyber security.

In *[Spring 2019]*, the authors give a useful guide regarding what answers decision makers shall seek to seven questions posed when they evaluate an AI/ML solution. They also mention what answers are expected to qualify the solution or otherwise.

In *[Kubovic 2019]*, the authors opine that a safer and more balanced approach to enterprise cyber security is by deploying a multi layered solution instead of relying only on one approach. They also suggest that limitations of ML technology are important to be considered e.g., an intelligent adversary who can break the rules by often changing the entire playing field without a warning. They recommend that human expertise along with ML can act as a solution to lower the number of false positives by achieving high detection rates. The authors suggest how AI/ML can be used: (a) creating new malware (b) creating new malspam and phishing content (c) helping spammers/phishers identify recurring patterns (d) improving malware's targeting by profiling victims (e) finding the most effective attack technique (f) finding new zero-day vulnerabilities (g) delegate various tasks between infected machines in botnet and (h) letting the nodes in the botnet learn collectively.

The UK report on "Growing the Artificial Intelligence industry in the UK" *[Hall 2017]* suggested that the Alan Turing Institute should become the national institute for AI and data science, becoming truly national and expanded beyond the universities, with a key stated aim that it centres its mission on AI.

The following paras describe international efforts in AI/MI in cyber security presented in various reports.

In *[Donegan 2019]*, it is observed that AI has potential to improve threat detection and enable critical, timesaving automation in cyber security operations and that AI can help security teams respond better to ambiguity and evolve to a more probabilistic model of security operations. But it cautions that controlled introduction is needed to mitigate the very real new risks that AI creates. They point out that advanced threat detection using AI can spot threats that are sophisticated enough to have evaded at least one set of other security controls. The report however mentions that poorly controlled AI can run the risk of breaching data protection regulation like General Data Protection Regulation (GDPR).

The report by Bresniker et al., *[Bresniker 2019]* observes that while cyber security plays a vital role in protecting the sensitive data in all the existing and emerging fields, cyber security threats are threatening the safe handling of sensitive data. Six interesting dimensions of AI/ML with cyber security on the areas of legal and policy issues, human factors, data, hardware, software and algorithms are to be more concentrated to give a different approach in order to enhance the data security and its implementation. The report emphasizes that AI/ML algorithms must be trained on large and diverse training datasets to ensure the effectiveness and accuracy of the algorithms and that these models trained using AI/ML must adapt quickly to dynamic threats and must act accordingly. AI/ML can also be used to create a hardware that functions securely and predictably against all sorts of attacks. They suggest that AI can help providing fast recovery from all types of attack like DDOS attack, poisoning attack, drive by download attacks, jamming attacks, malware attacks and other types of attacks in adversarial environment.

Defence Advanced Research Projects Agency (DARPA) in September 2018 has announced a multi-year investment of more than $2 billion in new and existing programs called the "AI Next" campaign *[DARPA 2018]*. The key areas of the campaign include automating critical DoD business processes, such as (a) security clearance vetting or accrediting software systems for operational deployment, (b) improving the robustness and reliability of AI systems (c) enhancing the security and resiliency of ML and AI technologies (d) reducing power, data, and performance inefficiencies and (e) pioneering the next generation of AI algorithms and applications, such as "explainability" and common sense reasoning.

This report recommends that AI systems shall be taught to adapt with dynamic environments by recognizing and reacting autonomously to the changes in real-world settings and defending against adversarial AI. It also suggests that a new generation of defences to thwart attempts, to deceive ML algorithms need to be investigated. The systems shall autonomously assess their own competence and communicate it to human teammates.

The DARPA program called CHESS (Computers and Humans Exploring Software Security) was intended to create powerful capabilities for the DoD by attending specifically to the following areas: (a) electronic resurgence initiative for real-time analysis of sophisticated cyber-attacks, (b) detection of fraudulent imagery, (c) construction of dynamic kill-chains for all-domain warfare, (d) human language technologies, (e) multi-modality automatic target recognition, (f) biomedical advances, and (g) control of prosthetic limbs. DARPA will advance AI technologies to facilitate fast accreditation of software systems prior to operational deployment through automation. The report also stresses that the failure modes of AI technologies are poorly understood and shall work to address this shortfall with focused R & D, both analytical and empirical. DARPA report also has recommended the use of adversarial AI.

NSF programs on cyber security are for developing next generation mathematical and statistical algorithms for threat detection, Secure and Trustworthy Cyberspace (SaTC), Cyber Physical Systems (CPS), cyber security innovation for cyber infrastructure and National AI Research Institutes *[NSF 2019]*.

The goals of the SaTC program are aligned with the federal cyber security Research and Development Strategic Plan (RDSP) of National Science and Technology Council (NSTC) and National Privacy Research Strategy (NPRS) to protect and preserve the growing social and economic benefits of cyber systems while ensuring security and privacy. This focuses on six topics: (1) scientific foundations; (2) risk management; (3) human aspects; (4) transitioning successful research into practice; (5) workforce development; and (6) enhancing the research infrastructure. The NPRS, which complements the RDSP, identifies a framework for privacy research, anchored in characterizing privacy expectations, understanding privacy violations, engineering privacy-protecting systems, and recovering from privacy violations.

The CPS program aims to develop the core research needed to engineer CPS, some of which may also require dependable, high-confidence, or provable behaviours. Core research areas of the program include control, data analytics, autonomy, design, information management, Internet of Things (IoT), mixed initiatives including human-in or on-the-loop, networking, privacy, real-time systems, safety, security, and verification. The CPS program seeks to reveal cross-cutting, fundamental scientific and engineering principles that underpin the integration of cyber and physical elements across all application domains and supports the development of methods, tools, hardware and software components along with validation of the principles via prototypes and testbeds.

The objective of the Cyber security Innovation for Cyber Infrastructure (CICI) program is to develop, deploy and integrate security solutions for research data protection.

The National AI Research Institutes program works together with several states and departments on the following topics: Trustworthy AI, Foundations of ML, AI-driven innovation in agriculture and the food system, AI-augmented learning, AI for accelerating molecular synthesis and manufacturing, AI for discovery in physics.

There were several initiatives by China *[China Initiatives]* in developing AI technologies and to remain competitive with developed nations.

1.3 National Efforts

There have been efforts by Indian government departments such as NITI Aayog, MeitY, Ministry of Commerce and Industry and industry association NASSCOM to bring out the discussion papers and reports on AI. The summary of reports of various committees formed by Indian agencies is presented briefly.

Ministry of Electronics & IT (MeitY), GoI had constituted four committees in February, 2018 with an aim of understanding various regulatory and technical challenges associated with AI along with areas where the technology could be implemented and to develop policy framework for the government departments as well as ministry. These are

A. Committee on Platforms and Data for AI
B. Committee on Leveraging AI for identifying National Missions in Key Sectors
C. Committee on Mapping Technological capabilities, key Policy enablers required for sectors, Skilling and re-skilling R&D
D. Committee on Cyber Security, Safety, Legal and Ethical issues

The conclusion *[MeitY 2019]* is that concerted effort to develop cyber security techniques and tools which use AI to defend against the attacks more effectively shall be put in. They stressed that research is needed to identify the new types of vulnerabilities in AI-based applications by taking advantage of the experience of other countries. They have suggested international collaboration to defend cyber-attacks. The other recommendations for technology development are as follows: (i) quantum algorithms may be used for designing faster ML tools, (ii) security and privacy of data involved in an AI system needs to be addressed, (iii) encryption schemes which are amenable to ML techniques need to be developed and finally (iv) mechanisms for use in such AI systems, robust to quantum attacks, need to be extensively studied.

The report of NITI Aayog *[NITI Aayog 2018]* on national strategy on AI aims to guide research and development in new and emerging technologies with a focus on India leveraging the AI to ensure social and inclusive growth. They have considered five sectors: a) healthcare, b) agriculture c) education d) smart cities and infrastructure and e) smart mobility and transportation. Their report proposes a two-tiered structure to address AI research aspirations of India:

a. Centre of Research Excellence (CORE) focused on developing better understanding of existing core research and pushing technology frontiers through creation of new knowledge.
b. International Centres of Transformational AI (ICTAI) with a mandate of developing and deploying application-based research. Private sector collaboration is envisioned to be a key aspect of ICTAIs.

The Artificial Intelligence (AI) Task Force was set up by the Ministry of Commerce and Industry, government of India to prepare India for the upcoming Industrial Revolution 4.0 and the resulting economic transformation, with an emphasis on AI. This task force has identified *[MOCI]* ten specific domains of relevance to India (a) manufacturing, (b) FinTech, (c) health, (d) agriculture, (e) technology for the differently-abled, (f) national security, (g) environment, (h) public utility services, (i) retail and customer relationships and (j) education.

These areas pertain to different ministries of GoI. It may be noted that the five areas stressed in *[NITI Aayog 2018]* form a subset of these.

The report of NASSCOM based on case studies submitted as part of NASSCOM AI Game Changer Awards 2018 states that advanced analytics and Computer Vision techniques are desired by several industries. It has also listed few private software companies that are advancing in AI by creating AI based platforms. They pointed out the role of AI in cyber security solutions in terms of performance, identifying zero day attacks etc., and that AI also plays a vital role in protecting the sensitive data in all the existing and emerging domains *[NASSCOM 2018].*

1.4 Genesis of the Report - Brainstorming Session & its Outcome

There is not much focus on AI in the context of cyber security at national level. AI techniques could be used to improve and provide state of the art cyber security solutions on one hand; on the other hand the AI systems themselves with remote monitoring and control through the internet/intranet systems cannot escape the cyber-attacks and it is expected that the impact would be quite serious. The present effort by O/o PSA through SETS is to focus on AI and cyber security.

A Brainstorming session on Artificial Intelligence & Cyber Security (AI & CS) was held at SETS, Chennai on 17th November 2018 under the chairmanship of Dr. K. VijayRaghavan, PSA to GOI & President, SETS. The following were the recommendations of this meet:

Recommendations:

- Training the trainers focussing on excellence at foundational level and having in place an amplifying mechanism. To adopt combined approach roping in faculty from reputed academic institutes and offer online training materials.
- Cyber security research program: To groom PhDs through a solution driven model where the focus is not just on doing incremental research but to create new frontiers and development. Lessons learnt from ISEA and networking initiatives in the past to be taken in cognizance.
- To create a High Quality Node (HQN) as a model with all AI & CS facilities to begin with. This model could be replicated to meet the needs of the nation in due course of time. The details related to this aspect can be done by taking inputs from Prof. B. Ravindran, IIT Madras.
- To create quality data for training of AI research. Creation of datasets in our country is essential and throwing open the same through the announcement of cyber challenges across the globe is required to bring in visibility and recognition for such dataset created by our country. Visual data analytics is very important to present the useful data alone to the user based on AI to enable immediate actions by humans.
- Creation of a national level test bed for carrying out cyber security and attack experiments to study and understand attack behaviour is very much required.
- Testing of AI systems is by itself an area to be researched thoroughly and needs sufficient focus towards the same.

Plan of action and way forward:

- To setup the PMG (Project Management Group) which shall play the role of nodal agency to collate and coordinate experts from across country in CS, mathematics, statistics, AI and sector specific leaders (NSM, CPS, banking, defence, ICS/SCADA etc.).
- PSA office shall provision seed resources towards Project Management and towards consultancy contract to group of people (or) to a nodal agency. Fund support shall be extended to monitor and implement the entire programme. However, PMG will work towards generating more funds beyond seed funding from programs such as NSM, CPS; cyber security funding agencies; or from user agencies like banking, defence.
- It is recommended to encourage world class collaborations in these technologies to develop core competencies in the country. As part of this recommendation to identify and choose at least top 5 experts abroad and invite them to spend quality time with researchers.
- To identify 4-5 young leaders (from government, academia, industry and R&D labs) to drive the AI & CS program and SETS to spearhead the program and manage the PMG nodal agency to have all facilities and resources to begin with.

In line with the above recommendations SETS has initiated the Project Management Group (PMG) for cyber security for AI (CyberSec4AI) and identified the experts who can become the taskforce to arrive at a national level plan and guide the PMG for the next steps. The following is the Task Force committee constituted under the chairmanship of Prof. B. Ravindran, IIT Chennai.

The Members of the Task Force are as follows:

1	*Prof. B. Ravindran, Department of CSE and Head, Robert Bosch Centre for Data Science and Artificial Intelligence, IIT Madras*	*Chairman*
2	*Dr. Sachin Lodha, Principal Research Scientist, TCS, Pune*	*Member*
3	*Dr. Prateek Jain, Microsoft Research & Adjunct Professor, IIT Kanpur*	*Member*
4	*Dr. P.V. Ananda Mohan, Technology Adviser, C-DAC Bangalore*	*Co-opted Member*
5	*Prof. Sandeep Shukla, Department of CSE, IIT Kanpur*	*Co-opted Member*
6	*Dr. Chester Rebeiro, Department of CSE, IIT Madras*	*Member*
7	*Prof. V.S. Subrahmanian, Dartmouth College, USA*	*International Expert & Member*
8	*Dr. N. Subramanian, Senior Director, C-DAC, Pune*	*Member*
9	*Shri. R. Pitchiah, Technology Adviser, SETS*	*Convenor*

Dr. N. Sarat Chandra Babu, ED-SETS was an invitee in all the taskforce meetings.

Terms of Reference of the Task Force on CybSec4AI are as follows:

i. Preparation of Task Force Report to evolve technology development strategy in the cross discipline of cyber security and artificial intelligence.
ii. To identify and recommend R&D components, product development and solutions in the domain of CybSec4AI.
iii. To identify the innovation support required for the start-up companies.

iv. To identify the gap in the skill development at various levels including operation and management of cyber security centres and recommend methodologies to meet growing demand of expertise in this field.

v. To recommend areas of international cooperation in the AI & CS domain.

vi. To identify and recommend implementation agencies for R&D, product development and solutions.

vii. Chairman may co-opt experts including international experts as members, if so desired.

viii. The Task Force would submit the report within a time frame of six months.

SETS conducted a ''National workshop on Deep Learning for Cyber Security" during 6-7 March, 2020 to bridge the gap between CS and AI as a part of CybSec4AI mission. The report of the workshop is given in Annexure IX.

1.5 Structure of the Report

Chapters 2&3 give a brief introduction of the conceptualization of the draft report and envisage that this draft report on CybSec4AI identifies both the R&D initiatives to be undertaken in the cross-discipline of AI and cyber security and specialized skill building in this domain. It gives insights to the different dimensions of the interrelation and intersection of AI and cyber security.

Chapter 4 focuses on the ongoing R & D in premier academic institutions, R & D institutions and industries in the area of cross discipline of AI and cyber security. Lists of cyber security research centres in Indian Institutions along with the research papers published by Indian institutions are surveyed and added in this chapter. An indicative list of product development carried out by industries and startups are also included in this chapter.

Chapter 5 discusses the motivations to have datasets for R & D in cyber security and includes the survey of global and national datasets along with the strategy to create datasets for R&D purposes, also honeynet based attack data capture & analysis. A list of R & D, academic institutions, and national cyber security organizations is proposed as possible implementing agencies. As a part of AI-CS infrastructure development, need and structure of a High Quality Node (HQN) is proposed in this chapter.

Chapter 6 proposes an AICS Education and Training (AICSET) skill development model which aims in providing training to the practicing engineers, cyber security analyst and operating personnel through intensive courses.

Chapter 7 mentions about an Intellectual Property Rights (IPR) model for collaborative R & D, possible R & D areas and major areas for possible international collaborations.

Chapter 8 R&D Programme is proposed to be initiated in four schemes for a period of 5 years, with a focus on basic research leading to research papers, collaborative projects with international institutions leading to joint research publications and prototypes, product development leading to advanced prototypes and patents and mission mode projects jointly with R & D labs, industry, and academic research institutions leading to industry-competitive tools, systems and solutions.

Chapter 9 spells out the recommendations of Task Force on CybSec4AI based on the inputs from the brainstorming session held, to identify the development strategy for India in the cross discipline of AI and cyber security.

Chapter 10 concludes the report.

2 AI for Cyber Security

"Advanced AI learns to understand cyber security, recognize patterns and connect dots between threats"

-Jeb Linton, Chief Security Architect, IBM Watson

2.1 Introduction

Cyberspace has tied distinct information structures from businesses, government operations to sensitive communications/transactions and critical digital control systems together. Recent years the world witnessed one of the worst ransomware attacks ever mounted through a malware called Wannacry that infected more than 300,000 computer systems in just four days. As per the analysis of Symantec, Wannacry was followed by its ransomware variant 'Petya' affecting countries badly in the Asian Pacific region. Cyber-attacks have become the modern weapon of the 21st century. A cyber-attack on critical infrastructure could be worse than war as it can cripple a nation without firing a single-shot.

The research community has been trying to address network security problems like phishing, website intrusions, and defacements, virus and Denial of Service (DOS) attacks for more than two decades by placing a significant emphasis on developing tools and techniques such as firewalls, anti-virus software, Intrusion Detection Systems (IDS), web application security and so on. Whenever an intelligent solution is proposed to resolve a network vulnerability, the attackers bring forth a smarter way to circumvent the proposed countermeasures. Current security approaches while being effective against novice attackers using popular and known attack techniques, they become ineffective against attackers devising novel techniques and methodologies. The new technologies such as cloud, IoT have increased the complexity in computing, communication and networking infrastructures preventing cyber-crime more tedious. New approaches such as AI techniques for cyber security are expected to solve the issue. This chapter explains the influence of AI on cyber security solutions and vice-versa.

AI is considered the silver lining to drive efficient cyber security services. AI-enabled cyber security provides two major classes of benefits: improved cyber security and less cyber security risks. AI which works by imitating human intelligence excels humans in terms of speed and learning. Intelligent machines are capable of learning continuously to keep them updated to the real-world scenario.

Threats are emerging at much higher rates than before. Many times, the threats are short-lived and sometimes advanced and persistent. Evasive attacks are very difficult to detect and predict. Limited investigation time and high cost of prevention discourage enterprises and governments to invest heavily in cyber security. Manually designing algorithms for threat detection consumes a large period. The absence of labelled data also becomes an issue. The use of ML and DL algorithms which are considered as a subset of AI, in cyber security will ensure that machines are protected 24/7.

Following are some of the advantages of using AI technology in cyber security applications:

i) Countering zero-day attacks

AI based mechanisms can be deployed to take counter-measure decisions in real-time. AI based systems can analyze incidents efficiently and identify root causes. They also identify methods to deploy incident response and learn trends smartly. Hence, they have the potential to predict the next move of an attack even before they take place (zero-day attacks).

ii) AI can handle the volume of data

Legacy applications cannot respond in real-time to cyber-attacks owing to the voluminous datasets in hand, which will take time to analyze. Such difficulties arise due to the lack of automation in the cyber-threat response mechanisms. AI can analyze large volumes of data at a fast pace (in real-time) and can automate the process of detecting advanced threats.

iii) AI-based cyber security systems learn over time

AI based cyber security solutions can learn over time about the regular traffic and behavior and can detect deviations from the normal behavior. In the process AI can detect malicious attacks based on the behavior of the application and the network as a whole.

iv) Malware Analysis

Advanced techniques like code obfuscation make detection and analysis of malwares a difficult task. ML techniques like Hidden Markov model and DL are being used to model malware behaviour.

***v) Anti*-gaming Capabilities**

AI-based systems can enable anti-gaming capabilities to defend themselves from the unusual behaviour of a server equipped with gaming capabilities. Also, AI-enabled servers can apply anti-gaming capabilities to an attacking server as an offensive defence strategy.

vi) Cyber Courses of Action (COAs)

AI can benefit cyber security through the use of automated techniques to generate Cyber Courses of Action (COAs) in response to cyber threats.

vii) Cyber security for critical infrastructure

Adopting AI would allow near-instant containment, recovery, and remediation in case of a cyber-attack in the context of Critical Infrastructure (CI). AI for CI would also help in identifying slightest deviations, true root cause of the issues and predicting failures. AI for CS in CI is no more a choice but a necessity.

viii) AI in Vulnerability analysis

Even though many organizations adopt a security aware software life cycle and follow secure coding practices, software developed still has a large number of vulnerabilities that can potentially be exploited. This is evident by the growing size of the Common Vulnerabilities and Exposures (CVE) database. The few automated tools that are available are insufficient to detect all vulnerabilities and therefore today, vulnerability detection relies considerably on human efforts. AI promises to alleviate human efforts significantly. The ability to learn and identify anomalies quickly can be leveraged to detect vulnerabilities leading to safer and more secure application software.

ix) AI for Offensive Technologies and Cyber Warfare

Today it is much agreed upon that wars will be fought more often in cyberspace rather than the battle field. The increasing use of ICT in defence and critical infrastructure has led to state-sponsored cyber-attacks that target a nation's critical infrastructure, spread fake news, and manipulate social-networks. A critical aspect in cyber warfare is reconnaissance, where sensitive information is learned about an enemy. This requires stealthy web crawlers and hackers to evade malware detection, bot detection, intrusion detection systems etc. Adversarial gaming in AI can be used to remain stealthy and bypass these detection mechanisms.

Countering these threats, would require detection tools to be tested for adversarial attacks. Traditionally testing is done using a mixture of benign data and infected data. Today, testing shall include adversarial inputs so that potential areas of weakness in the AI model are identified by analyzing the functionality of the model that arise during the design and implementation of the AI model. The process of identifying potential areas of weakness with the use of adversarial inputs, referred to as adversarial testing helps in reducing the vulnerability of the system, thus ensuring the integrity and confidentiality of the system. It is a must for us to wear black hats in training and testing our system robustness in the presence of an adversarial attack.

2.2 Malware Detection and Analysis

The diversity of requirements in conventional cyber security is as follows: access control, traffic analysis, spam detection, anomaly detection, intrusion detection, malware detection etc.

Malware detection was based on heuristic features – code fragments, hashes of code fragments or the whole file, file properties, and combinations of these features. Malware detection through standard signature-based methods is becoming difficult nowadays since all current malware applications tend to have polymorphic layers to avoid detection or use side mechanisms to automatically update themselves to newer versions at short periods of time to avoid detection by any antivirus software. Malware infected file samples can be obtained from Virus Haven collection or WildList collection etc. Clean files can be collected for example from operating system files. Malware detection can be at pre-execution stage or post-execution. Post-execution phase conveys information about behaviour or events caused by process activity in a system *[Gavrilut 2009].*

A major goal of malware analysis is to capture the additional properties to be used to improve security measures and make evasion as hard as possible. ML is natural choice to such a process of knowledge extraction. ML is used to support malware analysis of Windows executables, i.e. Portable Executables (PEs) and introduced malware analysis economics, regarding trade-offs between analysis accuracy, time and cost, which should be taken into account, when designing malware analysis environment *[Ucci 2019].*

2.3 Intrusion Detection Systems (IDS)

Several ML methods can be utilized in cyber security, primarily for IDS. Research shows that no method stands out clearly as the best approach for IDS *[Loaiza 2019]*.

Although some algorithms are accepted to be better performing than others, the performance of particular ML algorithm is application and implementation dependent. One can distinguish between three types of cyber analytics that support IDS: misuse based (or signature based), anomaly-based, and hybrid. Misuse based analytics are designed to detect known attacks without having a large rate of false positive alarms. They are not capable of detecting zero-day (never before seen) attacks. Anomaly-based analytics create a model of normal behavioral patterns and attempt to detect deviations from these patterns. They have the potential to detect novel attacks and generate signatures that can be utilized to detect similar future attacks. Hybrid approach combines these two approaches. The effective and functional IDS is most likely not constructed using single method or technique, but rather a hybrid of multiple subsystems utilizing multiple data sources *[Buczak 2016].*

2.4 Generative Adversarial Networks

Malware systems can be abused intentionally to misclassify malware, evade an anomaly detection system and even attack the underlying cryptographic system. Thus, a new adversarial model has emerged. More recently, Generative Adversarial Networks (GAN) have received considerable attention. In GAN, two neural networks contest with each other in a game (in the sense of game theory, often but not always in the form of a zero-sum game). Given a training set, this technique learns to generate new data with the same statistics as the training set. For example, a GAN trained on photographs can generate new photographs that look at least superficially authentic to human observers, having many realistic characteristics. Though originally proposed as a form of generative model for unsupervised learning, GANs have also proven useful for semi-supervised learning, fully supervised learning, and Reinforcement Learning (RL). GANs can easily fool IDS. Initially, GANs were used to create false images but more recently they have expanded to cyber-attacks on other file formats. They create synthetic data which looks similar to real data. In GANs there are two blocks: a generator and a discriminator which compete with each other to win a game. Based on discriminator knowledge gained, generator modifies its attack. GANs can for example can be used for password cracking for example *passGAN*. This could create passwords not governed by rules by studying distribution of passwords in tools such as HashCat, John the Ripper etc. GANs have been applied for steganography as well (SSGAN). The generated stego images were better for information insertion than original ones without impairing quality.

2.5 Attack Detection in CPS

Cyber Physical Systems (CPS) seamlessly integrate sensing, control, networking and computational algorithms into physical components, connecting them to the Internet and to each other *[NSF 2015]*. Smart Cyber Physical Systems (SCPS) refer to advanced CPS systems, which are more interconnected through various technologies like IoT, AI, Wireless Sensor Networks (WSN), and cloud computing to provide a wide range of innovative services and applications *[Horizon 2020]*. The increased levels of interconnectivity and autonomy have given rise to an increased number of cyberattacks.

By detecting malicious behaviors and attacks early, countermeasures and mitigation actions can be taken to minimize or prevent their impacts. Conventional IDS designed for IT systems are not enough as the model of physical systems is not considered. By modelling the physics or physical properties of physical processes, physical domain Attack Detection (pdAD) techniques have proven to be more effective in securing CPS from malicious attacks. Yan et al., presents a ML based attack detection scheme and its application to gas turbine control systems *[Yan 2019]*. The key ingredient of the proposed scheme is its feature generation that enables to derive salient features from noisy physical measurements by leveraging the physical domain knowledge. Kaloudi et al., describes a framework that includes the classification of several aspects of malicious uses of AI during the cyber-attack life-cycle and provides a basis for their detection in order to predict future threats and applying in a hypothetical scenario of a critical smart grid infrastructure *[Kaloudi 2020]*.

A Deep-Q-Network Detection (DQND) scheme has been proposed to defend data integrity attacks. DQND scheme applies a main network and a target network to learn the detection strategy during the learning stage with real data *[An 2019]*.

2.6 Trustworthiness of data

AI/ML systems will not spend time on untrustworthy data where the source of data is unknown. Digital signatures help to prove the ownership of the data. Data provenance and lineage techniques assure historical records. Interestingly, blockchain technology can give distributed trust for shared data. All-access to information or modification of data shall be governed by cryptographic contracts providing accountability and auditability. The use of homomorphic encryption and secure multiparty computation can help build trust and give assurance *[Yener 2019]***.**

2.7 Artificial Intelligence and Hardware Security

Deep Learning based Side Channel Analysis (SCA) threat models: The objective of SCA is to reveal the secret key of a system (algorithm). The analysis is performed to reveal secret keys by exploiting the relations between leakages obtained during the execution of an algorithm and a statistical model. In this regard, SCA could be transformed into a classification problem. ML models are mathematical functions that find patterns in data. Particularly, DL offers a high "degree of accuracy" compared to other methods, due to the ability of extracting potential information from leakage components even in the presence of noise, and also misaligned traces. In specific, these models have the ability to sneak potential information even from protected implementations.

2.8 Consideration for adoption of AI for CS

Building an AI-based cyber security system requires significant efforts by talented engineers and experts both in AI and in cyber security. The following brings out the considerations that are essential in the adoption of AI for cyber security.

Dataset consideration

- Input data shall include or encode features that allow meaningful assessment such as prediction or classification. The collection of data requires structured observations, experiments and case studies. Analysts shall be aware that an adversary may influence the decision-making process.
- An extensive database to train learning models is required which cannot be easily acquired. It is necessary to understand the context to build a model based on chosen scenarios – information of equipment, user and application.
- Methodology to visualize the data for better training performance - how to extract profitable features of data is also very important.
- Three types of datasets will be required: (a) training dataset (b) test dataset and (c) scale-up dataset.
- Distributed computing may need to be used to train using large datasets.
- Data sources could be firewall logs or weblogs. The logs are in real-time and could be coming from various sources especially in the IoT scenario. The logs shall contain details about IP addresses, sessions, users, data sizes of ingress and egress, etc.
- Besides network datasets, specialized test beds would be needed that can be configured to synthesize large amounts of datasets which can be then used to train AI models.

Reliability of AI-CS model

- The reliability of the operational results of AI systems is of great concern. For ML/DL algorithms, decisive factors vary depending upon model types and parameters. In the case of supervised learning, the adoption of different learning rates when applying gradient descent methods have a significant influence on classification and regression results. Further, samples for training shall be various and sufficient. On the other hand, for unsupervised learning, when the system performs clustering without labels, tiny divergence in the data may affect the results. For reinforcement learning, the accuracy of the results depends on the number of neural network layers and parameters. Whenever the labelled data is not enough, one can use transfer learning by taking the trained models of related domains/tasks, which also would help in improving the speed of learning.

Characteristics of AI tool

- It is required to be clear about what protection the AI tool has and how it uses the data, and how its results are protected during training as well as operation. The AI system shall be made resilient in case of attacks being carried out.
- AI systems shall not have bias or prejudice while handling sensitive information to avoid errors.
- An assessment of how an adversary may respond to a particular AI approach shall be thought of.
- AI tools have to be aware of the security policy of the organization.
- AI tools have to be continuously maintained by gathering data, labelling it, continuously investigating and learning about the cyber security field.
- AI systems shall use more than one algorithm to increase robustness. Different objective functions and different feature vectors need to be used.

Ethical AI-CS model

- Ethical implications shall be taken into account for example, data collection, conclusions of analytics etc., shall not violate right to privacy.

Other considerations

- Effective use of the analyst's time is also quite important. A daily investigative budget of a few hundred events shall reduce the false-positive rate up to <5 %.
- An outlier detection system shall detect rare and extreme events.
- Platforms shall quantify behaviour (also known as features) from raw data learned from unsupervised learning. With high volume and high-velocity data, this will be quite challenging.
- Feedback mechanisms for human operators shall coexist with AI systems so that analysts can confirm whether some events suspected are malicious or not. Based on the feedback, the learning systems shall reconstruct or improve the models.
- Behavioural signatures shall be analyzed. An attack typically needs several steps. This information necessary to quantify the signatures is buried deep in raw data. The so-called features or variables are extracted from this information on a per-entity per-unit-time basis. This can be called as activity tracking or activity aggregation.

2.9 Typical use cases

While use of AI for cyber security is left to the imagination of cyber security experts, few typical use cases where AI is being explored is given below, Figure 2.1 illustrates some possible areas in cyber security where ai can be used to make cyber security more reliable and robust.

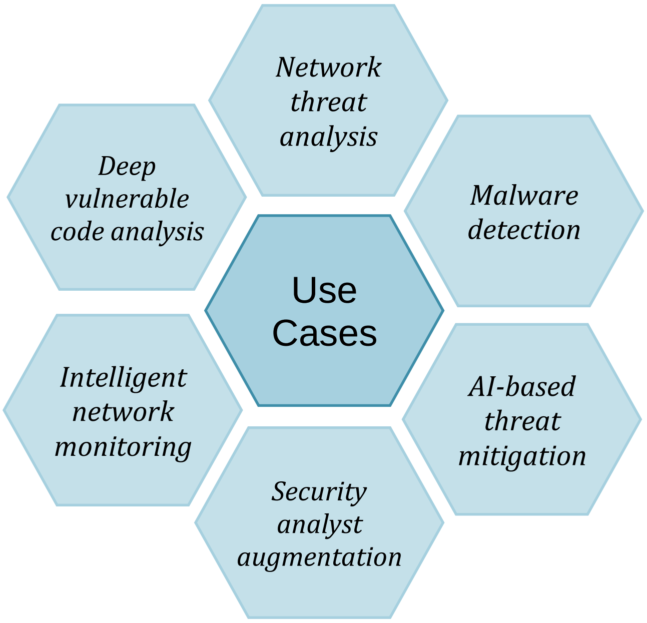

Network threat analysis
Deep vulnerable code analysis
Malware detection
Use Cases
Intelligent network monitoring
AI-based threat mitigation
Security analyst augmentation

Figure 2.1: AI driven cyber security

Network threat analysis: AI in cyber security monitors all incoming and outgoing network traffic to mine for suspicious activities and classify threat types.

Malware detection: AI in cyber security detects the presence of malware before malicious files are opened. It also identifies types of malware. This is critical because malware continues to evolve alongside other advancements, from bots and botnets to malvertising, ransomware and beyond.

AI-based threat mitigation: Cyber security technology and risks evolve in lockstep with AI. Today, companies must train machine learning algorithms to recognize attacks perpetrated by other machine learning algorithms.

Security analyst augmentation: Machine learning augments human analysts in two critical ways:

- ***AI automates repetitive tasks.*** For example, it triages low-risk alerts or tedious data enrichment tasks in order to free up analysts for higher-value or strategic decision-making.
- ***ML raises the baseline of threat intelligence.*** As a result, human analysts start with higher-order threats surfaced using machine learning to more rapidly analyze, curate, visualize and suggest potential actions.

This chapter on ***AI for cyber security*** looked at ‘building security machines that are intelligent’. It has briefly touched upon the need of such solutions in the ever-growing threat canvass, resulting benefits and various factors to be considered. The next chapter brings-out the importance of ***cyber security for AI*** in protecting the intelligent machines.

3 Cyber security for AI

"When functionality is all that matters, security is often overlooked"

- Alexandre D'Hondt

The last few years have seen a tremendous growth in the applications of AI. While many of the applications involve automating manual tasks, others intend to find solutions to problems that are otherwise difficult to solve. A considerable proportion of these applications ranging from driverless cars and intrusion detection to cyber-physical systems and medical diagnostics are critical in nature. These applications require highly reliable and accurate results.

A single false positive or false negative can be devastating in, for instance, a cancer detection application. Similarly, an incorrect detection of an object or a road sign can be catastrophic in a driverless car application. Unfortunately, with the current state-of-the-art of AI, obtaining the required accuracy and reliability is difficult. This problem is further compounded by attacks, by which an attacker can, with minor poisoning of data, trick the AI agent to provide inaccurate results. In the work of *[Sharif 2016]*, we can see that adding eye glasses to a person can cause facial recognition software to misclassify.

Adversarial attacks on automated agents are not new. For several years, adversaries have been finding new ways to fool spam filters, malware detectors, and other anomaly detection systems. However, with the rapid adoption of AI in sensitive applications, such attacks are bound to increase in number and become more sophisticated. Table 3.1 provides a summary of adversarial attacks that have been published in the recent years.

Table 3.1: Summary of adversarial attacks

Target Application	Research Papers
Images	
Image Classification	*[Goodfellow 2014], [Papernot 2016], [Carlini 2017], [Dong 2017], [Chen 2017]*
Semantic Image Segmentation	*[Xiao 2018], [Wei 2018]*
Natural Language Processing	
Machine Translation	*[Belinkov 2017]*
Text Generation	*[Liang 2017]*
Computing Infrastructure	
Cloud Services	*[Papernot 2017]*
Malware Detection	*[Katz 2017], [Krotov 2017], [Grosse 2016]*
Network Intrusion Detection	*[Huang 2018]*
IoT and Cyber Physical Systems	
Road Sign Recognition	*[Eykholt 2018]*
Spoofing Camera	*[Kurakin 2016]*
Machine Vision	*[Melis 2017]*
Face Recognition	*[Sharif 2016]*

3.1 Classification of AI Attacks based on attack motivation

Based on the motivation of the adversary, attacks on AI agents can be classified as confidentiality attacks, integrity attacks, availability attacks, and replication attacks. We discuss each of these attack classes below.

3.1.1 Integrity Attack

In safety-critical applications, integrity and availability attacks are arguably the most dangerous. An integrity attack causes a model to misbehave by poisoning the training data. For example, the AI model can be skewed by subtly retraining the model to re-categorize input data. This is especially possible in applications where models learn online. For example, Internet Search Engine Optimization (SEO), which continuously learns online, can be gradually retrained to produce incorrect results for a query. In a typical SEO attack, an attacker uses several techniques to move their pages up in search-engine rankings *[Grant Gross 2018]*. This attack is most effective in special events like a World Cup event or an upcoming election. During natural disasters, attackers use SEO poisoning methods to trick people sending aid to them instead of to the needy victims.

3.1.2 Availability Attack

In an availability attack, the availability of an AI model to output the correct result is compromised. Availability attacks work by subtly changing the input such that, to a human, the input seems unchanged, but to the AI model the input looks different, as a result, it produces the wrong result. An infamous attack is in driverless cars. For example, in *[Sitawarin 2018]*, the authors explore how slight modifications to traffic signals, such as by applying stickers, paint, and projecting images on the traffic signal using lasers, can cause the AI agent to make the wrong decisions. These wrong decisions can be catastrophic in such a case. It can potentially lead to crashes and cause traffic jams.

We now take an example of an Availability attack against a common ML classifier *[Andy Patel 2019]*. Such classifiers are used in wide applications such as sentiment analysis, malware analysis, and object recognition. During the learning, classifiers identify boundaries that can distinguish data. For example, execution time characteristics of several programs are used during the learning phase to determine a boundary that separates benign programs from malware.

During testing, a program is classified as benign or malware depending on which side of the boundary its execution time characteristics fall. By studying the boundary conditions carefully, an attacker can tweak the malware to have execution-time characteristics similar to that of a benign program. The malware would then get classified as a benign program – evading detection.

3.1.3 Replication Attack

In a black-box scenario, a replication attack would include attacks that reverse- engineer the model and its parameters. One common motivation of this is to steal intellectual property. In a typical replication attack, repeated queries are made to the target model using crafted inputs and the corresponding outputs are recorded.

The inputs may be all independently chosen or adaptively chosen based on the outputs observed. The outputs are used as a guide to create the new model. As the number of queries to the target model increases, the new model would more closely mimic the behaviour of the target.

One well studied application of a replication attack is in Physically Unclonable Functions (PUF). It is well studied that attackers, who can monitor sufficiently large number of challenge and response pairs are able to create a model that can mimic the PUF behaviour. This has far reaching consequences. Secret cryptographic keys can be stolen, devices can be cloned and the root-of-trust can be compromised.

3.1.4 Confidentiality Attack

The motivation of the attacker in a confidentiality attack is to obtain information about the data used during training. For example, by a series of queries to the AI model, an attacker can infer the medical condition of a targeted victim. One of the mechanisms to achieve this is the so-called Membership Inference Attack: where a data record and black box access to the model is used to identify if a record was in the model's training dataset. Membership inference is done by training custom inference models to recognize differences in the target model's predictions on the inputs that it trained on versus the inputs that it did not train on. Thus, the membership inference problem is converted to a classification problem.

3.2 Classification of AI attacks based on target

There can be threats in the *training phase*. These are called poisoning threats. Classification and regression depend on labels and poisoning which injects malicious samples may affect these processes. The solution is to ensure purity of the training data by separating adversarial samples from normal ones and removing them. Improvement of learning algorithms also can be carried out. For example, contextual data information can be used for this purpose.

There can be threats in the *testing phase* also. One of them is evasion attack during inference time. In this, adversarial samples– inputs carefully perturbed to look like real inputs- are fed which totally fool the classifier. These were well known in spam filtering, PDF malware detection etc.

Other kind of threat in testing phase is impersonate attacks. In this attack, data samples are imitated by inserting fake samples making it difficult to categorize the original samples with different labels. Solution is to retrain our systems using adversarial samples and test whether abnormalities are detected.

The third type of threats can occur in data privacy and security. Inversion attack is one such threat. By collecting ML operation results, one can perform an inverse operation. This can be of two types: white box where knowledge of the model is assumed to be known and black box where no knowledge of the model is available. The solutions could be based on homomorphic encryption *[Guan 2018]***.** This has found wide application in cloud environments. In this approach, we perform calculations on cipher text blocks. Public key crypto systems like RSA, El-Gamal can be used. Another technique is differential privacy which can be achieved by data encryption. It is very hard to obtain user information from what is available.

For example, dataset can be publicly shared describing patterns of the group but information about individuals is withheld.

Threats can be in other areas. In the case of mobile devices, malware in android asks users to give permissions to access user's data and system resources. Mobile embedded machine learning chips are available which can implement AI algorithms and do analytics locally. Side channel attacks could be carried out on wearable and mobile AI devices.

3.3 Classification of AI attacks based on attacker capabilities

Adversarial attacks are classified as white-box or black-box. In a white-box attack, it is assumed that the adversary has some knowledge and/or can tamper with the (1) labeled training data, (2) features extracted from the training data, and (3) classification algorithms. Having access to all these three (the strongest white-box model) would lead to the strongest adversary. In the black box model, the attacker knows none of these. Instead, it learns the model by performing a series of iterative queries on the target model. Consequently, the success of the attack changes based on the amount the adversary knows.

We take the example of the impact of adversarial evasion attacks on ML based Android malware classifiers *[Malware Classifiers]*, where fewer modifications of the features are sufficient to cause a false detection by the ML algorithm.

3.4 Handling AI Attack

Traditional security techniques, such as cryptography, have undergone over half a century of rigorous research and analysis. As a result, crypto-algorithms are backed with robust mathematical proofs of security. Tools ranging from Shannon's Entropy to Computational Security have evolved as a means to achieve this. Global competitions with scores of researchers extensively evaluate every crypto-algorithm before standardization. Periodically, these standards are reviewed and enhanced to meet with the growing adversarial power.

Similar arguments do not hold for any of the AI algorithms. While AI algorithms have been demonstrated to work in a range of applications, providing guarantees on their results is still an open challenge. Thus, most AI algorithms are highly vulnerable to attack.

The current countermeasures proposed are more temporary fixes rather than permanent solutions. It would require rethinking of traditional security for AI algorithms. For each algorithm and in each application, the threat space needs to be deeply studied and evaluated in the context of AI attacks.

With the eventual adoption of AI in mainstream safety critical applications, the need for securing the algorithms has never been higher. This would require substantial work on developing the theory, providing the guarantees in normal conditions and when under attack. Potential vulnerabilities in algorithms need to be identified and evaluated. Research should be carried out on designing new attacks that correspond to the identified threat and studying their impact on the system. New countermeasures should be proposed to improve the security of these algorithms.

3.5 Social Media Attacks

Social media has been a thorn in the side of enterprise security for some time now. Social media platforms are being used as a Trojan horse by hackers to enter enterprise. Up to 1 in 5 businesses have been infected with malware originated from social media and 1 in 8 have experienced security breaches as a result of social media directed cyber-attack *[McGuire 2019]*.

Fire et al., discuss Online Social Network (OSN) threats such as malware, phishing attacks, spammers, cross-site scripting, internet fraud as classical threats, and clickjacking, de-anonymization attacks, face recognition, fake profiles, identity clone attacks, inference attacks, information leakage, location leakage, socware as modern threats *[Fire 2014]*. The attackers can also combine classic and modern attacks. Threats against children such as online predators, risky behavior, and cyberbullying are also mentioned. A variety of solutions from social network operators, commercial solutions and academic solutions are also listed against various threats.

In OSNs, a Sybil is a fake account with which a user attempts to create multiple identities to make as many friends as possible with legitimate accounts. A Sybil account can lead to many malicious activities in an OSN. Al-Qurishi et al., surveyed research relating to Sybil attack defence schemes and techniques in OSN *[Al-Qurishi 2017]*. In general, the schemes are divided into four main categories: Graph-based schemes, ML based schemes, manual verification and prevention approaches. ML approaches are based on specific features, which make them vulnerable to Sybil attacks that are executed using different strategies. Further research is needed to stop fake users from causing negative effects, altering actions, and deriving quantifiable outcomes while using online platforms.

3.6 Ongoing Research on Secure AI

A number of defence mechanisms against evasion, poisoning, and privacy attacks have been proposed in the field of adversarial ML:

- The definition of secure learning algorithms *[Bruckner 2012]*
- The use of multiple classifier systems *[Biggio 2010]*
- The study of privacy-preserving learning *[Biggio 2014]*
- Ladder algorithm for Kaggle-style competitions
- Game theoretic models for adversarial machine learning and data mining *[Kantarcioglu 2011]*
- Sanitizing training data from adversarial poisoning attacks

One of the main flaws with current efforts on adversarial AI is that they seek to reduce the overall accuracy (usually measured through metrics such as Aurea under a Receiver Operating Curve or an F1-Score) of a ML classifier. However, in the real world, an attacker may not be interested in compromising the overall accuracy of a classifier. The attacker may only care about getting the attack through. This needs to be studied more carefully as a large number of attacks correctly flagged by a classifier may in fact provide camouflage for the one attack that the attacker wishes to succeed.

3.7 Available Software Resources

Some initial works have been initiated in this direction. Some open source tools *[Wikipedia]* are available, mainly for testing purposes and research.

- AdversariaLib- includes implementation of evasion attacks
- AdLib- Python library with a scikit-style interface which includes implementations of a number of published evasion attacks and defences
- AlfaSVMLib- Adversarial Label Flip Attacks against Support Vector Machines, poisoning attacks against support vector machines, and attacks against clustering algorithms
- deep-pwning- Metasploit for DL which currently has attacks on deep neural networks using TensorFlow. This framework currently updates to maintain compatibility with the latest versions of Python.
- Cleverhans- A TensorFlow library to test existing DL models versus known attacks
- foolbox- Python library to create adversarial examples, implements multiple attacks
- SecML- Python library for secure and explainable machine learning- includes implementation of a wide range of ML and attack algorithms, support for dense and sparse data, multiprocessing, visualization tools.

Events on AI-CS

- NIPS Workshop on Machine Learning in Adversarial Environments for Computer Security
- Special Issue on "Machine Learning in Adversarial Environments" in the journal of Machine Learning
- Dagstuhl Perspectives Workshop on "Machine Learning Methods for Computer Security"
- Workshop on Artificial Intelligence and Security, (AISec) Series

This chapter has brought out the need for incorporating cyber security for ai models to ensure reliability and robustness. The following chapter focuses on the current scenario of R&D being carried out in the cross discipline of AI-CS by various industries, institutions and start-ups in our nation.

> *Through 2022, 30% of all AI cyberattacks will leverage training-data poisoning, AI model theft or adversarial samples to attack AI-powered systems*
>
> *-Top 10 Strategic Technology Trends for 2020, Gartner research*

4 R&D Infrastructure in India

The task force felt the need to understand the national-level available infrastructure, solutions/products and training/education in the cross domain of AI and cyber security at Indian institutes, R&D labs and industry. A questionnaire was prepared and was sent to a cross section of academia, R&D labs and industry, spread over the country. This has helped in understanding the available infrastructure in the form of focussed research centres, on-going research, development and education in premier academic institutions, R&D institutions and industries. Some inputs are also obtained by googling through the net for the research papers published by Indian researchers in IEEE, ACM and Citeseer. An indicative list of solution/product development carried out by industry and start-ups is also provided in this chapter.

4.1 Ongoing Research & Development

A list of on-going research in premier academic institutes such as Indian Institute of Technology (IITs), Indian Statistical Institute (ISI), National Institute of Technology (NITs), Indian Institute of Information Technology (IIITs) and R & D labs is given below. This is not an exhaustive list as it is a summary of the responses received for the questionnaire on CybSec4AI.

i) Anomaly detection where techniques are developed to gather information from social media, blogs to detect online presence of users and gather intelligence about their malicious behavior. AI/ML based techniques are developed for automatically detecting vulnerabilities in software applications and to identify any exploits available for malicious attacks

ii) Areas of adversarial ML and malware analysis. The work is being carried out on deep fake and a classifier is being developed based on adversarial robustness to identify deep fake videos. This also uses pattern recognition and predictive analytics to process data and identify patterns.

iii) Attacks on IOT networks**,** mitigating the attacks by exploring physical layer in addition to computational complexity-based security algorithms. It is enhancing the throughput in dense traffic IoT networks by developing AI based channel access mechanisms and proposing adaptive communication strategies for optimal performance.

iv) Optimization methods and models for non-parallel support vector machines and on Digital forensic knowledge integration and intelligence.

v) Secure distributed computing, with emphasis on secure Multi-Party Computation (MPC) and verifiable computation, both at the theoretical as well as applied level.

vi) Areas of cyber threats detection, network security, Natural Language Processing (NLP), security management, identity and access management.

vii) Bitcoin prediction using AI/ML models. Feature extraction from block chain data and the prediction is done using DL models.

National R&D Labs:

Centre for Development of Advanced Computing (C-DAC) is working in the areas of cyber security and cyber forensics. C-DAC has developed various solutions for endpoint security (desktop & mobile), perimeter and network security, cloud & virtualization security, SCADA security, steganography, identification and authentication, biometrics and face-recognition, web application security, honey-net & attack analysis, cryptanalysis and cyber forensics. Many of these solutions are deployed and being used by user agencies. It has been using AI/ ML techniques in some of the tools developed like malware resist, dynamic firewall. It has also offered a number of PG level training programs in the related areas such as AI and big data, cyber security through its Advanced Computing Training School (ACTS). C-DAC is spear-heading the Government's Information Security Education and Awareness (ISEA) programme towards sharing, disseminating and creating awareness related to cyber security across the country.

It has expertise in dataset creation and cyber threat analysis**.** It has been closely working with key agencies like CERT-In to share the data and carryout joint analysis of the same to detect newer/variants of attacks. C-DAC Mohali is executing a project "Scalable Attack data capturing and analysis framework for Cyber Threat Intelligence Generation". Cyber Threat Intelligence will contain the evidence-based knowledge, including context, indicators that will be in actionable formats and will be directly utilized by any security solutions. The attack data captured and collected through this system can be further normalized and labelled and make it available for public accessibility for applying computational algorithms such as AI/DL/ML algorithms to determine what is malicious and what is benign.

Some of the Indian institutions (IITs, NITs and IIITs) and R&D labs working in the areas of artificial intelligence and cyber security are given in Annexure VII.

4.2 Cyber Security research Centres in India

Society for Electronic Transactions and Security (SETS)

SETS is an R&D lab under the office of the Principal Scientific Adviser to the Government of India. The main charter of SETS is to pursue basic and applied research in the areas of cyber security, network security, security of embedded systems and cryptology. SETS has come out with usable products/solutions in the areas mentioned which includes Integrated Threat Management Appliance (ITMA), e-Abhedya, KDE for QKD IP Core, puzzle based DDOS Mitigation software, True Random Number Generator, crypto-accelerator, crypto libraries (hardware/software), Deep Packet Analysis and Tuning Engine and working on many more products.

Pratap Subhramanyam Centre for Digital Intelligence, Security, Hardware, and Architecture (PS-CDISHA, IIT Madras)

The PS-CDISHA was established in 2020 in the department of Computer Science and Engineering at IIT Madras. One of its primary research areas is in computer security with focus on research, education, and training.

The centre, which has developed India's first RISC V microprocessor, SHAKTI, is now designing a secure variant with a generous grant from the Ministry of Electronics and Information Technology (MeitY), New Delhi. The secure RISC V microprocessor would support logic locking, memory encryption, secure boot, side-channel resistant crypto-accelerators, memory protection, fine-grained compartments, and support for safe programming languages like OcamL. The centre is currently developing a secure AI-coprocessor based on the SHAKTI core.

The centre boasts of a scalable and reconfigurable malware testbed, Jugaad, with 300 Intel Atom boards. ML and DL models are used to detect attack threats at three different layers – network, Operating System, and hardware, making it the most comprehensive malware analysis testbed. With close ties with TAU Israel and K7 Anti-virus, India, the centre's malware research focuses on early detection of malware and its propagation characteristics. With considerable strengths in binary exploitation, the centre has dissected a number of malwares including Stage-fright and Mirai.

An important focus in the centre is the design of tools for security assessment of applications. The centre has expertise in formal verification tools like SMT solvers, symbolic execution, and model checkers to assess software and hardware codes. The centre has developed a formally proven public-key crypto library for the Defence Research and Development Organization (DRDO). Further, with University of Florida and IIT Kharagpur, the centre has won an unrestricted award from CISCO for fault attack vulnerability research and has developed compiler plugins to detect and patch fault attack vulnerabilities in software and hardware.

Robert Bosch Centre for Data Science and AI (RBC DSAI), IIT Madras

RBCDSAI at IITM is the one of the largest inter-disciplinary Data Science and AI centre in India focussing on applications of AI in various engineering disciplines. RBCDSAI has made significant contributions in fundamental AI areas such as DL, natural language processing, reinforcement learning, causal reasoning, etc. In particular, the centre is engaged in research in mitigating adversarial attacks on deep networks for classification, ranking and reinforcement learning. They also work on learning with poisoned data and unreliable data sources.

The Interdisciplinary Centre for Cyber Security and Cyber Defense of Critical Infrastructures (C3I Centre), IIT Kanpur

The C3I Centre was established in 2016 in the department of computer science and engineering of IIT Kanpur for research, education, training and awareness campaigns on cyber security of critical infrastructures. This centre is funded by the Science and Engineering Research Board (SERB), Department of Science and Technology, Government of India. C3I centre focuses on some of these critical infrastructures, power grid, industrial automation, water treatment, agriculture, and IT.

At C3I, India's first industrial scale testbed for power grid is being built. They have a test bed for distribution automation with SCADA control in operation for 2 years, and more than 10 responsible disclosures of vulnerabilities in PLC, SCADA, and other components have been made to the industry partners.

Machine learning based detection of threats with high accuracy (above 95%) and with low false positive rate (less than 5%) has been implemented and is being integrated into the threat intelligence monitoring system. This will help any SOC (Security Operation Centre) to obtain full visibility, situational awareness and actionable threat intelligence. This system can be customized for other facilities outside the centre as well. Malware detection and classification systems were built using machine learning. C3I Centre has built many server honeypots (for SSH, FTP, HTTP, SQLDB etc.,) as well as for IoT protocols such as CWMP and MQTT etc.

National Centre of Excellence in Technology for Internal Security (NCETIS), IIT Bombay funded by MeitY

The objective of the centre is to develop state-of-art-technologies which are vital to our security agencies (police and para military forces) and for providing rescue and relief operations. This centre has been set up for addressing internal security needs of the nation on a continuous basis and to deliver technologies/prototypes required for internal security and to promote domestic industry in internal security area within the country. NCETIS will act as a national nodal facility, to cater to the requirements of homeland/internal security technology requirements and technologies for disaster management.

The centre has been mandated to develop technology prototypes, in PPP mode through industry collaborations on Broadband Wireless Communications System for Public Protection and Disaster Recovery (B-PPDR), ground penetration radar, video surveillance and analytics, robotic all-terrain vehicle, GIS for spatial data security, handheld explosive detectors and thermal imaging.

Information Security Research and Development Centre, IIT Bombay

The Centre was initiated as part of the Information Security Education and Awareness (ISEA) program funded by MeitY. It focuses on research in information security: operating systems, language-based security, cloud security, IoT security, android security, privacy, blockchain etc.

Centre for Excellence in Cyber Systems and Information Security, IIT Delhi

The main mission of this centre is to conduct basic research in cyber systems, education and training. The focus is on areas such as critical infrastructure cyber security, visual analytics for cyber security, wireless security, mobile security, network security, bio-inspired technologies for enhancing cyber security, cyber security in Online Social Media (OSM), cyber forensics, detection of cybercrime and data fraud and low-cost system.

Security Research Group, IISc., Bangalore

This group works on diverse topics in security ranging from symmetric and asymmetric-key cryptosystems, secure multi-party computing protocols, post-quantum cryptography, information-theoretic security, operating system and application security, wired and wireless network security, cloud platform security, security of mobile devices and apps, security of cyber-physical systems and the Internet-of-Things (IoT).

Cyber Security Education and Research Centre, IIIT Delhi

Besides education in cyber security, the focus of this centre is on security and trust in distributed systems, privacy and anonymity in networks and complex systems, software and multimedia security, and cyber polices.

Centre for Security, Theory, and Algorithmic Research, IIIT Hyderabad

The centre supports various programs in the area of theoretical Computer Science and Information Security. The centre runs its own M.Tech. in cyber security and offers various programs such as M.S and PhD in Information Security. The main focus of the centre is on theoretical cryptography, network security, and malware analysis. Some of the works include secure VoIP, firewalls, network traffic analyzer, etc.

Centre of Excellence in Cyber Security, an initiative by Government of Karnataka
The centre in cyber security was formed by the Government of Karnataka, as part of the technology innovation strategy, to promote the cyber-safe and conducive environment for industry collaboration, address the skill gaps, build awareness and facilitate innovation in this emerging technology field of cyber security.

4.3 Papers published by Indian Institutions

A survey of the papers published by Indian researchers during 2017-2020 in IEEE, ACM and other journals and conferences based on the web search has revealed that the Indian authors have worked in the following areas:

ML/DL algorithms used as a defensive mechanism:

- Data integrity attacks on smart grids
- Malicious attacks on IoT devices and networks
- Classification of SNMP dataset
- Secure shell traffic analysis
- Network traffic prediction
- Design of network defense system against port scanning attacks
- Generative adversarial networks for dynamic network embedding

ML/DL algorithms used in detection of Adversary:

- Detection and classification of Android Malware applications
- Image forgery detection
- Detection of DOS attacks
- Malware and ransomware detection
- Network intrusion detection and database intrusion detection
- Spam detection
- Malware text classification

Cyber security for AI:

- Security of DL algorithms from side channel based reverse engineering
- Robustness against adversarial attacks on DL systems

List of research papers published from India are provided in the Annexure VI.

4.4 Industrial Research

4.4.1 Research on cyber security and privacy

An Indian company, Tata Consultancy Services, which has an R&D centre on cyber security and privacy is working on following research themes as part of their company's business 4.0 Vision.

01. Secure, Safe and Private AI/ML
Suitably tweaked training data can lead to untrustworthy AI systems. Alternatively, systematic querying to existing models can help invert them, learn training data, and thus breach privacy.

With small noise addition to the test data, the model can be tricked into targeted misclassification, completely undermining the confidence in such systems. This has serious safety repercussions, with everything being connected e.g., health, automotive domains.

02. Integrity and Confidentiality of Data in Cloud

There is an explosion in the cloud adoption for the agility, flexibility and economies of scale it offers. However, security and privacy of data in the cloud continue to remain major concerns, more so, with stricter data protection regulations coming up across the globe. In terms of the famous Confidentiality-Integrity-Availability (CIA) triad of security, cloud certainly scores high on the 'Availability' aspect, but 'Confidentiality' and 'Integrity' of data in cloud need serious attention and effective technology solutions.

03. Securing Internet of Things

While IoT has a lot of disruptive potential, catastrophe awaits in its new vulnerabilities and dangers. Here even proven security and privacy methods and tech require a fresh look. For example, standard crypto does not work owing to low power, low CPU, low memory in many IoT devices. Further, software/firmware updates and their patch management in the IoT context is a management nightmare and IoT devices are probably not best suited for the user-id + password-based authentication, a standard practice otherwise.

04. Privacy by Design

Privacy is typically an afterthought, and added later as a bolt on. This turns out to be a more expensive and less effective solution compared to embedding privacy in the whole software/system design and consideration right at the start. With the emerging regulations like GDPR, the cost of privacy failure has risen significantly, so organizations cannot afford to ignore it any more. Interestingly, GDPR itself preaches and endorses privacy by design approach!

05. Strengthening the 'Carbon' Layer

Humans are always considered as the weakest link when it comes to security. With everything going digital, both in personal and professional spaces– the cognitive load on humans is growing alarmingly, so they are expected to behave even more poorly when it comes to security and privacy related decision making, especially in the professional setting. Additionally, malicious insiders continue to be a big pain for organizations, given the ease of attacks and what is at stake.

Other research projects being carried out in 2019-2020 by industries include secure and private learning, applications of fully homomorphic encryption, infonomy, lightweight cryptography, fair usage privacy controls, privacy assurance, speech masking for privacy, provenance-based access control, privacy preserving secure biometrics, human aspects in cyber security and user and entity behaviour analytics.

4.4.2 Product Development and Systems Development

1. The presentation by Data Security Council of India (DSCI) on the "Indian Cyber Security Product Landscape-Key Findings" shows that the products were focused on technologies such as AI/ML, automation, analytics, encryption, blockchain, quantum cryptography, and threat intelligence. It was reported that 30% of the products were ML enabled.
2. Based on the questionnaire on CybSec4AI circulated to academic research institutions and R & D labs in our country, the following is the list of on-going products/systems development:
   - A tool has been developed to analyze malicious behavior in social media where, ML techniques are used extensively to extract entity (or keywords), sentiment analysis over extracted text and infer preferences of the user.
   - Anomaly based detection techniques are implemented for network scan, DoS, DDoS and flooding attacks.
   - Model based anomaly detection (protocol and pattern-based models), behavior based analysis and context aware detection is used for detecting incidents.
   - Cyber Threat Monitoring with actionable Cyber Threat Intelligence (CTI) is carried out through large scale threat monitoring by deploying threat capturing sensors at different locations.
   - Malware detection using AI/ML, phishing URL detection, adversarial attack detection on AI/ML models and their defence mechanisms, etc.
   - ML is actively used to classify user-browsed URLs into respective categories. This allows the administrator to restrict access to certain categories of websites.
   - Features extracted from scanned objects, e.g., Windows Portable Executable objects or android packages, etc., can be used to
     - cluster similar-looking objects for family-based generic signature detection
     - classify into malware and benign
   - Understanding the risks of using machine learning models in cyber security and their weakness to model poisoning techniques, etc., and coming up with defensive measures to mitigate such scenarios through adversarial nets.

4.4.3 Industry practice/perspectives on CybSec4AI

AI for cyber

Industry perspective in using AI for cyber security is to use ML to analyze and find vulnerabilities in the given system, to suggest and secure the given system and to induce attacks on the given system. The expected challenges are unknowns presented in ML data, source of data and amount of data.

Some properties of data considered are as follows: clean/corrupted; closeness between training data and production data (is the data representative? e.g., Aadhaar); features that better represent the signal in data; unknowns present in ML models. The interpretation of the model arrived at: to understand whether the model outcome is explainable; how it learns; what it learns and how it behaves for unseen data (can the model generalize?).

It is essential to know the model functionality/behavior: whether it is genuine/misleading; are there blind spots (worrisome for 0/1 setting); is model transferability possible; does it fall under distributed models.

Cyber for AI

ML models are often not trained keeping in mind privacy and security of the data and models. The main goals of an adversary are as follows:

i. To retrieve sensitive private training data (privacy breach, membership inference)
ii. To poison the data that is being used in learning (model integrity compromise)
iii. To fool the models to make incorrect predictions (model integrity compromise)
iv. To duplicate/extract the proprietary models (model inversion)

Open questions:

i. How to detect when an AI model is fooled?
ii. How to find robust defences invariant to changing environments?
iii. What processes need to be followed to mitigate the misuse of AI?
iv. What architecture and policies to be implemented to ensure security of AI and privacy of data?
v. How to ensure trustworthiness and transparency of the models?

Annexure VIII provides further details.

4.5 AI-CS Start-up Ecosystem in India

Annexure V provides details of few start-ups using AI to drive cyber security innovatively.

4.5.1 Challenges faced by Start-ups

Some of the challenges faced by these start-ups are listed below:

- Customers are hesitant to change to a solution offered by start-ups, thus increasing customer acquisition cost.
- Many of the start-ups are bootstrapped and the sustenance of the start-ups in the markets is also questionable. Identifying investors and raising funds although the product has received a positive response is a difficult task. Lack of good mentors and a supportive environment also impacts sustenance.
- Skilled cyber security professionals are scarce and acquiring them is expensive. Hiring qualified cyber security professionals with working knowledge on AI is even scarcer. Most of the start-ups focus on recruiting experienced professionals, as the time and cost of training a fresher are high.

4.5.2 Skilling up for the new era

The Union Budget 2020 has allocated a fund of Rs.3000 crores for skill development. The IT/ITeS industry is at the forefront of the ongoing talent skilling and re-skilling initiatives. The government, industry, and academia have concerted courses, educational programmes and training to improve the number of skilled professionals.

Future Skills initiative of NASSCOM on new technologies is one such initiative. Its participants are member companies and have tied up with academia to offer its courses *[Futureskills]*. The government has several skilling and re-skilling programmes, while companies have taken the lead in building their workforces both on their own and in partnership with the government and academia. The goal is to make India the global hub for digital talents.

4.5.3 Promoting innovation in the cyberspace

- The objective of DSCI is to act as a catalyst for start-ups working in the cyber security space to come up with more innovative product ideas and address real risks, build resilience, increase trustworthiness and create a conducive environment for businesses. The initiative is an attempt to provide support to product companies in various aspects by bringing these new players nearer to established security leaders, innovators and other stakeholders on a common platform for idea sharing, guidance and collaboration.
- ***Bharat Electronics Limited*** (BEL) encourages start-ups to participate in supplies for its procurements in the areas of ML, AI and cyber security. BEL has recently relaxed the norms and eligibility criteria: prior turnover and prior experience. BEL has conducted a national level seminar on the role of start-ups in defence *[BEL]*.
- The Government Union Budget 2020 has a major focus on emerging technologies like AI *[Union Budget 2020]*. The proposal aims at bringing fundamental structural reforms and digital governance such as setting up investment advisory cell online to help young entrepreneurs with faster clearance and launch of seed funds to support early-stage start-ups come as a major booster for the sector.
- ***The Karnataka Government*** is in the process of developing a new IT policy that addresses IT start-ups regulatory challenges. The new policy is also believed to offer incentives for start-ups and entrepreneurs. The district innovation hub, an initiative by the Karnataka Government is to be set across different cities in Karnataka. The Karnataka government has interacted with Germany for start-up engagement, collaboration and innovation. The department of Information Technology; Biotechnology; and Science & Technology, Government of Karnataka, have signed a Memorandum of Understanding (MoU) with Bahrain Economic Development Board to promote cooperation in AI, cyber security, FinTech that would engage start-ups and technological firms establish partnerships with corporates, R&D, government agencies *[Karnataka Government]*.
- ***Incubation Centres at IITs***: Several IITs such as IIT Madras, IIT Delhi have set up incubation cells that fund and promote the development of technology start-ups in various domains including cyber security and AI. These incubation cells have played a major role in converting innovative ideas into commercially viable products. Many publicly sponsored incubators and accelerators are being run by academia and R&D.

5 AI-CS Infrastructure Development

This chapter discusses on the motivations to have datasets for R&D in cyber security, survey of global and national datasets and strategy to create datasets for R&D purposes. A list of R&D, academic institutions, and national cyber security organizations are proposed as implementing agencies. As a part of AI-CS infrastructure development a High Quality Node (HQN) is also proposed in this chapter.

5.1 Trends & Approaches towards Dataset creation

5.1.1 Motivations to have datasets for R&D in cyber security

Evolving high-speed network communication fabric over both wired and wireless medium and adoption of 5G, IoTs and the trend towards programmable Software Defined Network (SDN) aims to offer seamless dependable network service. However, there are also increasing concerns of equally developing scenarios from security concerns with the kind of newer type of attack surface that is originating and the highly sophisticated nature of these cyber security breaches. Hence, what is required is to enhance the defence mechanisms and try out various possibilities over virtual environments that are near-life network environments before adopting the same.

One of the major research challenges is the unavailability of a comprehensive network environment that can facilitate mimicking real-world scenarios for various verticals like Health networks/SCADA networks/smart cities network environment etc. Another important challenge is to create representative datasets which can reflect modern network traffic scenarios, vast varieties of low footprint intrusions and depth structured information about the systems, network, applications, devices and user patterns.

Though there are plenty of publicly available datasets on the internet to evaluate the intrusion detection systems, most of them have some limitations. The most popular datasets are DARPA 98/99 and KDD99 datasets and they are used by the researchers. However, numerous current studies showed that for the current network threat environment, these datasets do not inclusively reflect network traffic and modern low footprint attacks. As highlighted by Moustafa et al., these publicly available datasets have redundant and missing records in the training sets *[Moustafa 2015]*. Though there were attempts to normalize these datasets they lack in their comprehensive representation of modern data reflecting the attack environment.

Creation of attack datasets are challenging as it has to ensure its comprehensive coverage of various attacks, accuracy, completeness of data, specificity etc. Rodofile et al., discusses specific characteristic of attack datasets towards aiding attack dataset creation for SCADA environment such as ability to parse, replicate, sniff, inject, modify, replay, flood kind of capabilities for SCADA protocols *[Rodofile 2017]*.

Key challenges include the following:

- Non-availability of research data that is comprehensive, relevant and accurate for network and cyber security researchers.

- Lack of standardization for attack data sharing: Lack of standard procedures for data sanitization, formatting and sharing, clear legal and ethical policies for attack data sharing.
- Processed Metadata: The attack data available online is mostly incomplete and lacks attack semantics and its meta data covering various classes of attacks for multiple domains like health specific, SCADA specific, government specific, banking and financial institution specific etc.
- Region-specific data: In addition to all these, it would be useful to also have region specific data including cyber-attack datasets that are country specific along with context and other variables.

Researchers have started using the various AI driven algorithms to build the intelligent systems to determine the security threats. Hence, this is the right moment to create a national level infrastructure exclusively for creating datasets and enabling researchers try out various experiments over the wide-area network reflecting the real-world attack scenarios.

5.1.2 Survey of Global and national datasets

The trend of utilizing the intelligence gathered from large scale collection of attack data is becoming popular among the security community. Such intelligence offers better situational awareness and enables researchers and defenders to deeply understand the latest attacks and develop better preventive measures. Apart from providing situational awareness, the attack dataset has usability in the areas of 1) cyber threat prediction, 2) cyber security research, 3) testing & validation of security solutions and 4) cyber security training. A large repository of labelled attack data provided in machine digestible format, enables researchers and security experts to develop better machine learned models for detection and mitigation of the cyber-attacks.

Researchers use existing datasets (i.e., DARPA, KDD, CAIDA etc.) or setup their own test beds and simulate the attack scenarios to generate attack data. The conventional databases such as DARPA, KDD are more than 15-year-old and hence have limited attack coverage and have obsolete attack classes, where as in case of test bed approach due to the limited methods and guidelines available for generating anomalous behaviour, the attack data generated lacks completeness.

Few of the attack datasets being used at global level include the following:

- Canadian Institute for cyber security datasets are available for specific purposes like android malware datasets, android adware dataset, DDoS dataset, VPN & non-VPN dataset, ToR and non-ToR dataset etc.
- UGR'16 dataset covering various applications like SSH, SMTP, DNS, HTTP, HTTPS, POP3, SMB etc.
- Stanford Large Network Dataset Collection (Stanford Network Analysis Project-SNAP) covering wide range of various domains like social networks, communication networks, web graphs, internet networks, online communities, twitter and meme tracker, face-to-face communication networks etc.

Open Malware datasets:

EMBER: https://github.com/endgameinc/ember
ADFA-WD: https://www.unsw.adfa.edu.au/unsw-canberra-cyber/Cyber Security/ADFA-IDS-Datasets/
NGIDS – Linux: https://cloudstor.aarnet.edu.au/plus/s/f2fgM0lG1AMBKpu/download
Microsoft Malware Prediction: https://www.kaggle.com/c/microsoft-malware-prediction
Zeus: http://www.secrepo.com/pe_static/zeus.zip
Malware Detection - N Saravana: https://www.kaggle.com/nsaravana/malware-detection
Benign & Malicious PE Files: https://www.kaggle.com/amauricio/pe-files-malwares
Malware Analysis Datasets – Sections: https://www.kaggle.com/ang3loliveira/malware-analysis-datasets-pe-section-headers
Malware Analysis Datasets – PE: https://www.kaggle.com/ang3loliveira/malware-analysis-datasets-top1000-pe-imports
Mal-API-2019 - Ferhat OzgurCatak: https://www.kaggle.com/focatak/malapi2019

There are a wide number of sources with varied licensing conditions, accuracy and clarity offering datasets for research purposes.

5.2 Strategy to create Datasets for R&D purposes

5.2.1 Honeynet Based Attack Data Capture & Analysis

Based on the support from MeitY, C-DAC has developed and deployed honeynet based attack data capturing & analysis solutions being used to create standard labelled attack datasets. The attack data captured through deployments of honeypot sensors along with the analysis techniques is widely used for capturing, collection and generation cyber threat labelled events.

Generation of cyber threat dataset system automatically collects data from honeypots, malware analysis reports, and active honeypot systems and stores it in a structured format. The generated dataset contains several types of dataset such as malware, URLs, domains, IPs, botnet traffic, etc.

Dataset Creation: This involves a variety of methods, tools and techniques for feature extraction, static analysis, dynamic analysis, temporal analysis, sequence analysis and DL, etc. It is required to have a mechanism in place to automate the process of feature extraction from large amount of historical data and evolving feature space. DL algorithms facilitate addressing challenges regarding feature engineering and high-volume data and hence development of the system for the detection of attacks in Honeypot context is essential.

Attack Categorization*:* It is required to mark contextual information in the collected data through Honeypot sensors with threat indicators along with severity of attacks.

Semantic Analysis: The data captured by the honeypot sensors will be processed, classified and labeled. This labeled data will be converted into machine digestible formats and saved in the attack data repository. This data will be made accessible to the public with the help of a web portal.

5.2.2 Collaborative approach for dataset creation

For the creation of cyber security dataset to apply the AI based algorithms for security threat detection, there is a need to collect the cyber-attack data through collaborative approach. The attack data can be collected through multi-institutional support by creation of distributed sensor networks of attack data capturing, collection, enrichment and analysis to generate standard attack datasets.

The attack data capturing & analysis system provides solutions to the problems such as effective capturing of targeted cyber-attacks, lack of data repository of India specific cyber-attacks, effective and timely dissemination of cyber threat intelligence to cyber security agencies. Also, collection of datasets pertaining to various domains are essential, like Health-related datasets, Banking and Finance (BFSI) sector, SCADA sector and e-governance domain etc. To address these issues following processes are involved for dataset creation as shown in Figure 5.1:

- Capturing of targeted cyber-attacks
- Development of tools and techniques for cyber-attack analysis
- Development of system for effective and timely dissemination of cyber threat information in machine digestible formats

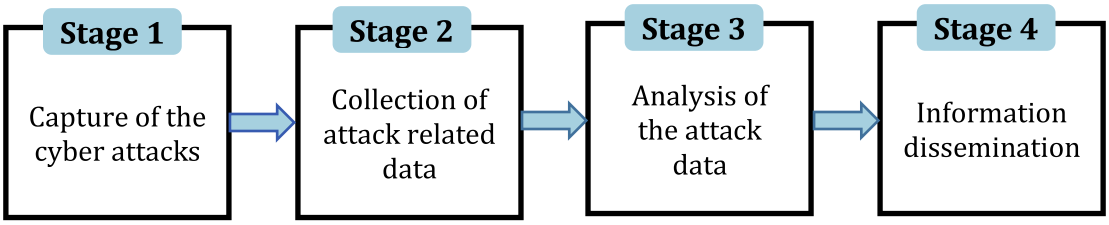

Stage 1
Capture of the cyber attacks
Stage 2
Collection of attack related data
Stage 3
Analysis of the attack data
Stage 4
Information dissemination

Figure 5.1: Capturing and analysis system workflow

For the purpose of attack data collection, the distributed network of honeynet sensors as threat capturing sensors can be placed at various geo-locations including C-DAC centres, NKN, CSIR-4PI, IITs, NITs, other academic institutions and ISPs. The threat capturing sensor could be placed either inside their networks with publicly exposed to the internet or by creating a separate network for attack data collection. The collected attack datasets will be further normalized and stored into big data-base to address the scalability issues. Then the security analysis techniques can be applied to extract the labelled datasets for the purposes of applying AI based detection models. Figure 5.2 demonstrates the model of various processes involved in dataset collection which includes distributed network of attack data capturing, collection and analysis.

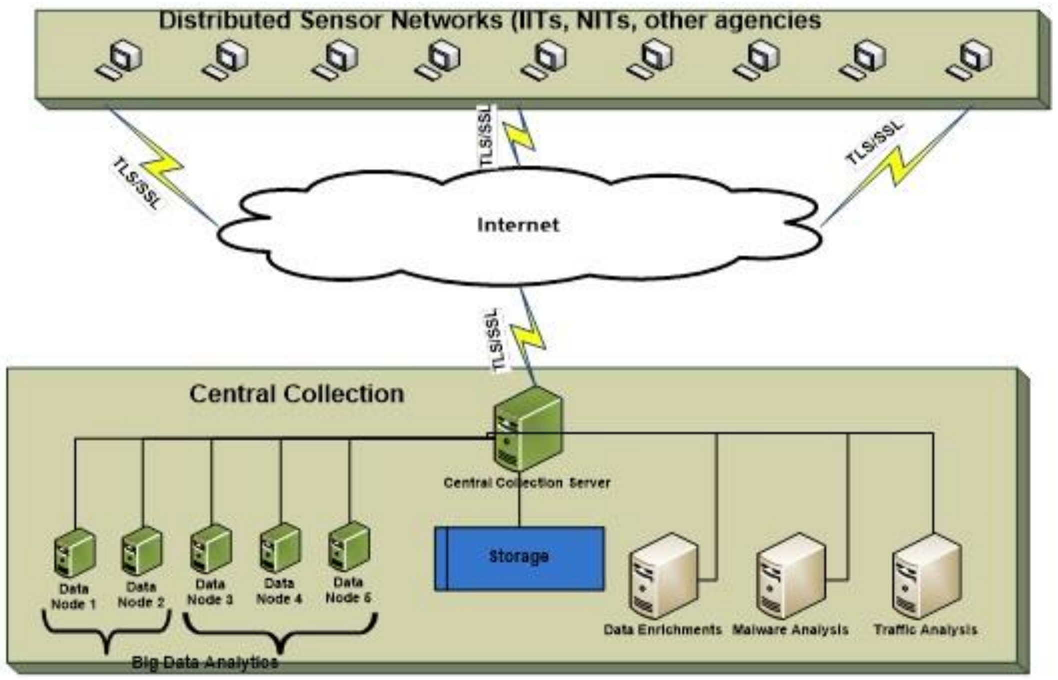


Figure 5.2: Distributed network of attack data capturing, collection and analysis

5.2.3 Social media attacks data collection

The enormous growth of social media usage has led to an increasing accumulation of data, which has been termed social media big data. Social media platforms offer many possibilities of data formats, including textual data, pictures, videos, sounds, and geolocations. Generally, this data can be divided into unstructured data and structured data *[Baars 2008]*. The social media analytics process involves four distinct steps, data discovery, collection, preparation, and analysis. The following Figure 5.3 is adapted from *[Stieglitz 2018]*. An end-to-end real-time situational awareness system which aims to retrieve security related information from social networking site Twitter.com is discussed *[Rodriguez 2020]*. It is proposed to create a research platform for data collection and analysis from social media sites for providing assistance to security analysts for evaluating risks and take action.

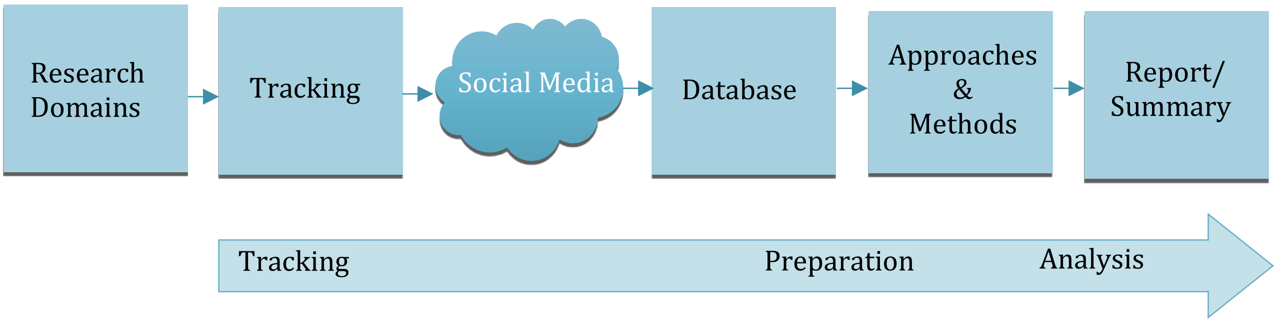


Figure 5.3: Social media data collection for analytics

5.2.4 Proposed list of Participating Agencies

The list of participating agencies for collecting the cyber-attack data through collaborative approach could be:

- CERT-In
- NKN
- C-DAC
- SETS
- CSIR-4PI
- Selected IITs, IIITs, NITs and other academic institutions
- Selected ISPs and start-up companies working in this space

Any other agencies interested shall also be roped in depending on the need such as sector specific agencies for healthcare, power sector, finance and e-governance sectors.

5.3 AI-CS infrastructure

The High Quality Node (HQN) of CyberSec4AI is proposed as a distributed node encompassing R &D labs, reputed academic institutes and industries as a team to implement the development strategy along with a shared facility to provide leadership, research, best practices and support in Education and Training in the area of AI and cyber security to meet the growing needs of manpower, infrastructure and indigenous systems, solutions and products. SETS may be considered as a secretariat to implement the CybSec4AI mission in addition to jointly implementing an HQN.

5.4 High Quality Node

An HQN of CyberSec4AI is recommended which can have a number of players addressing the issues and challenges together.

Long term objective: To serve as a High Quality Nodal centre for research, design, development and engineering of tools, systems, solutions and products for CybSec4AI.

Short term objectives:

- To do research, design and development of AI/ML based intrusion detection, anomaly detection systems, solution and products.
- To do research, design and development of deep learning based malware classification and model malware systems.
- To do research and development in the applications of AI/ML across "Cyber Kill Chain", i.e., identifying, preventing, responding to and recovering from attacks.
- To do R&D in feature engineering using knowledge about particular domains of Critical Infrastructure such as smart grid, smart cities.
- To create a dataset creation platform for CybSec4AI in association with NKN, Cert-In, NCIIPC, C-DAC and premier academic research institutions.
- To function as secretariat for implementing CybSec4AI mission.

Project duration and budget: Five years with a total outlay of Rs.50 crores.

Key objectives:

- Building core competencies, state-of- the- art infrastructure
- Strong collaborations (with govt., academia & industry–rope-in startup cos.)
- Define delivery models to the identified verticals
- Work on development and use of tools
- Data driven security models and quality datasets
- Compliance to standards (Setting effective AI and security related standards with preventive and recovery strategies)
- Regular hardware & software audits (ensuring the health of the systems)
- Promotion of innovative applications (periodic training and technology update)

Specific Research Activities:

- CS-AI can focus on a few selected domains (like SCADA, banking & health) to develop AI systems and identify and address the cyber security challenges and provide solutions (CS-AI).
- AI-CS can identify different areas where AI techniques can be applied to enhance cyber security, zero-day attacks hardware security and big data security.
- Addressing reliability of the operation results of ML systems
- Development of solutions for data filtration, outlier detection system, devising feedback mechanism to coexist with AI systems, behavioural signature analysis etc.
- Methods for automatic data collection, cleaning, labelling, categorization based on classes of attacks, anonymization etc.

It is recommended to set-up a national portal "CyberSec4AI-India" which can be launched to create a common platform for sharing ideas, expertise, datasets etc., to the various stakeholders such as researchers, faculties and developers. This will help in improving the current research going on thus enhancing our nation to compete globally.

The HQN involving R&D institutions, premier academic institutions and industry is proposed to implement the development strategy in the cross-discipline of cyber security and AI to meet the growing needs of manpower, infrastructure and indigenous systems, solutions and products. Select R&D labs and reputed academic institutes could jointly work in establishing HQN to arrive at the AI-CS infrastructure for R & D, necessary datasets, imparting training to the practising engineers and scientists. Based on the Taskforce's understanding of the present activities, one such possibility is IIT-Madras, IIT-Kanpur, SETS and C-DAC. Figure 5.4 provides an illustration of the key features of HQN and intersection of AI and CS. It also indicates the intersection of CS enabled AI and AI enabled CS.

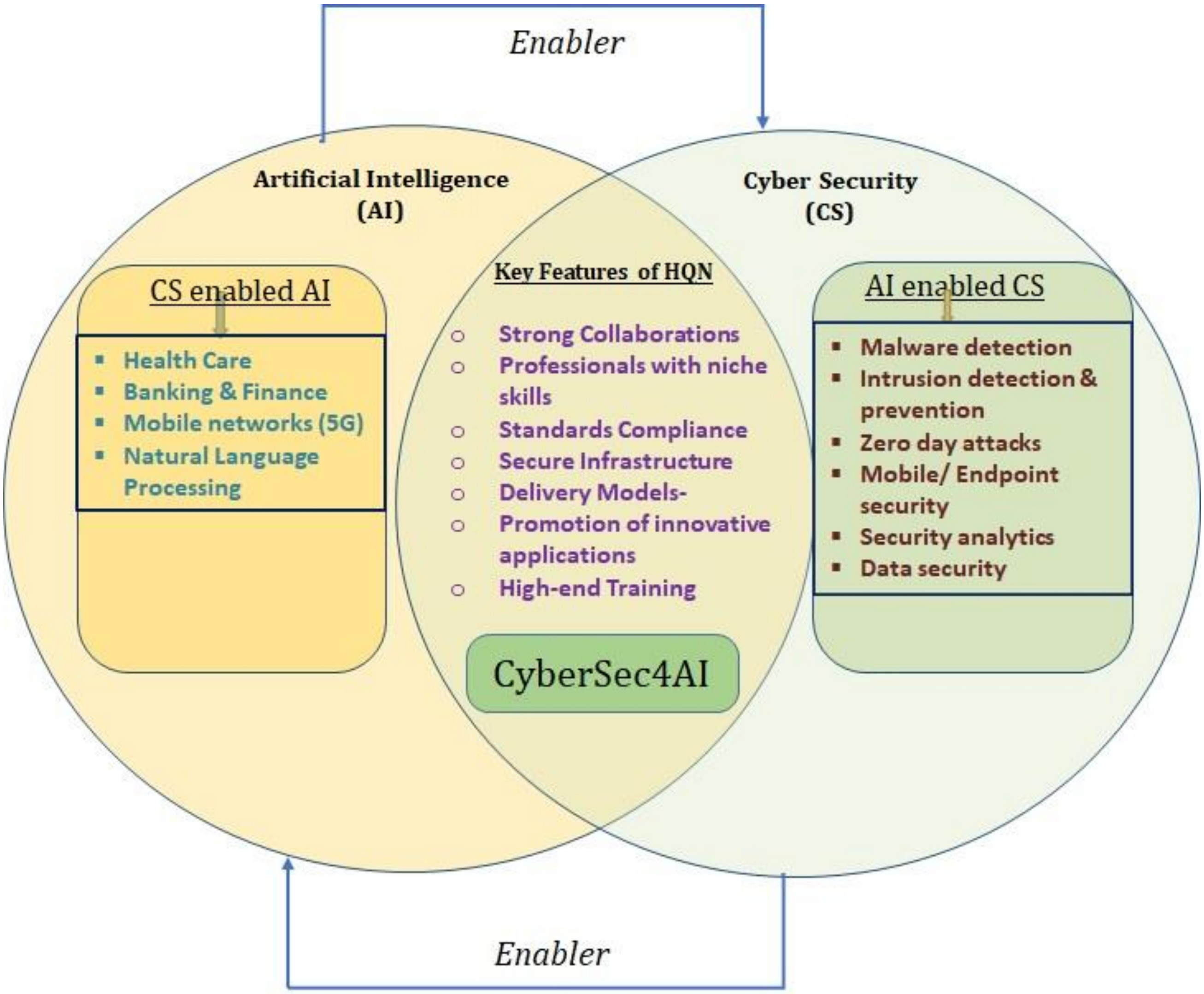


Figure 5.4: HQN - Intersection of AI and CS

6 AI-CS Education & Training: Capacity Building

> “There are risks and costs to a program of action - but they are far less than the long range cost of comfortable inaction.”
>
> – John F. Kennedy

6.1 Introduction

AI is changing the game for cyber security. It acts as a cornerstone in analysing massive quantities of risk data, in speeding up response times and in augmenting the capabilities of under-resourced security operations. As our nation has an amazing growth in the digital space, there arises a need of AI and cyber security experts to provide orchestrated responses against the data breaches.

In a report titled 'The Future of Jobs 2018', the World Economic Forum (WEF) said around 54 per cent of the global workforce had to be re-skilled or up-skilled to work in disruptive and digital technologies spawning the virtual world. In a report “Growing Cyber Security industry roadmap for India 2016”, NASSCOM and DSCI, charted out a vision 2025 to “grow the Indian cyber security products and services industry to US$35 billion, create one million cyber security jobs and 1000 cyber security start-ups by 2025”. However, the curriculum in colleges is being upgraded to incorporate such requirements in the syllabus both at the UG and PG levels, and researchers of various institutes can also be encouraged to work in AI and CS domains. Thus, the focus of this section is to illuminate on the methodology of addressing the human resource gap in AI for cyber security and cyber security for AI and to provide a framework on how these human resources can be generated over the next few years to meet the needs at national level.

The diffusion of digital technologies in India through digital India programme, the resultant application of latest technologies such as AI, ML and DL into different sectors of economy and the availability of big data from a variety of sources, is providing a larger threat canvas to Cyber attackers. This warrants a larger pool of skilled human resources at various levels to address the security needs of AI systems and also the human resources who are experts in AI and can use these techniques for enhancing AI based cyber security solutions. Thus, India is witnessing a growing demand for persons with AI and cyber security skills to draw actionable insights in securing the networks and systems. Moreover, India with more than 50% of the IT/ITES outsourcing and knowledge outsourcing, has an opportunity to provide AI & CS services to the world.

There are some efforts happening from NITI Aayog in the area of AI and to some extent in the area of cyber security. MeitY has launched “Information Security Education & Awareness” (ISEA) programme over the last 10 years exclusively on cyber security focussing on research, education, training, and awareness with the support of top notch academic institutes (IITs, NITs), national level R&D labs (C-DAC) and training centres (NIELIT). Recognising the importance of research and its commercial adoption in success of AI, NITI Aayog has proposed setting up of Centre of Research Excellence (CORE) to focus on developing the better understanding of existing core research.

India having one of the highest concentrations of IT manpower in the world, the manpower needs to be trained in the areas of AI and CS to take benefit out of the existing resources. It is also imperative to focus on the quality human resources. The following sections would address the human resources building requirement in AI and CS in India, in a planned and holistic way.

It is proposed to address two aspects of skills development: cyber security for AI and AI for cyber security. The model supports nearly 100 PhDs and 100 M.S fellowships in the area of AICS. It aims in providing Industry oriented courses, online courses and Faculty Development Program (FDP) for faculties from engineering colleges, universities, polytechnics and high school. The proposed model is depicted in Figure 6.1.

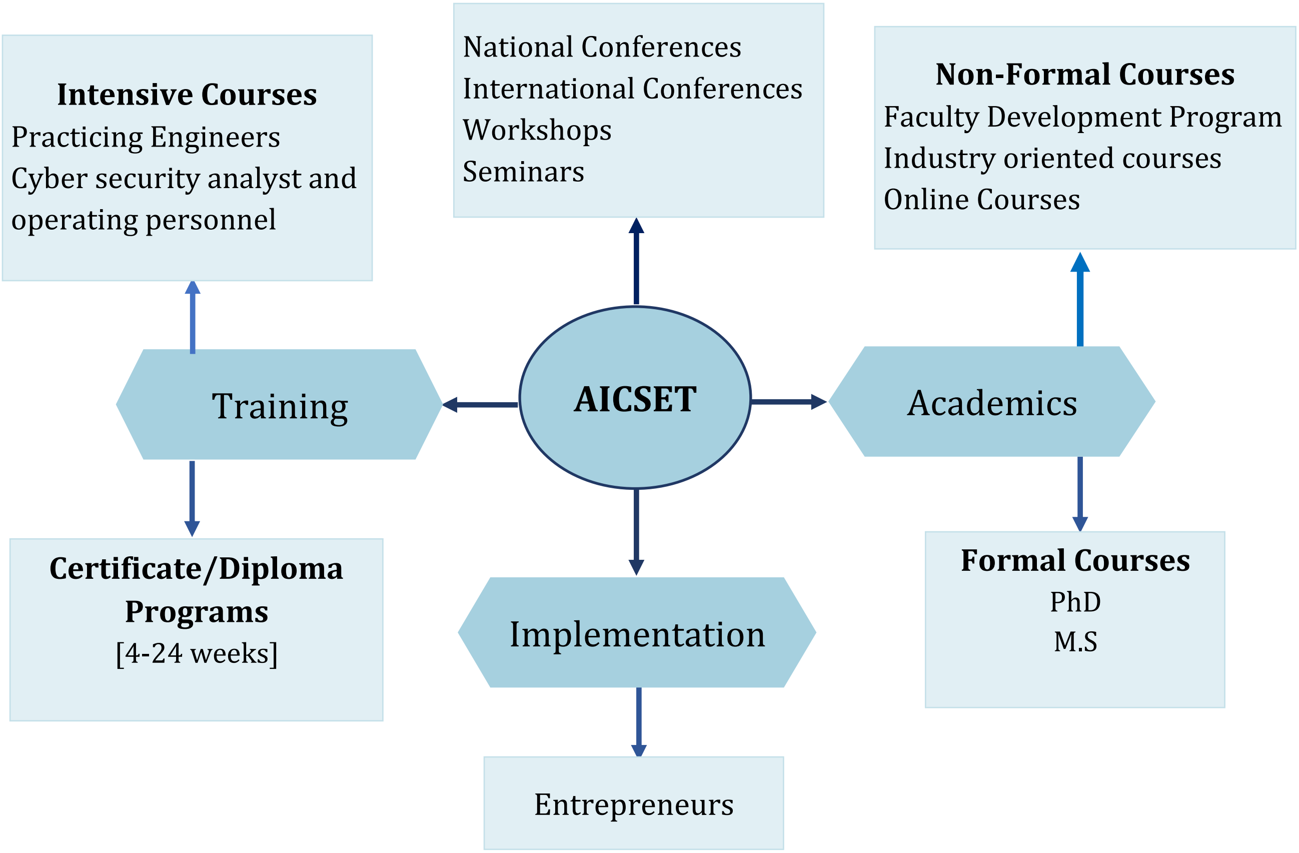


Figure 6.1: AI and CS Education and Training (AICSET)

6.2 Objectives

Following are the objectives of AI and CS education and training model

1. *Capacity building in the area of AI-CS is to address the human resource requirement in the country which mandates:*
    - Generation of core research manpower in the AI-CS domain to undertake basic/fundamental research, applied research, research in the area of product/solution design, and development in specific areas of national strategic importance to build indigenous capability.

- Introduction of AI-CS curriculum in formal courses like B.Tech. CS, M.Tech. CS, ME, MS in Physics/Mathematics/Statistics, MCA, MBA Information Systems, and post-graduate diploma courses etc., faculty training, modular short-term knowledge oriented courses, certified assessment schemes, and internships.
- Non-formal courses & short-term courses for professionals in classroom format and some as Massive Open Online Course (MOOC) offerings.

2. *Capacity building of domain specialists from govt. research labs, line ministries and industry verticals (e.g., banking, healthcare, agriculture, etc.) in AI-CS that requires to;*
   - Conduct awareness programs on innovation opportunities of AI-CS
   - Training programs on advanced AI-CS technologies
3. *Building expertise in use of AI/ML/DL tools for cyber security*
4. *Capacity building of higher secondary school teachers and students in AI-CS concepts & skills;*
   - Collaborate with NCERT, CBSE to develop training programs for high school teachers on AI-CS and practical applications
   - Develop and provide Free and Open-Source Software (FOSS) tools-based school edition AI-CS tool kits for use by school students
5. *Conduct international, national conferences, workshops and seminars*
6. *Develop a website to serve as a portal for all information about CyberSec4AI activities, events, resources etc., and include discussion forums, newsletters, technology forecasts and market trends, etc.*

6.3 Salient features

In order to achieve the objectives as indicated, and to meet the growing demand for skilled manpower for the various job roles in AI-CS, it is recommended to launch a national level AI-CS Capacity Building Program (AI-CS-CBP) for a five-year duration.

The main components are:

- Organization model for implementing the national AI-CS-CBP activities
- Generation of core research manpower in AI-CS through PhDs and Postdoctoral Fellowships
- Formal M.Tech. degree focused on AI-CS and adding AI-CS courses/topics appropriately at the UG level
- Faculty development training programs
- Short-term non-formal courses, certification schemes, MOOC programs for capacity building of working IT professionals
- Capacity building and AI-CS benefits awareness programs for domain specialists in govt. research labs, line ministries and industry verticals
- International, national conferences, workshops, seminars
- AI-CS-CBP portal- to serve as a single window for all information about AI-CS-CBP activities, events, resources etc., and include discussion forums, newsletters, etc.

Each of these components is briefly described in the sections given below.

6.4 Structure for Implementation

A virtual organizational structure is proposed for the implementation of AICS-CBP and is illustrated below:

Identifying and operating Resource Centres (RCs)

It would consist of select premier academic institutes and research labs in technology, management and statistical domains who have significant research activities and deep expertise in these areas. Examples of Resource Centres (RCs) are: IITs, IIITs, NITs, IISc., SETS, C-DAC etc. An RC would be an academic/engineering institute/government R&D lab having deep expertise in multiple areas relevant to AI and CS. An RC would have sizeable number of trained faculty in the respective areas of expertise. Institutions, organisations and government R&D labs, offering formal/non-formal courses and wanting to enhance its course offerings and the skills of its faculties/members or start new courses in AI-CS topics, will also be involved, in association with the RCs.

Generation of core research manpower in AI-CS

Select RC will be identified for AI-CS research and development through PhD and Postdoctoral Research Fellowships. The AI-CS R & D RC are proposed to be setup at premier academic institutions like IITs, IIITs, and R&D labs, which could be identified to undertake research in domain specific AI-CS research and products/solution development.

Some thematic areas for AI-CS research could be: AI for network security; AI for big data security; cyber security for AI based domain specific applications etc.

The RC will also need to undertake faculty development and development of high-end quality manpower.

The faculty of RCs could also become the pool of prospective candidates for Doctoral/Postdoctoral programs offered at select RCs.

The key responsibilities of RCs would be to get involved in carrying out R&D, academic activities and to be in constant interaction & co-ordination with the AI-CS-RCs, national/international community and industries working in this area. They are also required to guide the research initiatives, PhD students and conduct advanced faculty training courses in emerging and niche areas of AI-CS. They will be involved in the syllabus formulation for formal courses in AI-CS, development of learning material, offering MOOCS courses, online labs, design and offer certification schemes in AI-CS, in identified specialized areas for use by RCs and they have to co-ordinate in conducting of international, national conferences/seminars.

6.5 Formal Courses, Faculty Training, Short-term courses and Certification schemes

It is proposed to scale up formal courses and faculty update programs in the area of AI and cyber security at various levels. This would mainly include B.Tech., M.Tech., Doctoral/Postdoctoral, short-term specialized/modular courses, etc. This is proposed to be achieved through a mechanism/structure comprising of RCs.

a. ***Some of the formal Courses that could be introduced at various levels in AI-CS***
   - M.Tech. in AI-CS
   - Retrofitting of AI-CS courses in B.Tech./BE and M.Tech./ME courses in Computer Science/Information Systems, MCA, MS in Mathematics/Statistics/Information Systems.
   - Research programme leading to Postdoctoral/PhD.

b. ***Non-formal courses & Short-term courses for Professionals***

   A suitable methodology for launching non-formal courses and short term courses for updating skills/competencies of working professionals through RCs needs to be worked out. Some of the non-formal modular courses/short term training programs could be offered as MOOC by RCs. These could include:
   - Short-term courses– 6 months certificate course (modular and Integratable to PG level)
   - Certification programmes aimed at developers, security analysts, network system administrators, on specific AI-CS technologies

   Specialised courses on CybSec4AI for practising engineers and scientists need to be organized with national and international experts to jumpstart the development work in building indigenous tools, systems and products.

   The intensive course can be offered for a period of two weeks, for practising engineers from R&D labs, industry and researchers working in this domain. Crash course on ML would be taught for one week followed by case studies on cyber security with lab sessions and demonstrations.

   The draft syllabus of the intensive courses to be organized includes the following: Bayesian methods, regression, clustering, principal component analysis, support vector machines, neural networks, introduction to reinforcement learning, spam detection, security and privacy of machine learning systems, decision trees, boosting methods, bagging methods including random forest, text analytics and graph convolutional networks.

   In the first phase, four 1week courses could be planned to enhance the skills development of working professionals. An illustrative list of short term training programs are:
   - AI essentials for cyber security professionals
   - Cyber security for AI professionals
   - AI and CS for computer/IT professionals
   - Foundations of maths/statistics to AI/CS professional

c. ***Training the trainers***

   It is recommended that similar programs like T10kT (supported by National Mission on Education through ICT) for training a large number of trainers, could be launched for CybSec4AI. It is a tried-and tested model. Use of an online and blended approach is proposed that allows participants to complete a significant part of training online, thus reducing the time which must be spent on face-to face synchronous interaction.
   Details are provided in Annexure I.

d. ***Capacity building of domain specialists in government research labs, line ministries and industry verticals through professional training programs in AI-CS***

- Conduct awareness programs on AI-CS driven innovation opportunities, benefits and impact of AI-CS in domain verticals
- Training programs on AI-CS technologies

e. ***AI-CS Portal***

A web portal would be developed to provide a single window for all activities, events, newsletters on AI-CS-education, training & research activities. It can serve as a portal for knowledge sharing and dissemination in AI-CS for various stakeholders.

6.6 PhD Fellowships, Internships and Travel Support

a. PhD Fellowships

It is proposed to adopt PhD fellowship models similar to EU consortium "Privacy & Us" that seem to be working well for properly defined research topics. Details are provided in the Annexure IV.

Implementing a similar concept for Cyber4AI topic within India can be considered, albeit at a bigger scale, say, for 100 PhDs, across multiple Indian universities/institutes with support from both government and industry. Governance could be light, as periodic meet ups would lead to a lot of peer reviews and oversight.

Details of no. of PhD fellowships support requirement per year as provided by responses to the questionnaire on CybSec4AI is provided in Annexure III.

b. Internships to Graduate & PG Students

It is important to offer internships to graduate & PG students in streams of statistics, mathematics, electronics and communication, computer science and computer application, to give them exposure in practical projects employing AI-CS. RCs and PIs could offer graduate and post-graduate students internship projects involving application of AI-CS skills and knowledge.

c. Travel support for international conferences

It is felt that the participation of the research community in the international events such as conferences, workshops is growing, though the numbers are still small compared to the similar participation from countries like China. In order to increase the participation of Indian researchers, faculties and practising engineers in international events, it is recommended that adequate travel support for attending international conferences in the specialized domain of CybSec4AI may be provided as a part of the mission.

6.7 National Workshops, Seminars and National & International Conferences

The Resource Centres (RC) would organize international & national conference, workshops and seminars for knowledge dissemination in AI-CS:

- One national and one international conference annually
- Quarterly national workshops/seminars

a. National level workshops to be organized

1. Since CybSec4AI is still an emerging research area, workshops may be organized with invited experts from India and international institutions. It is proposed to organize two international level conferences and five national level conferences/workshops in cyber security and AI.
2. It is also proposed to explore the possibility of having special sessions on CybSec4AI in the following conferences
   - International Conference on Data Sciences (CODS)
   - International Conference on Information and Communication Systems (ICICS)
   - International Conference on Security, Privacy and Applied Cryptography Engineering (SPACE)

b. International level conferences to be organized

It is proposed that at least one international conference be conducted per year. Experts in related areas would be invited to give lectures. Contributed papers will also be published after peer review.

6.8 Budget

Following Table 6.1 shows the budget estimated for a period of five years which is to be supported by GOI for the successful implementation of all the components listed in above sections under AI-CS-CBP.

Table 6.1: Estimated Budget for the period of 5 years

S.No.	Budget Head	Number (for Five Years)	Cost for Five Years (in Rupees Crores)
1	Intensive Training	20	10.00
2	Training the Trainers	2500 trainers	100.00
3	PhD Fellowships	100	7.00
4	MS Fellowships	100	3.00
5	Travel Support (International) for researchers, faculties and practising engineers	75	5.00
6	National Conferences/ Workshops	20	3.00
7	International Conferences	05	2.00
8	Infrastructure set-up for 10 Resource Centres (RCs)	10	25.0
	Total		***155.00***

This chapter on "AI-CS Education & Training: Capacity building" focuses on the AI-CS Capacity Building Program (AI-CS-CBP) model to bring the mission of the Task Force report into action. It has listed out various components for the implementation of the AI-CS-CBP model. The model tries to develop the required skills for the spontaneous growth of research and training in the area of AI-CS at all levels from high school to Postdoctoral Fellowships, also including training the trainers for the development of "intelligent security" model for our nation.

7 R&D under CybSec4AI

7.1 Introduction

Based on the discussions held during Task Force meetings and inputs received from R&D labs, academia, industry, published research work and the study of international and national reports, it is proposed to initiate research projects in the following areas, leading to development of systems, products and solutions to meet the growing challenges of CyberSec4AI. However, additional areas of R&D can also be initiated when the need arises.

7.2 Proposed R&D areas

7.2.1 Adversarial ML and Robust AI

There has been growing interest in the field of adversarial machine learning to identify vulnerabilities in the learning-based classifiers and the design of suitable counter measures. Recent research work on Test Time Evasion (TTE), Data Poisoning (DP), Back door DP, Reverse Engineering (RE) attacks and defences against attacks are surveyed *[Miller 2019].*

R&D areas in this topic will be theoretical study of existing defence against adversarial examples and practical demonstration of crafting adversarial examples for a ML-based malware detection system.

Some of the developments could include:

i. Develop attack strategies against ML both at training time (poisoning) and at test time (evasion)
ii. Propose systematic methodologies for security evaluation of learning algorithms against attacks
iii. Design suitable defense mechanisms to mitigate these threats
iv. Develop models for threats against learning algorithms and deep networks
v. Theoretical study of existing defence against adversarial examples
vi. Practical demonstration of crafting adversarial examples for a ML-based malware detection system

7.2.2 Malware Detection and Anomaly Detection

It is proposed to conduct research, design and development of AI/ML based malware detection, anomaly detection systems, solution and products in the following areas:

- Semi-supervised anomaly detection
- Unsupervised anomaly detection
- Explainable models for anomaly detection
- Human-in-the loop anomaly detection

The complexity and dynamics of intrusion detection can be addressed efficiently by representation learning and function approximation capabilities of DL and optimal sequential decision making the capability of reinforcement learning collectively. Thus, there is a gap for future research where the capabilities of Deep Reinforcement Learning

(DRL) can be exploited fully to solve complex and sophisticated intrusion detection problems *[Nguyen 2019]*. Exploration of model-based DRL methods and model-free methods for cyber defense is an interesting future study.

7.2.3 Cyber Attack Projection, Prediction and Forecasting

The cyber security attack surface is massive and growing rapidly. It is no more a human scale problem. In a proactive approach, there is a need to preemptively infer the upcoming malicious activities so that reactive action can be taken before any harm is done. In such a case, it is possible to predict the next steps and is known as attack projection. A similar task is intention recognition i.e., estimate the goal of the adversary, which can help in predicting adversary's next move *[Husak 2018]*. Research in attack prediction in a collaborative environment such as collaborative intrusion detection could be taken up.

It is important to study AI based cyber-attacks and to map them onto a framework that includes classification of several aspects of malicious uses of AI during the attack life cycle and providing a basis for their detection in order to predict future threats.

It is proposed to encourage research and develop AI/ML based approaches in the cyber-kill-chain i.e., identifying, preventing, responding to, and recovering from attacks.

7.2.4 Automation of Cyber Security Operation Centre (SOC)

The growing attack surface includes amateur threats, such as phishing, sophisticated DDoS attacks, and skilled state actors. Prevention is nearly impossible. Automated probing will eventually find weakness. Advanced persistent threats show that hackers are patient. The biggest problem in cyber security is not better end-point detection but how to enable the analyst to keep pace with the sheer volume of alerts being generated *[Bresniker 2019]*. The behavior, thought process and actions of cyber security analysts need to be observed. It is proposed to explore capturing and understanding the actions of cyber security analysts by using a grand challenge. Game theoretic models for dynamic security paradigms such as moving target defence could also be explored.

7.2.5 Differential Privacy (DP)

R&D areas like theoretical study of DP techniques and their applicability to specific ML systems, development of ML tools incorporating DP techniques, theoretical study of privacy vs. cost trade-off and representative implementation of a system, including data test bed, DP-enabled ML tools and sample statistics/summaries which can be published will be pursued.

R&D areas in this topic proposed to be taken up are as follows:

i. Theoretical study of DP techniques and their applicability to specific ML systems
ii. Development of ML tools incorporating DP techniques
iii. Theoretical study of privacy vs. cost trade-off
iv. Representative implementation of a system, including data test bed, DP-enabled ML tools and sample statistics/summaries which can be published

7.2.6 Cryptographic Techniques in ML

Cryptography and cryptanalysis can propel people towards using `Machine Learning as a Service'. Many cryptographic primitives, like homomorphic encryption, secret sharing and multiparty computation, have been used in ML systems for achieving the above goal. It is desirable to search for crypto solutions solving the problems of protection against active adversary, model proving and access control in the context of machine learning.

It is desirable to develop crypto solutions solving the following problems:

i. Can cryptographic techniques be applied in the ML system so that the adversary cannot manipulate the data to fool the model?
ii. Can a framework be designed to prove that the model that has been arrived at based on given data is correct?
iii. Can cryptographic techniques be used to trace unauthorized data/model accesses?

R&D areas in this topic are as follows:

i. Modelling of unauthorized access to data/model
ii. Complete implementation of a ML system where the data, process and result are all secure
iii. Explore whether some ML procedure can be replaced with a cryptography friendly one and gauge the trade-off between cost and security
iv. Design specialized crypto primitives for application to ML systems

7.2.7 Resilience of ML classifiers

Multi-sensor fusion is the basis of most modern cyber defence systems in which a variety of sensors are deployed throughout the defending organization and the sensors readings are analysed collectively in an attempt to identify attacks *[Katzir 2018]*. The classifiers are susceptible to manipulation by an adversary as a mathematical game. Inherent resilience of different classifier algorithms is to be explored. The random forest classifier demonstrated superior resilience, suggesting that ensemble-based classifiers are inherently more resilient to adversarial attacks *[Loaiza 2019]*.

It is proposed to initiate research projects to study the resilience of ML classifiers used for cyber security.

7.2.8 Attacks related open questions

The main goals of an adversary are:

i. To retrieve sensitive private training data (privacy breach, membership inference)
ii. To poison the data that is being used in learning (model integrity compromise)
iii. To fool the models to make incorrect predictions (model integrity compromise)
iv. To duplicate/extract the proprietary models (model inversion)

Research projects may be initiated to address the following open questions:

i. How to detect when an AI model is fooled?
ii. How to find robust defences invariant to changing environments?
iii. What processes need to be followed to mitigate the misuse of AI?

iv. What architecture and policies to be implemented to ensure security of AI and privacy of data?
v. How to ensure trustworthiness and transparency of the models?

7.2.9 AI based Social Media Attacks

There have been studies on adopting ML algorithms to carry out complex social engineering attacks. There are essentially two kinds of Facebook fakes. One is bot account that is created and operated remotely via software. The other sock-puppet is a false account that is operated by a human being pretending to be someone or something they are not.

Social Bots: Attackers could leverage ML based techniques to build intelligent botnets made of autonomous intelligent bots. These intelligent bot scans decide on the fly what shall be done according to the context, mission and targets. Intelligent botnets could allow bots to conduct reconnaissance on their surrounding environment and make decisions on their own. DeepPhish, an AI algorithm that produces new synthetic URLs by learning patterns from most effective URLs in historical attacks has been presented *[Kaloudi 2020]*. DeepPhish uses LSTM model to implement a phishing URL classifier *[Bahnsen 2018]*.

An example of weaponization of social media platforms is the automated spear phishing framework on Twitter. Seymour et al., used a RNN to demonstrate the automation of attack payload in the phishing process and leveraged data science to target users with personalized phishing messages *[Seymour 2016]*.

A spam filtering approach has been proposed using ensemble learning techniques with regularized deep neural networks as base learners to show better performance in comparison with decision trees, naive bayes, support vector machine etc. *[Barushka 2018]*.

Garg et al., proposed a hybrid DL based anomaly detection for suspicious flow detection in Software Defined Networking (SDN) in the context of social media *[Garg 2019]*. It consists of an anomaly detection module that leverages improved restricted boltzmann machine and gradient descent-based support vector machine to detect the abnormal activities. The performance in terms of detecting malicious events such as identity theft, profile cloning, confidential data collection etc., is analysed.

The above research papers and the papers referred in chapter 3, indicate some of the important research work that have been carried out on social media attacks using AI/ML techniques. It is proposed to initiate research projects on social media attack and mitigation techniques using AI/ML algorithms to address new challenges.

7.2.10 Use of AI for CPS Security including IoT Security

The concept of digital twin is at the forefront of the Industry 4.0 revolution facilitated through advanced data analytics and the Internet of Things (IoT) connectivity. IoT has increased the volume of data usable from manufacturing, healthcare, and smart city environments. A digital twin is a virtual instance (or twin) of a physical system that is continually updated with the latter's performance, maintenance, and health status data throughout the physical system's life cycle *[Madni 2019]*.

Recent research efforts are indicating that the data integrity attacks are able to bypass the bad data detection mechanism and make the system operator obtain the misleading states of the system, leading to massive losses in cyber physical systems such as smart grid. Kaloudi et al., evolved a framework to analyze AI based cyber-attacks in hypothetical scenarios of a critical smart grid infrastructure *[Kaloudi 2020]*.

It is proposed to carry out research leading to evolving architectures to address resilience and security of AI based systems, which are likely to be used in critical infrastructure such as smart grid, smart cities, etc.

7.2.11 Wireless Security in the context of ML enabled 5G/6G services

AI/ ML technologies are increasingly being applied to network management and cyber security solutions. With the growing complexity of SDN/NFV and IMT2020/5G networks and beyond, ML may be well applicable for automatic network orchestration and network management. Wang et al., studied how AI/ML can be leveraged for the design and operation of beyond 5G wireless networks *[Wang 2020]*. The 5G and beyond wireless networks are critical to support diverse vertical applications connecting heterogeneous devices and machines, which directly increase vulnerability for various spoofing attacks and an intelligent authentication is designed to enhance security performance *[Fang 2019]*. It is proposed to initiate research projects to address growing security challenges in ML based network management systems in 5G and B5G networks.

7.2.12 Development components for building AI based cyber security solutions

The following development components/areas are suggested for building industry competitive AI based solutions/products.

- End-system profiling and detection of malicious activities
- Network anomaly detection
- Cyber threat analysis
- Securing Edge/IoT/SDN environments in the context of 5G
- Botnet traffic detection
- Incident analysis and forecasting
- Malware analysis
- SCADA & mobile communications
- Fraud detection
- Privacy preserving systems
- Situational awareness
- Insider threat detection
- Bio-inspired approaches for cyber security
- Digital forensics

7.3 IPR model for collaborative R&D

> "The real question is, when will we draft an artificial intelligence bill of rights? What will that consist of? And who will get to decide that?"
>
> - Gray Scott

It is recommended that an IPR model for collaborative R&D could be evolved similar to the Australian Cyber Security Cooperative Research Centre (CSCRC) where industry, government and research partners work to create new products, services and systems that deliver a secure and resilient national Cyber Security capability. The details are provided in Annexure II.

Possible R&D areas proposed by Indian Institutions and R&D Labs:

Following are the possible R&D areas proposed by Indian Institutions as per the questionnaire on CybSec4AI:

- Adversarial AI
- Pattern recognition
- Anomaly detection, misbehaviour detection, automatic vulnerability detection, intrusion detection, fraud detection, user/machine behavioural analysis, encrypted traffic monitoring and analysis, early detection of malware & anomalies
- Digital evidence analysis, block chain
- Cloud security, IOT security, privacy
- Cryptanalysis using AI algorithms
- Can cryptographic techniques be applied in the ML system so that the adversary cannot manipulate the data to fool the model?
- Can a framework be designed to prove that the model that has been arrived at based on correct data can also be evolved?

Annexure III provides the details of the currently existing international collaborations with Indian institutions and R&D labs.

7.4 Major areas/institutions identified for International collaboration

A suggestive list of research departments working on cyber security and AI in various universities abroad are identified, initially, for possible research collaborations with Indian academic and research institutions and are given in Annexure X.

8 Time schedule and Budget requirement

It is proposed to initiate R&D Programme in four schemes with a (i) focus on basic research leading to research papers, (ii) collaborative projects with international institutions leading to joint research publications and prototypes; (iii) product development leading to advanced prototypes and patents and (iv) mission mode projects jointly with R & D labs, industry, and academic research institutions leading to industry-competitive tools, systems and solutions.

(i) **High-end academic research with international research collaborations**
Deliverables- Papers in reputed journals, conferences, PhD and M.S students
Project Outlay- Rs.50 Lakhs to 100 Lakhs

(ii) **Academic and R&D labs with international research institution collaborations**
Deliverables- Prototypes and patents, joint research publications
Project Outlay- Rs.100 Lakhs to 200 Lakhs

(iii) **Product development R&D labs with industry/start-up**
Deliverables- Advanced prototypes and engineering and product development for stipulated domains
Project Outlay- Rs.200 Lakhs to 500 Lakhs

(iv) **Mission mode projects such as building industry- competitive indigenous tools, developing domain specific models involving end users/verticals with industry/start-ups**
Deliverables- Cyber security tools, models and systems using ML/AI for stipulated domains
Project Outlay- Up to Rs.50 Crores

The mapping of outcome of the R&D schemes to Technology Readiness Levels (TRLs) are illustrated in Figure 8.1. Software Technology Readiness levels of NASA are taken as reference for mapping *[GAO 20-48G-2020].*

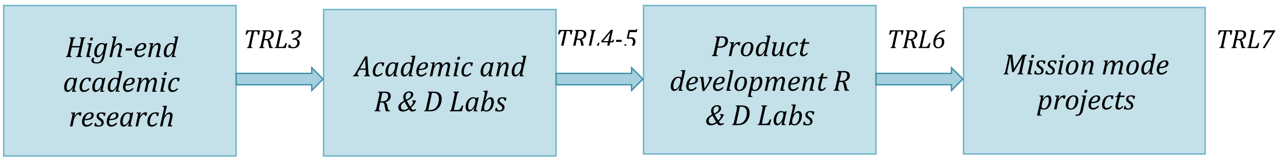


Figure 8.1: Mapping of the outcome of various R & D schemes proposed to the Technology Readiness Levels

8.1 Budget Requirements

I. R&D budget:

1. Mission mode projects for development of indigenous tools and models by R&D institutions and industry/start-ups: 5 Projects
   Total Cost: Rs.250 Crores.

2. Product development in R & D labs and industry/start-ups leading to advanced prototypes, engineering and patents: 20 Projects
   Total Cost: Rs.100 Crores.
3. Academic and R & D labs with international research collaborations leading to prototypes and patents: 10 Projects
   Total Cost: Rs.30 Crores.
4. High-end academic research collaborations leading to research patents/papers in reputed journals, conferences and research students: 100 Projects
   Total Cost: Rs.50 Crores.

II. ***PhD fellowships, training and skill building***

1. Number of PhD fellowships: 100 Nos
   Total cost: Rs.7 Crores.
2. Number of M.S fellowships: 100 Nos
   Total cost: Rs.3 Crores.

III. ***Travel support for attending international conferences for researchers, faculties and practising engineers***

75 travel scholarships for 5years
Total Cost: Rs.5 Crores for 5 years.

IV. ***Support for organizing national and international conferences***

Number of national conferences: 4 per year
Total cost: Rs.3 Crores for 5 years.
Number of International conferences: 1 per year
Total cost: Rs.2.5 Crores for 5 years.

V. ***Intensive training***

For practicing engineers/scientist: 4 per year
Total cost: Rs.10 Crore for 5 years.

VI. ***Training the trainers (similar to National Mission on Education through ICT)***

Number of trainers to be trained: 2500 for 5 years
Total Cost: Rs.100Crores for 5 years.

VII. ***High Quality Node:*** Rs.50 Crores

VIII. ***Infrastructure set-up for Resource Centres (RCs)***

Number of RCs: 10
Total cost: Rs.25 Crores.

Total Budget: Rs.635 Crores

It is estimated that the mission program on AI4-CS needs a budget support of Rs.635 Crores from GOI, towards all the components enlisted in the report, spread over a period of 5 years. Table 8.1 provides the detailed expenditure model on various recommended components distributed year wise.

Table 8.1: Projected budget requirement

S.No.	Budget Head	Amount in Rupees Crores					
		I Year	II Year	III Year	IV Year	V Year	Total
1	Mission mode projects	40	60	60	50	40	250
2	Product development	15	25	25	20	15	100
3	Academic and R&D projects with international collaboration	5	6	7	7	5	30
4	High end academic research	8	12	12	11	7	50
5	PhD and M.S fellowships	1.5	2.5	3	2	1	10
6	Travel support for attending conferences	0.75	1	1.25	1	1	5
7	Support for organizing national and international conferences	0.75	1	1.25	1	1	5
8	Intensive training for practising engineers	1.75	2.25	2	2	2	10
9	Training the trainers	15	25	25	20	15	100
10	High Quality Node (HQN)	7.5	8.5	10.5	11	12.5	50
11	Infrastructure set-up for Resource Centres (RCs)	4	6	6	5	4	25
Total		99.25	148.75	152	129.5	103.5	635

9 Recommendations of the Task Force

The "Task Force on CybSec4AI" was formulated to identify the development strategy for India in the cross discipline of AI and cyber security. The task force deliberated the need for initiating R&D schemes resulting in system developments, solutions and products in the cross discipline of AI and cyber security. The human resources required to address the growing requirements of engineers and scientists in this emerging area have been identified and schemes to accelerate the skill building are proposed. The necessary R&D infrastructure required for data creation, simulation and testing facilities are also proposed.

9.1 A High Quality Node of CyberSec4AI

A High Quality Node (HQN) needs to be created, that can have a number of players together addressing the issues and challenges of CyberSec4AI. This HQN shall have:

- A vision for the node with actionable mission
- Efficient and multitalented manpower base (professionals with niche skills)
- Ambitious roadmap for the next 5 years for knowledge creation through R & D and its dissemination through education and training

The focus shall be on the following aspects:

- Building the core competencies, state-of- the- art infrastructure
- Strong collaborations (with govt., academia & industry – rope-in start-up companies)
- Define delivery models to the identified verticals
- Work on development and use of tools
- Data driven security models and quality datasets
- Compliance to standards (setting effective AI and security related standards with preventive and recovery strategies)
- Regular hardware & software audits (ensuring the health of the systems)
- Promotion of innovative applications (periodical training and technology updation)

The HQN can be implemented as a distributed node involving R & D institutions, academic institutions and industry. The identified institutes (one such possibility is IIT-Madras, IIT-Kanpur, SETS and C-DAC) could jointly work together in establishing the HQN to meet the requirements of the AI-CS infrastructure for R & D, creation and maintenance of necessary datasets, imparting training to the practising engineers and scientists.

SETS may be considered as the secretariat to implement the CybSec4AI mission in addition to jointly implementing the High Quality Node along with the other identified institutes .

9.2 R&D Program

It is proposed to initiate R&D Programme under four schemes with (i) a focus on basic research leading to research papers, (ii) collaborative projects with international institutions leading to joint research publications, prototypes, (iii) product development leading to advanced prototypes and patents and (iv) mission mode projects jointly with R&D labs, industry, and academic research institutions leading to industry competitive tools, systems and solutions.

The priority areas identified for R&D are as follows:

(i) ML for Cyber Security

Use of machine learning for the following applications will be of great interest: malware detection, intrusion detection, deep-fake detection, classification method for software vulnerability based on deep NN, security of machine learning applications, impersonation, inversion, defending machine learning models, safety verification, model explainability, adversarial approach for explainable AI, privacy ensuring techniques in ML such as ML on encrypted data, homomorphic encryption, secure multi-party computation, differential privacy, adversarial ML, poisoning and evasion attacks.

Botnet traffic detection, incident analysis and forecasting, malware analysis, SCADA and mobile communications, fraud detection, privacy preserving systems, situational awareness, insider threat detection, bio-inspired approaches for cyber security and digital forensics.

Model based DRL methods and model free methods for cyber defence are interesting areas to explore.

(ii) ***Attack resistant AI system development***

AI can help defend against cyber-attacks but also facilitate dangerous attacks. Hackers can take advantage of AI to make attackers smarter and more sophisticated to bypass detection methods to penetrate computer systems and networks. Attackers can poison the data pool used for training deep ML methods (i.e., ML poisoning) or attackers can manipulate the states or policies, falsify part of reward signals in Reinforcement Learning to trick the agent into taking sub optimal actions, resulting in the agent being compromised. These kinds of attacks need to be detected, prevented, to develop attack resistant AI systems.

(iii) Use of AI for CPS Security including IoT Security

Wikipedia defines digital twin as a replica of a living or non-living physical entity. In the modern connected world, digital twins integrate IoT, AI/ML, and software analytics, to create living digital models that update as their physical counter parts change. It increases the number of vulnerabilities for software, risk of Intellectual Property theft and exposure of critical processes. A well-defined Secure Software Development Lifecycle (SDLC) management process that includes all aspects of lifecycle, from inception to system retirement is suggested *[Hearn 2019]*.

As recent research efforts indicate that data integrity attacks are able to bypass the data detection mechanism and provide misleading system states to the system operator, thus leading to massive losses in cyber physical systems such as smart grid. Research programmes leading to new architectures and solutions to address emerging AI based cyber threat landscape in Smart CPS including critical infrastructures to be initiated.

(iv) Wireless Security in the context of ML enabled 5G/beyond 5G services

AI/ML technologies are increasingly being applied to network management and cyber security solutions. The growing complexity of SDN/NFV and IMT2020/5G networks and beyond, ML may well be applicable for automatic network orchestration and network management. There is a need to initiate R&D projects to address cyber security challenges of software driven 5G solutions and network management, which are increasingly being designed with ML algorithms.

(v) Securing ML tools

The protection mechanism needed for the ML tools during the training as well as in the operation phase are to be explored and the ML systems need to be made more resilient.

- ML system shall not have bias or prejudice while handling sensitive information to avoid errors.
- An assessment of how an adversary may respond to a particular ML approach
- ML tools have to be aware of the security policy of the organization.
- ML tools have to be continuously maintained by gathering data, labelling it, continuously investigating and learning about the cyber security field.
- Ethical implications shall be taken into account: for example, the ML systems shall not violate right to privacy.
- AI systems shall use more than one algorithm to increase robustness. Different objective functions and different feature vectors need to be used.

(vi) Social Media Attacks and Mitigation

With the unprecedented popularity of social media networks such as Facebook, WhatsApp, Twitter, etc., the network today has become more vulnerable to several new security attacks. As social networks are also being used by enterprises for HR and marketing activities, the attackers find ways to spread malware into corporate networks. Bots manipulate social influence in cyber space. There have been efforts to automate bots using ML algorithms to launch several new attacks on enterprise systems. Research & development on social media attacks and mitigation techniques using AI/ML algorithms need to be initiated.

9.3 Skilling and Training

Development of AI-CS and CS-AI systems in different domains requires significant efforts. Building such systems require well trained and skilled human resources. Chapter 6 of this report brought out the acute need of skilled and trained human resources to work in the cross disciplines of AI and CS.

9.3.1 A framework for training the trainers needs to be worked out focusing on excellence at the foundational level and having in-place amplifying mechanisms similar to T10kT programme offered by IIT Bombay (supported by National Mission on Education through ICT-NMEICT), which permits a large number of teachers to benefit from each of the specialized courses. Use of online and blended approach allows participants to complete a significant part of training online, thus reducing the time and travel costs and can scale to the large number of faculty. They in turn can reach to a larger cross-section of students, bringing the multiplier effect. The main focus would be to work with engineering universities/colleges across the country to enhance the AI and CS skills of faculty in the cross disciplines of AI and cyber security. Training the trainers in this specialized domain of AI and CS could be entrusted to ISEA programme of MeitY with the technical support from High Quality Node (HQN).

9.3.2 Specialised courses on CybSec4AI for practising engineers and scientists need to be organized with national and international experts to jumpstart the development work in building indigenous tools, systems and products. It is important to teach ML with as much data context as possible.

It is suggested that a domain context be chosen which the students understand and apply ML within that context and also making students learn the concepts, algorithms and the necessary maths behind machine learning models to be fully enable them to develop, deploy and optimize the efficacious models.

9.3.3 PhD and M.S Fellowships in the specialized domain of CybSec4AI.

To meet the growing needs of researchers and high-end system developers including modelling experts, it is proposed to institute PhD and M.S Fellowships in the specialized domain of CybSec4AI. About 200 research fellowships are proposed for a period of five years.

9.3.4 Since CybSec4AI is still an emerging research area, it is necessary to periodically organize workshops/symposia/conferences/training programs on AI-CS and CS-AI as knowledge dissemination exercise. International experts can take part in such programs to give exposure on the state-of-the-art technologies to scientists, engineers and faculty.

9.4 International research collaborations

International research collaborations need to be initiated with leaders in chosen areas. Some of the areas identified are differential privacy, adversarial examples and protection mechanisms, and cryptographic techniques in machine learning. This would also help in organizing international conferences/seminars/training programs with the contributions coming from the International experts.

9.5 Creation of National level test bed

CyberSec4AI test bed for carrying out AI based cyber security attack, response and defence experiments which would help in studying, modelling and understanding attack behaviour as well as arriving at tried-and-tested defence mechanisms is very much required. This could be established in the High Quality Nodal Centre (HQN).

9.6 Repository of datasets

Collection of extensive datasets from structured observations, experiments and case studies to train learning models is essential and concerted efforts need to be undertaken at national level. Setting up of a national repository of datasets along the lines of the linguistic data consortium for creating quality data for AI research needs to be considered. Announcement of cyber challenges may be one way to strengthen and bring visibility and recognition for such datasets created under this program.

It is necessary to understand the context in which data is generated as the context would help to build an appropriate model based on chosen scenarios. These datasets are domain specific and each such dataset shall consist of three parts, viz., (a) training dataset, (b) test dataset and (c) scale up dataset.

Data visualization techniques need to be understood and implemented for better training performance - how to extract profitable features of data is also very important.

Datasets will be suitably massaged so that privacy and confidentiality expectations of individuals and organizations are met.

It is also proposed to create a research platform for data collection and analysis for providing assistance to cyber security analysts for evaluating risks and take action plans.

9.7 AI-CS and CS-AI Standards

It is necessary to arrive at standards in this domain to provide interoperability and accessibility. An expert group may be formed to interact with international and national standards bodies such as Institute of Electrical and Electronics Engineers (IEEE), Bureau of Indian Standards (BIS) to arrive at new standards or adapt to the national level standards.

Efforts may be taken to ensure that cyber security aspects are taken into consideration when developing standards for AI/ML as well as evolving AI/ML standards may also be considered while designing AI enabled cyber security systems.

9.8 CyberSec4AI-India Portal

A national portal “CyberSec4AI-India” needs to be launched to create a common platform for sharing ideas, expertise, datasets, etc., to various stakeholders such as researchers, faculty, practitioners and developers.

Necessary travel support need to be provided to the faculty, research students, practicing engineers and scientists in R&D labs for attending and presenting research papers in reputed conferences.

9.9 Duration of the Mission & Budget Requirements

The global proliferation of AI systems and the potential use of AI technology in many sectors including critical infrastructure to improve their efficiency of operation has made it imperative that India as a nation has to scale up its efforts to remain competitive. It has made our nation, to gear itself up in building national level competency in terms of skilled human resources to develop the technology indigenously to meet the growing challenges in the intersection of AI and cyber security. While AI/ML helps in building sophisticated tools and solutions to address expanding cyber threats, it introduces many new cyber security challenges, which are to be addressed by initiating R&D program leading to the development of cyber security technologies in terms of tools, products and solutions, and various related aspects of CyberSec4AI. It is estimated that this program needs GOI support of Rs.635 Crores towards all the components enlisted in the report, spread over a period of 5 years.

10 Conclusion

The digital era empowered by AI is enabling immeasurable innovations in every discipline of science and engineering, and our data and information both at individual and organizational levels are exposed to countless threats.

The proliferation of new IoT devices, the software driving the IoT devices, and the next generation communication networks driven by sophisticated software will exponentially increase the attack surface to a point, where no single human or even a team of humans will be able to respond to the attacks.

Moreover, the increasing deployment of IoT devices in critical infrastructure, the emerging massive Machine to Machine communication (massive M2M) driven by next generation cellular communication systems (5G) and exponential growth of social media will cause diversity of data to explode exponentially making it infeasible for humans to mount timely cyber defences.

Malicious hackers and adversarial nation states will likely take advantage of these developments quickly and develop new types of attacks. Advances in AI will enable AI systems to generate new types of attacks automatically. The combination of new devices/software coupled with power of AI will make level of cyber-threat rise to unimaginable scales.

This report proposes R&D initiatives in "AI for cyber security" to defend against these new threats in an effective and timely manner, and "cyber security for AI" as the defensive-AI will also likely to be attacked by malicious actors. The development of human resources required at national-level to develop technologies in the intersection of AI and cyber security domains is also addressed in this report.

References

[Abaid 2017]
Abaid, Z., Kaafar, M. A., & Jha, S. (2017). Quantifying the impact of adversarial evasion attacks on machine learning based android malware classifiers. In *2017 IEEE 16th International Symposium on Network Computing and Applications (NCA)* (pp. 1-10).

[AI Farms]
Source: https://time.com/5518339/china-ai-farm-artificial-intelligence-Cyber Security/

[AI Policy]
Source: https://futureoflife.org/ai-policy-china/?cn-reloaded=1

[Alkasassbeh 2016]
Alkasassbeh, M., Al-Naymat, G., Hassanat, A., & Almseidin, M. (2016). Detecting distributed denial of service attacks using data mining techniques. *International Journal of Advanced Computer Science and Applications*, *7*(1), 436-445.

[Al-Qurishi 2017]
Al-Qurishi, M., Al-Rakhami, M., Alamri, A., Alrubaian, M., Rahman, S. M. M., & Hossain, M. S. (2017). Sybil defense techniques in online social networks: a survey. *IEEE Access*, *5*, 1200-1219.

[An 2019]
An, D., Yang, Q., Liu, W., & Zhang, Y. (2019). Defending against data integrity attacks in smart grid: A deep reinforcement learning-based approach. *IEEE Access*, *7*, 110835-110845.

[Andy Patel 2019]
Source: https://blog.f-secure.com/adversarial-attacks-against-ai/

[Baars 2008]
Baars, H., & Kemper, H. G. (2008). Management support with structured and unstructured data—an integrated business intelligence framework. *Information Systems Management*, *25*(2), 132-148.

[Barushka 2018]
Barushka, A., & Hájek, P. (2018, May). Spam filtering in social networks using regularized deep neural networks with ensemble learning. In *IFIP International Conference on Artificial Intelligence Applications and Innovations* (pp. 38-49).

[Bahnsen 2018]
Bahnsen, A. C., Torroledo, I., Camacho, L. D., & Villegas, S. (2018). DeepPhish: Simulating Malicious AI. In *2018 APWG Symposium on Electronic Crime Research (eCrime)* (pp. 1-8).

[Beer 2017]
Beer, F., Hofer, T., Karimi, D., & Buhler, U. (2017). A new attack composition for network security. In *10. DFN-Forum Kommunikationstechnologien*. Gesellschaft fürInformatik eV.

[Beigi 2014]
Beigi, E. B., Jazi, H. H., Stakhanova, N., & Ghorbani, A. A. (2014). Towards effective feature selection in machine learning-based botnet detection approaches. In *2014 IEEE Conference on Communications and Network Security* (pp. 247-255).

[BEL]
Source: http://bel-india.in/ContentPage.aspx?MId=32&CId=1413&LId=1&link=1413

[Belinkov 2017]
Belinkov, Y., & Bisk, Y. (2017). Synthetic and natural noise both break neural machine translation. *arXiv preprint arXiv:1711.02173.*

[Berman 2019]
Berman, D. S., Buczak, A. L., Chavis, J. S., & Corbett, C. L. (2019). A survey of deep learning methods for Cyber Security. Information, 10(4), 122.

[Bhatia 2014]
Bhatia, S., Schmidt, D., Mohay, G., & Tickle, A. (2014). A framework for generating realistic traffic for Distributed Denial-of-Service attacks and Flash Events. *Computers & security*, *40*, 95-107.

[Bhattacharya 2014]
Bhattacharya, S., & Selvakumar, S. (2014). Ssenet-2014 dataset: A dataset for detection of multiconnection attacks. In *2014 3rd International Conference on Eco-friendly Computing and Communication Systems* (pp. 121-126).

[Bhuyan 2015]
Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2015). Towards Generating Real-life Datasets for Network Intrusion Detection. *IJ Network Security*, *17*(6), 683-701.

[Biggio 2010]
Biggio, B., Fumera, G., & Roli, F. (2010). Multiple classifier systems for robust classifier design in adversarial environments. *International Journal of Machine Learning and Cybernetics*, *1*(1-4), 27-41.

[Biggio 2014]
Biggio, B., Corona, I., Nelson, B., Rubinstein, B. I., Maiorca, D., Fumera, G., & Roli, F. (2014). Security evaluation of support vector machines in adversarial environments. In *Support Vector Machines Applications* (pp. 105-153). Springer, Cham.

[Biggio2018]
Biggio, B., & Roli, F. (2018). Wild patterns: Ten years after the rise of adversarial machine learning. *Pattern Recognition*, *84*, 317-331.

[Bresniker 2019]
Bresniker, K., Gavrilovska, A., Holt, J., Milojicic, D., Tran, T. (2019). Grand Challenge: Applying Artificial Intelligence and Machine Learning to Cyber Security, *IEEE Computer*, *52*(12), 45-52.

[Brown 2018]
Brown, S. (2017). An analysis of loss-free data aggregation for high data reliability in wireless sensor networks. In *2017 28th Irish Signals and Systems Conference (ISSC)* (pp. 1-6).

[Bruckner 2012]
Bruckner, M., Kanzow, C., & Scheffer, T. (2012). Static prediction games for adversarial learning problems. *Journal of Machine Learning Research*, *13*(Sep), 2617-2654.

[Buczak 2016]
Buczak, A. L., & Guven, E. (2015). A survey of data mining and machine learning methods for Cyber Security intrusion detection. *IEEE Communications surveys & tutorials*, *18*(2), 1153-1176.

[CAIDA]
Source: http://www.caida.org/data/overview

[Carlini 2017]
Carlini, N., & Wagner, D. (2017). Towards Evaluating the Robustness of Neural Networks, IEEE Symposium on Security and Privacy. *IEEE Computer Society*.

[Chen 2017]
Chen, P. Y., Zhang, H., Sharma, Y., Yi, J., & Hsieh, C. J. (2017). Zoo: Zero-th order optimization based black-box attacks to deep neural networks without training substitute models. In *Proceedings of the 10th ACM Workshop on Artificial Intelligence and Security* (pp. 15-26).

[China Initiatives]
Source: https://oecd.ai/dashboards/policy-initiatives?conceptUris=http:%2F%2Fkim.oecd.org%2FTaxonomy%2FGeographicalAreas%23China

[Cordero 2019]
Cordero, C. G., Vasilomanolakis, E., Wainakh, A., Mühlhäuser, M., & Nadjm-Tehrani, S. (2019). On generating network traffic datasets with synthetic attacks for intrusion detection. *arXiv preprint arXiv:1905.00304*.

[Creech 2013]
Creech, G., & Hu, J. (2013). Generation of a new IDS test dataset: Time to retire the KDD collection. In *2013 IEEE Wireless Communications and Networking Conference (WCNC)* (pp. 4487-4492).

[DARPA 2018]
Source: https://www.darpa.mil/work-with-us/ai-next-campaign

[Delvaux 2019]
Delvaux, J. (2019). Machine-Learning Attacks on PolyPUFs, OB-PUFs, RPUFs, LHS-PUFs, and PUF–FSMs. *IEEE Transactions on Information Forensics and Security*, *14*(8), 2043-2058.

[Donegan 2019]
Donegan, P. (2019). AI in Cyber Security: Filtering out the noise, HardenStance White Paper sponsored by Fortinet, Juniper, Kpan and Nokia

[Dong 2017]
Dong, Y., Liao, F., Pang, T., Su, H., Hu, X., Li, J., & Zhu, J. (2017). Boosting adversarial attacks with momentum. arxiv preprint. *arXiv preprint arXiv: 1710.06081*.

[Drias 2015]
Drias, Z., Serhrouchni, A., & Vogel, O. (2015). Taxonomy of attacks on industrial control protocols. In *2015 International Conference on Protocol Engineering (ICPE) and International Conference on New Technologies of Distributed Systems (NTDS)* (pp. 1-6).

[Eykholt 2018]
Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., & Song, D. (2018). Robust physical-world attacks on deep learning visual classification. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1625-1634).

[Fang 2019]
Fang, H., Wang, X., & Tomasin, S. (2019). Machine Learning for Intelligent Authentication in 5G and Beyond Wireless Networks. *IEEE Wireless Communications*, *26*(5), 55-61.

[Fire 2014]
Fire, M., Goldschmidt, R., & Elovici, Y. (2014). Online social networks: threats and solutions. *IEEE Communications Surveys & Tutorials*, *16*(4), 2019-2036.

[Futureskills]
Source: https://economictimes.indiatimes.com/industry/services/education/nasscoms-futureskills-will-train-students-from-next-year/articleshow/66342396.cms?from=mdr

[Fuller 2020]
Aidan Fuller, Zhong Fan, Charles Day and Chris Barlow (2020), "Digital Twin: Enabling Technologies, Challenges and Open Research", IEEE Access (preprint).

[Garcia 2014]
Garcia, S., Grill, M., Stiborek, J., & Zunino, A. (2014). An empirical comparison of botnet detection methods. *Computers & security*, *45*, 100-123.

[Garg 2019]
Garg, S., Kaur, K., Kumar, N., & Rodrigues, J. J. (2019). Hybrid deep-learning-based anomaly detection scheme for suspicious flow detection in SDN: A social multimedia perspective. *IEEE Transactions on Multimedia*, *21*(3), 566-578.

[GAO 20-48G-2020]
Technology Readiness Assessment Guide, GAO US Government Accountability Office, January 2020.

[Gavrilut 2009]
Gavrilut, D., Cimpoesu, M., Anton, D., & Ciortuz, L. (2009). Malware detection using machine learning. In *2009 International Multiconference on Computer Science and Information Technology* (pp. 735-741).

[Gepp 2015]

Gepp, A., & Kumar, K. (2015). Predicting financial distress: a comparison of survival analysis and decision tree techniques. *Procedia Computer Science*, *54*, 396-404.

[Global Sharing]

Source: https://www.nature.com/articles/d41586-019-01681-x

[Gogoi 2012]

Gogoi, P., Bhuyan, M. H., Bhattacharyya, D. K., & Kalita, J. K. (2012). Packet and flow-based network intrusion dataset. In *International Conference on Contemporary Computing* (pp. 322-334). Springer, Berlin, Heidelberg.

[Goodfellow 2014]

Goodfellow, I. J., Shlens, J., & Szegedy, C. (2014). Explaining and harnessing adversarial examples. *arXiv preprint arXiv:1412.6572*.

[Grant Gross 2018]

Source: https://securityintelligence.com/how-seo-poisoning-campaigns-are-mounting-a-comeback/

[Gringoli 2009]

Gringoli, F., Salgarelli, L., Dusi, M., Cascarano, N., Risso, F., & Claffy, K. C. (2009). Gt: Picking up the truth from the ground for internet traffic. *ACM SIGCOMM Computer Communication Review*, *39*(5), 12-18.

[Grosse 2016]

Grosse, K., Papernot, N., Manoharan, P., Backes, M., & McDaniel, P. (2016). Adversarial perturbations against deep neural networks for malware classification. *arXiv preprint arXiv:1606.04435*.

[Guan 2018]

Guan, Z., Bian, L., Shang, T., & Liu, J. (2018). When machine learning meets security issues: A survey. In *2018 IEEE International Conference on Intelligence and Safety for Robotics (ISR)* (pp. 158-165).

[Hall 2017]

Hall, W., & Pesenti, J. (2017). Growing the artificial intelligence industry in the UK. *Department for Digital, Culture, Media & Sport and Department for Business, Energy & Industrial Strategy. Part of the Industrial Strategy UK and the Commonwealth*.

[Hearn 2019]

Hearn, M., Rix, S. (2019).Cyber Security considerations for Digital Twin Implementation, *Industrial Internet Consortium (IIC) Journal of Innovation*.

[Hofstede 2014]

Hofstede, R., Hendriks, L., Sperotto, A., & Pras, A. (2014). SSH compromise detection using NetFlow/IPFIX. *ACM SIGCOMM computer communication review*, *44*(5), 20-26.

[Horizon 2020]
Horizon 2020 work programme 2014-2015, Leadership in enabling and industrial technologies. Source: https://ec.europa.eu/programmes/horizon2020/en/h2020-section/leadership-enabling-and-industrial-technologies

[Huang 2018]
Huang, C. H., Lee, T. H., Chang, L. H., Lin, J. R., & Horng, G. (2018). Adversarial attacks on SDN-based deep learning IDS system. In *International Conference on Mobile and Wireless Technology* (pp. 181-191). Springer, Singapore.

[Husak 2018]
Husak, M., Komarkova, J., Bou-Harb, E., & Celeda, P. (2018). Survey of attack projection, prediction, and forecasting in Cyber Security. *IEEE Communications Surveys & Tutorials*, *21*(1), 640-660.

[IEEE Confluence 2017]
The result of an intensive three-day IEEE Confluence 6-8 October (2017). Artificial Intelligence and Machine Learning Applied to Cyber Security
Source: https://www.ieee.org/about/industry/confluence/feedback.html

[Jazi 2017]
Jazi, H. H., Gonzalez, H., Stakhanova, N., & Ghorbani, A. A. (2017). Detecting HTTP-based application layer DoS attacks on web servers in the presence of sampling. *Computer Networks*, *121*, 25-36.

[Kantarcioglu 2011]
Kantarcioglu, M., Xi, B., & Clifton, C. (2011). Classifier evaluation and attribute selection against active adversaries. *Data Mining and Knowledge Discovery*, *22*(1-2), 291-335.

[Karnataka Government]
Source: https://bahrainedb.com/latest-news/bahrain-and-karnataka-sign-memorandum-of-understanding-promoting-cooperation-in-fintech-ai-iot-and-cyber-security/

[Katz 2017]
Katz, G., Barrett, C., Dill, D. L., Julian, K., & Kochenderfer, M. J. (2017). Towards proving the adversarial robustness of deep neural networks. *arXiv preprint arXiv:1709.02802*.

[Kaloudi 2020]
Kaloudi, N., & Li, J. (2020). The AI-based cyber threat landscape: A survey. *ACM Computing Surveys (CSUR)*, *53*(1), 1-34.

[Katzir 2018]
Katzir, Z., & Elovici, Y. (2018). Quantifying the resilience of machine learning classifiers used for Cyber Security. *Expert Systems with Applications*, *92*, 419-429.

[Kent 2015]
Kent, A. D. (2015). *Multi-Source Cyber-Security Events, Los Alamos National Laboratory*.

[Kent 2016]
Kent, A. D. (2016). Cyber Security data sources for dynamic network research. In *Dynamic Networks and Cyber-Security* (pp. 37-65).

[Kolias 2015]
Kolias, C., Kambourakis, G., Stavrou, A., & Gritzalis, S. (2015). Intrusion detection in 802.11 networks: empirical evaluation of threats and a public dataset. *IEEE Communications Surveys & Tutorials*, *18*(1), 184-208.

[Krotov 2017]
Krotov, D., & Hopfield, J. (2018). Dense associative memory is robust to adversarial inputs. *Neural computation*, *30*(12), 3151-3167.

[Kubovic 2019]
Kubovic, O., (2019). Machine-Learning era in Cyber Security: A step towards a safer world or the brink of chaos? ESET ASIA PTE LTD, Singapore.

[Kurakin 2016]
Kurakin, A., Goodfellow, I., & Bengio, S. (2016). Adversarial machine learning at scale. *arXiv preprint arXiv:1611.01236*.

[Liang 2017]
Liang, B., Li, H., Su, M., Bian, P., Li, X., & Shi, W. (2017). Deep text classification can be fooled. *arXiv preprint arXiv:1704.08006*.

[Lippmann 2000]
Lippmann, R. P., Fried, D. J., Graf, I., Haines, J. W., Kendall, K. R., McClung, D., Weber, D., Webster, S.E., Wyschogrod, D., Cunningham, R.K. & Zissman, M. A. (2000). Evaluating intrusion detection systems: The 1998 DARPA off-line intrusion detection evaluation. In *Proceedings DARPA Information Survivability Conference and Exposition. DISCEX'00* (Vol. 2, pp. 12-26).

[Loaiza 2019]
Loaiza, F.L., Birdwell, J.D., Kennedy, G.L., & Visser, D. (2019). Utility of Artificial Intelligence and Machine Learning in Cyber Security. IDA Non-Standard NS D-10694.

[Madni 2019]
Madni, A. M., Madni, C. C., & Lucero, S. D. (2019). Leveraging digital twin technology in model-based systems engineering. *Systems*, *7*(1), 7.

[Mahoney 2003]
Mahoney, M. V., & Chan, P. K. (2003). An analysis of the 1999 DARPA/Lincoln Laboratory evaluation data for network anomaly detection. In *International Workshop on Recent Advances in Intrusion Detection* (pp. 220-237). Springer, Berlin, Heidelberg.

[Malware Classifiers]
Abaid, Z., Kaafar, M. A., & Jha, S. (2017). Quantifying the impact of adversarial evasion attacks on machine learning based android malware classifiers. In 2017 IEEE 16th International Symposium on Network Computing and Applications (NCA) (pp. 1-10).

[Martinez 2019]
Martinez, D. R., et al (2019). Artificial Intelligence: Short History, Present Developments, Future Outlook. Project Report, MIT Lincoln Laboratory Report

[MeitY 2019]
Cyber Security, Safety, Legal and Ethical Issues, Ministry of Electronics and Information Technology (MeitY), Government of India.
Source: https://meity.gov.in/artificial-intelligence-committees-reports

[McGuire 2019]
McGuire M., (2019). Social Media Platforms and the Cyber Economy, Into The Web of Profit, University of Surrey, Sponsored by Bromium.

[Melis 2017]
Melis, M., Demontis, A., Biggio, B., Brown, G., Fumera, G., & Roli, F. (2017). Is deep learning safe for robot vision? adversarial examples against the icub humanoid. In *Proceedings of the IEEE International Conference on Computer Vision Workshops* (pp. 751-759).

[Miller 2019]
Miller, D. J., Xiang, Z., & Kesidis, G. (2019). Adversarial learning in statistical classification: A comprehensive review of defenses against attacks. *arXiv preprint arXiv:1904.06292*.

[MOCI]
The Artificial Intelligence Task Force, Ministry of Commerce and Industry, Government of India.
Source: https://dipp.gov.in/whats-new/report-task-force-artificial-intelligence

[Morris 2014]
Morris, T., & Gao, W. (2014). Industrial control system traffic datasets for intrusion detection research. In *International Conference on Critical Infrastructure Protection* (pp. 65-78). Springer, Berlin, Heidelberg.

[Moustafa 2015]
Moustafa, N., & Slay, J. (2015). UNSW-NB15: a comprehensive dataset for network intrusion detection systems (UNSW-NB15 network dataset). In *2015 military communications and information systems conference (MilCIS)* (pp. 1-6).

[NASSCOM 2018]
Artificial Intelligence Primer 2018, NASSCOM's Annual Bigdata and Analytics Summit, 2018

[Newman 2019]
Newman, J.C. (2019). Towards AI Security, Global Aspirations for a more resilient future, Centre for Long-Term Cyber Security, Centre for Long-term Cyber Security, UC Berkeley. cltc.berkeley.edu

[Nguyen 2019]
Nguyen, T. T., & Reddi, V. J. (2019). Deep Reinforcement Learning for Cyber Security. arXiv preprint arXiv:1906.05799.

[NITI Aayog 2018]
National Strategy for Artificial Intelligence #AIFORALL, Discussion Paper, June 2018, NITI Aayog, Government of India.

[NSF 2019]
Source: https://www.nsf.gov/pubs/2019/nsf19603/nsf19603.htm

[NSF 2015]
Source: https://www.nsf.gov/pubs/2015/nsf15541/nsf15541.htm

[Pang 2005]
Pang, R., Allman, M., Bennett, M., Lee, J., Paxson, V., & Tierney, B. (2005). A first look at modern enterprise traffic. In *Proceedings of the 5th ACM SIGCOMM conference on Internet Measurement*.

[Papernot 2016]
Papernot, N., McDaniel, P., Jha, S., Fredrikson, M., Celik, Z. B., & Swami, A. (2016). The limitations of deep learning in adversarial settings. *IEEE European symposium on security and privacy* (pp. 372-387).

[Papernot 2017]
Papernot, N., McDaniel, P., Goodfellow, I., Jha, S., Celik, Z. B., & Swami, A. (2017). Practical black-box attacks against machine learning. In *Proceedings of the 2017 ACM on Asia conference on computer and communications security* (pp. 506-519).

[Qian 2006]
Qian, J., Xu, C., & Shi, M. (2006). Redesign and implementation of evaluation dataset for intrusion detection system. In *International Conference on Emerging Trends in Information and Communication Security* (pp. 451-465).

[Ring 2017]
Ring, M., Wunderlich, S., Grüdl, D., Landes, D., & Hotho, A. (2017). Creation of flow-based datasets for intrusion detection. *Journal of Information Warfare*, *16*(4), 41-54.

[Ring 2019]
Ring, M., Wunderlich, S., Scheuring, D., Landes, D., & Hotho, A. (2019). A survey of network-based intrusion detection datasets. *Computers & Security*.

[RIPE NCC]
Source: https://labs.ripe.net/datarepository/data-sets

[Rodofile 2017]
Rodofile, N. R., Radke, K., & Foo, E. (2017). Framework for SCADA cyber-attack dataset creation. In *Proceedings of the Australasian Computer Science Week Multiconference* (pp. 1-10).

[Rodriguez 2020]
Ariel Rodriguez and Koji Okamura (2020), Social Media Data Mining for Proactive Cyber Defense, Journal of Information Processing, Vol.28 pp 230-238. March 2020, Information Processing Society of Japan.

[Saad 2011]
Saad, S., Traore, I., Ghorbani, A., Sayed, B., Zhao, D., Lu, W., Felix, J. & Hakimian, P. (2011). Detecting P2P botnets through network behavior analysis and machine learning. In *2011 Ninth annual international conference on privacy, security and trust* (pp. 174-180).

[Sangster 2009]
Sangster, B., O'Connor, T. J., Cook, T., Fanelli, R., Dean, E., Morrell, C., & Conti, G. J. (2009). Toward Instrumenting Network Warfare Competitions to Generate Labelled Datasets. In *CSET*.

[Santanna 2015]
Santanna, J. J., van Rijswijk-Deij, R., Hofstede, R., Sperotto, A., Wierbosch, M., Granville, L. Z., & Pras, A. (2015). Booters - an analysis of DDoS-as-a-service attacks. In *2015 IFIP/IEEE International Symposium on Integrated Network Management (IM)* (pp. 243-251).

[Santikellur 2019]
Santikellur, P., Bhattacharyay, A., & Chakraborty, R. S. (2019). *Deep Learning Based Model Building Attacks on Arbiter PUF Compositions* (p.10). Cryptology ePrint Archive, Report 2019/566. 2019.

[Seymour 2016]
Seymour, J., & Tully, P. (2016). Weaponizing data science for social engineering: Automated E2E spear phishing on Twitter. *Black Hat USA*, *37*, 1-39.

[Seymour 2018]
Seymour, J., & Tully, P. (2018). Generative Models for Spear Phishing Posts on Social Media. *arXiv preprint arXiv:1802.05196*.

[Sharif 2016]
Sharif, M., Bhagavatula, S., Bauer, L., & Reiter, M. K. (2016). Accessorize to a crime: Real and stealthy attacks on state-of-the-art face recognition. In *Proceedings of the 2016 acmsigsac conference on computer and communications security* (pp. 1528-1540).

[Shiravi 2012]
Shiravi, A., Shiravi, H., Tavallaee, M., & Ghorbani, A. A. (2012). Toward developing a systematic approach to generate benchmark datasets for intrusion detection. *Computers & security*, *31*(3), 357-374.

[Singh 2015]
Singh, R., Kumar, H., & Singla, R. K. (2015). A reference dataset for network traffic activity based intrusion detection system. *International Journal of Computers Communications & Control*, *10*(3), 390-402.

[Sitawarin 2018]
Sitawarin, C., Bhagoji, A. N., Mosenia, A., Chiang, M., & Mittal, P. (2018). Darts: Deceiving autonomous cars with toxic signs. *arXiv preprint arXiv:1802.06430*.

[Song 2011]
Song, J., Takakura, H., Okabe, Y., Eto, M., Inoue, D., & Nakao, K. (2011). Statistical analysis of honeypot data and building of Kyoto 2006+ dataset for NIDS evaluation. In *Proceedings of the first workshop on building analysis datasets and gathering experience returns for security* (pp. 29-36).

[Sperotto 2009]
Sperotto, A., Sadre, R., Van Vliet, F., & Pras, A. (2009). A labeled dataset for flow-based intrusion detection. In *International Workshop on IP Operations and Management* (pp. 39-50). Springer, Berlin, Heidelberg.

[Spring 2019]
Spring, J. M., Fallon, J., Galyardt, A., Horneman, A., Metcalf, L., & Stoner, E. (2019). Machine Learning in Cyber Security: A GUIDE,Carneige Mellon University, Software Engineering institute, Technical report CMU/SEI-2019-TR-005, http://www.sei.cmu.edu

[Stieglitz 2018]
Stieglitz, S., Mirbabaie, M., Ross, B., & Neuberger, C. (2018). Social media analytics–Challenges in topic discovery, data collection, and data preparation. *International journal of information management*, *39*, 156-168.

[Tavallaee 2009]
Tavallaee, M., Bagheri, E., Lu, W., & Ghorbani, A. A. (2009). A detailed analysis of the KDD CUP 99 dataset. In *2009 IEEE symposium on computational intelligence for security and defense applications* (pp. 1-6).

[Ucci 2019]
Ucci, D., Aniello, L., & Baldoni, R. (2019). Survey of machine learning techniques for malware analysis. *Computers & Security*, *81*, 123-147.

[UMassTraceRepository]
Source: http://traces.cs.umass.edu/index.php/Network/Network

[Union Budget 2020]
Source: https://www.theweek.in/news/biz-tech/2020/02/01/it-sector-cheers-union-budget-2020-for-focusing-on-ai-digitisation.html

[Vectra 2019]
VECTRA White Paper, (2019). How to augment Security Operation Centres with artificial intelligence, vectra.ai

[Veeramachaneni 2016]
Veeramachaneni, K., Arnaldo, I., Korrapati, V., Bassias, C., & Li, K. (2016). AI^ 2: training a big data machine to defend. In *2016 IEEE 2nd International Conference on Big Data Security on Cloud (Big Data Security), IEEE International Conference on High Performance and Smart Computing (HPSC), and IEEE International Conference on Intelligent Data and Security (IDS)* (pp. 49-54).

[Viegas 2017]
Viegas, E. K., Santin, A. O., & Oliveira, L. S. (2017). Toward a reliable anomaly-based intrusion detection in real-world environments. *Computer Networks*, *127*, 200-216.

[Vilela 2014]
Vilela, D. W., Ed'Wilson, T. F., Shinoda, A. A., de Souza Araújo, N. V., de Oliveira, R., & Nascimento, V. E. (2014). A dataset for evaluating intrusion detection systems in IEEE 802.11 wireless networks. In *2014 IEEE Colombian Conference on Communications and Computing (COLCOM)* (pp. 1-5).

[Wang 2020]
Wang, C. X., Di Renzo, M., Stanczak, S., Wang, S., & Larsson, E. G. (2020). Artificial Intelligence Enabled Wireless Networking for 5G and Beyond: Recent Advances and Future Challenges. *IEEE Wireless Communications*, *27*(1), 16-23.

[Wei 2018]
Wei, X., Liang, S., Chen, N., & Cao, X. (2018). Transferable adversarial attacks for image and video object detection. *arXiv preprint arXiv:1811.12641*.

[Wikipedia]
Source: https://en.wikipedia.org/wiki/Adversarial_machine_learning

[Xiao 2018]
Xiao, C., Deng, R., Li, B., Yu, F., Liu, M., & Song, D. (2018). Characterizing adversarial examples based on spatial consistency information for semantic segmentation. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 217-234).

[Yan 2019]
Yan, W., Mestha, L. K., & Abbaszadeh, M. (2019). Attack Detection for Securing Cyber Physical Systems. *IEEE Internet of Things Journal*, *6*(5), 8471-8481.

[Yener 2019]
Yener, B., & Gal, T. (2019). Cyber Security in the Era of Data Science: Examining New Adversarial Models. *IEEE Security & Privacy*.

[Zhou 2018]
Zhou, Z., Tang, D., Wang, X., Han, W., Liu, X., & Zhang, K. (2018). Invisible mask: Practical attacks on face recognition with infrared. *arXiv preprint arXiv:1803.04683*.

[Zuech 2015]
Zuech, R., Khoshgoftaar, T. M., Seliya, N., Najafabadi, M. M., & Kemp, C. (2015). A new intrusion detection benchmarking system. In *The Twenty-Eighth International Flairs Conference*.

Annexure I: Delivery model for Training the trainers- T10kT

The detailed information about the T10kT program is provided below, which is taken from the IIT Bombay homepage: https://www.it.iitb.ac.in/nmeict/home.html

The T10kT project (supported by National Mission on Education through ICT) at IIT Bombay, permits thousands of teachers to benefit from each of these programs. Use of online and blended approach allows participants to complete a significant part of training online, thus reducing the time which must be spent on face-to face synchronous interaction.

A Teaching Learning Centre (TLC) has been setup at IIT Bombay under Pandit Madan Mohan Malaviya National Mission on Teachers and Teaching. Our Faculty Development Programs (FDPs), approved by AICTE are now conducted under the aegis of this scheme.

About T10kT

The 'Train 1000 Teachers' program was initiated by IIT Bombay in 2009, under the project 'Empowerment of Students/Teachers,' sponsored by the National Mission on Education through ICT (MHRD, Government of India). The main focus of this program is to work with Engineering Colleges in the country to enhance the teaching skills of faculty in core Engineering and Science subjects. This project attempts to address a critical subset of important issues and adopts an approach to address these utilizing modern technologies. It uses an ICT enabled process involving both synchronous and asynchronous mode to actually reach out and engage a large number of teachers, and through them, a much larger number of students. Indeed, actual involvement of these important stakeholders in the entire process, scaled to very large numbers using ICT, can be said to be a major contribution of the project.

In 2013, this program was scaled further, to train up to 10,000 teachers at a time, using 385 established Remote Centres across India. IIT Kharagpur is our partner Institution in the second phase. Together, we have a mandate to train 1, 50,000 teachers over the next three years.

The 'Train 10,000 Teachers' (T10kT) uses the AVIEW framework developed by Amrita University, which provides an interactive social environment through the internet, for e-learning, and communicating and interacting with people at different places. It is used to deliver and transmit live lectures given by faculty at IIT, to all RCs.

Methodology

We identify engineering institutes across the country, which could act as Remote Centres (RCs). The criteria used are proximity to a number of other colleges, infrastructure, good faculty, and number of engineering disciplines.

Once the topic of the workshop is decided, we invite expert faculty from these Centres to a five-day Coordinators' training workshop held at IIT Bombay, two months before the main workshop. The training workshop is conducted by IIT faculty members, who are subject experts in the topic of the main workshop. This workshop is conducted to finalize the syllabus of the main workshop, and to train the workshop coordinators in the methodology used by the IIT faculty.

Each Remote Centre hosts the main two-week ISTE workshop with partial funding from IIT Bombay. The workshop is attended by interested participants of the nearby colleges. The lectures of the IIT subject experts are transmitted live from IIT Bombay/IIT Kharagpur during the workshop. The participants attend lectures in the morning and participate in assignments/tutorials in the afternoon. The Workshop Coordinators supervise the conduct of tutorials and Labs at their Centres. They liaise between the participants at their Remote Centres, and IIT Bombay/IIT Kharagpur. The participants of the workshop receive certificates from IIT Bombay/IIT Kharagpur and ISTE, after satisfying all the stipulated requirements. All the lectures and tutorial sessions are also recorded for future use.

The final edited audio-visual content, along with other course material, is released in Open Source to be freely used later by all teachers, students, and other learners.

Figures

- Total Workshops Conducted: 69
- Total Coordinator Workshops Conducted: 26
- Total Main Workshops Conducted: 25

For example,

- Cyber security (URL:https://www.it.iitb.ac.in/nmeict/workshopDetails.html?workshopid=o58Z2itp91VLPLmH1suCtA&category=ryjCPtOBugIe3qrCMGqopQ),
- Computer networking (URL:https://www.it.iitb.ac.in/nmeict/workshopDetails.html?workshopid=GydhXjgE7sJZw8Dd5uuMRw&category=ryjCPtOBugIe3qrCMGqopQ),
- DBMS (URL:https://www.it.iitb.ac.in/nmeict/workshopDetails.html?workshopid=kOfYToT7yaxssTHH1To5Xw&category=ryjCPtOBugIe3qrCMGqopQ)

- Total self sustained main workshops conducted: 18

Annexure II: IPR model for collaborative R&D

Australian Cyber Security Cooperative Research Centre (CSCRC) could be an interesting case-study in general, and for our context in particular. The detailed information taken from the URL: https://www.cybersecuritycrc.org.au/ is provided below.

Mission Statement

"To be an independent and collaborative centre where industry, government and research partners work together to create new products, services and systems that deliver a secure and resilient national cyber security capability, and enhance cyber expertise for the nation, making Australia a safer place to do business."

About

The CSCRC, with more than 20 participants, operates through a public, not-for-profit company (Cyber Security Research Centre Ltd). The CSCRC is focused on delivering industry-driven cyber security research outcomes that have impact and address real-world cyber security problems with innovative solutions.

The CSCRC has been granted $50M of funding over 7 years from the Australian Government's Cooperative Research Centres program.

The CSCRC connects its participants that are from industry, government as well as leading cyber security research organizations (including six universities) to develop and potentially commercialize products and services that improve the cyber security posture of Australia. CSCRC will also play an important role in raising cyber security awareness more broadly and exploring key cyber security issues with respect to the legal, policy and regulatory implications of Cyber Security risks.

One key CSCRC aim is to attract, inspire, mentor and develop the next generation of cyber security professionals by offering the best and brightest student's scholarships through our participating universities. Students will enjoy the added advantage of also working within the CSCRC network that includes industry participants.

IPR Model

Here are some relevant aspects of the CSCRC IPR model copied from clauses 23 and 24 of its participants agreement:

Centre IP

Ownership of Project IP

The Company will be the owner of the legal interest in the project IP. The beneficial interest in the project IP will be owned by: (i) those project participants who have made contributions to the project; and (ii) the company, in proportion to that part of the project that is funded from common wealth funding or other funding made available to the project by the company.

Variations to Project Shares

Each project party's project share will be recalculated by the company in accordance with the following formula as at 1st July of each Financial Year and will apply to the following Financial Year:

$$PS = PC/TPC \times 100\%$$

Where,

- PS is the Project Share for a qualifying project party, expressed as a percentage of all Project Shares;
- TPC is the cumulative total of all qualifying project parties' project contributions made (not project contributions owing nor future project contributions promised) up to the most recent 30 June; and
- PC means the cumulative Project Contributions of a qualifying project party made (not project contributions owing nor future project contributions promised) up to the most recent 30 June.

Use of Centre IP

A participant wishing to use centre IP for any purpose, whether for research or otherwise, not authorized under clause 23.14 must obtain a license from the company covering such use. The grant of any such license will be at the company's absolute discretion and must include provisions governing ownership and utilization of further intellectual property developed from such use. If the proposed use of centre IP relates to the activities then the company will not unreasonably withhold its consent to such use.

Registration in Company name

For the avoidance of doubt, if patenting or other forms of intellectual property protection of centre IP is pursued, such registrations to be in the company's name and at the company's cost (unless agreed otherwise); and where required, will identify

- The inventors of the centre IP
- The beneficial owners of the centre IP

Costs of protection

The centre account may be drawn upon by the company to meet all costs associated with applying for, maintaining and prosecuting patent or any other form of intellectual property protection associated with centre IP (including any action for infringement of the centre IP) and the application, maintenance and prosecution of any actions which may be associated with any such intellectual property and such drawings will be taken to be expenses incurred in the performance of the activities.

Project Participants' right to use Project IP

Subject to this agreement, each project participant has a non-exclusive royalty-free right to use the project IP (excluding the right to sublicense):

- for the purpose of undertaking a project in accordance with this agreement; or
- for internal research, education and teaching purposes other than commercialization; or in the case of defence
- for any other purpose, other than commercialization, within the project participant's respective

Utilization of Centre IP

A. Right and responsibility of company

The company has the exclusive right to utilize the centre IP at its discretion (including a right to sublicense)

B. Commercialization Expenses

Without limiting clause 27.8, the participants agree that the company may use the commercialization income with respect to particular project IP to pay commercialization expenses or reimburse the company for previously incurred commercialization expenses with respect to that project IP. For the avoidance of doubt, the participants acknowledge that the company may in any financial year use commercialization income received that financial year to reimburse commercialization Expenses with respect to that project IP incurred in previous financial years. Any commercialization expenses relating to various project IP may be apportioned between the various projects IP by the company (in accordance with project shares for that project IP) for the purposes of this clause 24.9.

C. Payment of Net Commercialization Income

By 31 August each year, the company must remit to each project participant its share of net commercialization income with respect to their beneficial interest in the respective project IP for the previous Financial Year corresponding to project shares as at 1 July that year, subject to any adjustments pursuant to clauses 24.3 or 24.4. The company will retain its project share for reinvestment into its activities. The company will provide information to the participants to support the calculation and payment of net commercialization revenue.

D. Survival

This clause 24 survives expiration or earlier termination of this agreement.

Additional comments:

All IP generated in the research is owned by the CRC. Each project participant has a non-exclusive royalty-free right to use the project IP for non-commercial purposes to create proof of concepts and provide demos of such project IP to its clients (or prospective clients) either alone or in combination with other participant or third-party products, solutions, services or offerings. Multiple parties involved in a particular project, need to act in good faith towards project outcome shares.

This seems like a simple, reasonable IP arrangement that is worth looking into.

Annexure III: Research Components & Tools for CS4AI

(Responses from Institutions, Organizations & Industry)

A questionnaire on CybSec4AI was circulated among few key institutions and industry in order to obtain a view of the research components, datasets and other tools being used by them in their cyber security and AI research.

''

The questionnaire sent to these institutes is presented below:

Questionnaire on CybSec4AI

- Datasets for research:
  1. **Describe datasets being used for research**:
  2. **Are they Proprietary or Open source?**
  3. **Whether willing to share these datasets?**
- Tools for design and development AI/ML systems:
  1. **The tools used for development of AI/ML models and systems**
  2. **Have you developed any toolset for design and development of AI/ML systems? If yes, please provide details of in-house developed tools.**
  3. **Have you established any specialized infrastructure?**
- Ongoing research: **Provide details of ongoing research in the areas of AI for CS;**
- Product Development and Systems Development: **Provide details of the products/system developments based on AI for CS applications.**
- Support requirement for PhD fellowships: **Specify the number of PhD Fellowship requirements.**
- Intensive training: **What are the courses being offered or proposed for providing training in managing and operating cyber security systems**
- Travel support for international conferences: **Indicate the travel requirements per year by the identified members working in the AI-CS domain.**
- Possible R&D areas proposed: **Provide details on the proposed R&D areas using AI-CS technology.**
- Industries/Start up engagements: **Name the start-ups and industries you are engaged with for carrying out research/activities in AI-CS domain.**
- Existing international collaborations: **Please specify details w.r.t the existing international collaborations is the AI-CS area.**
- Project funding: **Provide the estimate requirements of project funding per year for the AI-CS related activities.**

''

The responses received from various premier academic institutes, R&D labs, and industry for the above questionnaire are examined, consolidated and summarised in the following in the areas of datasets used for research, tools used for design and development, product development & systems developments under taken, PhD fellowship support, intensive training, travel support, industry engagements, international collaborations etc.

Datasets used for Research

Academia, organisations and industry confirmed that publicly available and custom datasets are being used for CybSec for AI research. There are institutions who also took advantage of datasets taken from Wikipedia, Twitter data available with TREC conference calls, NIST USA, NIH GEO gene expression datasets, Amazon customer review data, MICCAI BraTS & RSNA, Kaggle, UCI Machine Learning repository, MNIST handwritten dataset, LDC Database, Pascal VOC, RightWhale, image forgery and image manipulation datasets etc. Healthcare datasets were also collected in collaboration with hospitals and also from their online sources. Efforts were also made to collect agricultural datasets. Organisations collect datasets from various sources of databases. Intrusion detection evaluation datasets like CICIDS2017, CIDDS-001, 2017-SUEE-data-set, Wassa-2017, video image retrieval and analysis tool video dataset, labelled datasets developed comprising of classified data like URL analyser and classifier database, suspicious and obfuscated URLs, botnet traffic, binary image, malware hashes, port scan etc., BTC-USD dataset from Yahoo finance industries have been used. Freely available datasets as well as those obtained the task specific datasets from customers under non-disclosure agreements are also used.

Most of the Institutions, organisations and industries generally use open source data for their CS & AI research. However, few institutions have created their own custom data for their own applications. Few have also depended on proprietary datasets. While most of them were willing to share these datasets, few have also expressed that owing to security hazards, the same could be shared only with explicit sample sharing agreement while few industries also informed that due to legal considerations, it would be difficult to share proprietary datasets.

Tools for design and development AI/ML systems

The institutions had indicated that they have used tools like Python, R, RStudio, C++ , Java, MATLAB, LabelImg, ROS (Robot Operating System), PyTorch, TensorFlow, Jetson board, Hadoop, Spark, Eclipse, NS2, OPnet, Neural Network tools, Scilab, AI-based open source software for data analytics etc., for the design and development of AI systems. Tools were also being developed based on generative adversarial learning. Few of them are in the process of development and others have already set up high-performance computing centres, developed frameworks, ML algorithms for several applications, tools like Matlabfsolve, interpolation, CUDA Development Toolkit, Microsoft Visual Studio, NVIDIA Visual Profiler, few other open source tools etc. They have established GPU based servers for ML/DL applications and other infrastructures like EEG wireless recording system, high configuration machine for large computations, IOT security labs, scientific computing labs and soft computing labs, heterogeneous computing facility etc.

For the development of AI/ML models and systems, organisations used tools such as Python, Dot Net ecosystem, Weka, Scikit for Machine Learning etc. They have developed open source tools like Python and other tools like CUDA Development Toolkit, Microsoft Visual Studio, and NVIDIA Visual Profiler etc., for design and development of AI/ML systems. Infrastructure such as big data scalable framework is established under NCCC Project for large scale data storage.

As for industries, development frameworks such as TensorFlow, Keras, Weka, Pytorch, and OpenCV along with several python libraries were majorly used. The infrastructure created was in the form of internal cloud which can be used across the companies and caters both GPU and CPU workloads. They also had some local data storage and processing.

Open Research

The institutions and organisations are exploring open research in AI, machine learning, deep learning, bio-engineering and cyber security areas using various applications and techniques, adversarial machine learning, malware analysis, applied statistics, AI/ML-based techniques for automatically detecting vulnerabilities in software applications, AI-based aerial/terrestrial traffic sensing using LiDAR point cloud processing, high throughput crop phenotyping using UAV based sensor like hyperspectral, multispectral and RGB camera, IoT enabled AI based guided and automated diagnostic system, Brain-Controlled IoT environments (BCE), algorithm for enhancing Cyber Security in IoT infrastructure, AI-based channel access mechanisms and adaptive communication strategies, secure distributed computing-secure Multi-Party Computation (MPC), cognitive computing, Natural Language Processing (NLP), predictive analysis, packet processing, flow-based analysis, modelling decision trees, behavioural based analysis, context aware detection, DIGIFAI toolset etc.

The institutions and the R&D labs have expressed the above as key areas for possible R&D research in the AI-CS domain.

Product Development and Systems Development

Different tools are being developed by institutions and industries using ML based techniques to develop effective system over current applications.

Suraksha Vyuh - A real-time video analytics product, Jigyasa - an offline video analytics product, Drishti - quality and compliance product, automatic number plate recognition are few such tools developed by IIT Bombay. IIT Indore is working on a high scale international project entitled 'Digital Forensic Knowledge Integration and Intelligence (DIREKT-Intel)' under MHRD SPARC scheme for designing and developing algorithms for aiding cyber forensic investigations. IDS, IPS & firewall system, Security Monitoring Unit (SMU) of C-DAC Bangalore and cyber threat monitoring/management system of C-DAC Mohali are few worth mentioning developments using AI-CS applications.

Research backed indigenous products are developed by TCS, especially in privacy space; viz., TCS MasterCraft Data Masker®, TCS MasterCraft Dynamic Data Masker®, TCS MasterCraft Volume Data Generator®, TCS Crystal Ball™ etc., K7 Labs uses ML actively to classify user-browsed URLs into respective categories.

PhD fellowships support

The responses received from questionnaire circulated in about 14 institutions and organisations indicates a requirement average of around 5 PhD Fellowships per year, i.e. a maximum of 12 and minimum of 1 no. of them, to support their AI-CS technology ventures.

Intensive Trainings

The faculties from various institutions in the country have offered to deliver the courses in the following AI-CS areas:

- Deep learning/Machine learning
- Probability and statistics
- Cryptography
- Cyber Physical Systems (CPS)
- Network security/IOT security/hi-performance computing and hardware security
- Security and data communications
- Data science for cyber security
- Image and text forgery
- Obscene content detection in social media
- AI in cyber security & forensics/hands-on cyber security & cyber forensics/document forensics
- Biometrics
- Application layer DOS attacks

Travel support for international conferences

The responses for travel support requirements during a year as specified in the questionnaire by around 14 Institutions and organisations indicates an average budget of Rs.12 Lakhs per year for about 5 members per year. This helped in arriving at the national level requirements.

Industries/ Start-up engagements

The Industries/ Start up engagements of various Indian institutions and R&D labs as specified in the questionnaire includes Ernst and Young (E&Y) and Payatu Technologies, Express Analytics, SKIOT start-up, HoneyPot sensors deployed at various geo-locations covering different sectors/organizations, Anavadya Softech, TCS Co-Innovation Network (COIN™) etc.

International collaborations

The questionnaire indicates responses from the Indian institutions and organizations w.r.t their existing international collaborations in the AI-CS field and it includes many reputed and recognized Universities and key firms worldwide. The universities are Fordham University, New York University, Duke University, University of Florida, University of Tokyo, Ritsumeikan University, Nanyang Technological University, University of Technology Sydney, Edge Hill University, University of Tasmania, Norwegian University of Science and Technology, Hong Kong Polytechnique University, German National Research Centre for Information Technology, University of Kent, University of Reading, Trent University, Purdue University, University of Bristol, University of Newcastle, University of Lisbon, Portugal, University of the Peloponnese, Yale Institute of Network Science.

Few key Firms are Australian Cyber Security Cooperative Research Centre, Blavatnik Interdisciplinary Cyber Research Centre, Cyber@ Ben-Gurion University of the Negev and Thales Research.

Project funding

The project funding requirements section of the questionnaire fetched responses from around 14 Indian institutions and organizations which indicated a budgetary average of Rs.7 crores estimate and an average duration of 4 years for the project and it ranged from Rs.50 lakhs minimum to a maximum of Rs.10 crores and the minimum duration being 1.5 years and maximum 8 years.

Task Force is extremely grateful and acknowledges the co-operation in providing responses to the questionnaires, by the **Institutions**: IIT Bhilai, IIT Goa, IIT Hyderabad (WiNet Lab), IIT Indore, Machine Intelligence Unit, Indian Statistical Institute-Kolkata, IIIT Hyderabad, DR SPM International Institute of Information Technology (IIIT)-Naya Raipur, IIIT Bangalore, NIT Calicut, NIT Meghalaya, NIT Puducherry and NIT Surathkal, **Organisations**: C-DAC Bangalore, C-DAC Kolkata, C-DAC Mohali, C-DAC Mumbai and **Industries:** TATA Consultancy Services Ltd/ TCS Research/ Cyber Security and Privacy, K7 Labs (research labs at K7 Computing).

Annexure IV: Model for PhD Fellowships – Privacy & Us - EU

"Privacy & Us" is an interesting consortium model in EU that seems to be working well for properly defined research topic. Here is information about the same taken from its URL: https://privacyus.eu/

With the rapid accumulation and processing of personal data by numerous organizations, it is of paramount importance to protect people from adverse uses of their data, while allowing them to enjoy the benefits the use of these data can possibly provide. This is the question of protecting citizens' privacy, while enabling them to make informed decisions regarding their actions with privacy implications.

The Privacy & Us Innovative Training Network (ITN) will train thirteen creative, entrepreneurial and innovative early stage researchers (ESRs) to be able to reason, design and develop novel solutions to questions related to the protection of citizens' privacy, considering the multidisciplinary and inter-sectoral aspects of the issue. ESRs will be trained to face both current and future challenges in the area of privacy and usability. Privacy & Us offers a combination of research-related and transferable competence skills that will enhance the career perspectives of the ESRs in both the academic and nonacademic sectors.

The ESRs will receive comprehensive training and engage in inter-sectoral and multidisciplinary collaboration. Through this collaborative effort, the project will make a significant contribution and impact to the ESRs future careers. It will also contribute to shaping future privacy policies and practices in Europe and will significantly advance the state of the art in privacy and usability research.

This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Sklodowska-Curie grant agreement No 675730, within the Marie Sklodowska-Curie Innovative Training Networks (ITN-ETN) framework with a total funding of EUR 3,376,517.

This consortium consists of five universities and four non-academic organizations. This deliberate and balanced consortium ensures that the project outcomes and impacts are maximal with respect to the EU policy regarding data security and privacy. The market potential for the generated knowledge is also a key element of Privacy & Us. Our industrial partners are two SMEs, one large enterprise and a data protection authority.

Annexure V: Start-ups using AI to drive Cyber Security

This Annexure includes a list of start-ups working in AI-CS application domains:

- ***Antivirus Solutions***: Start-up- Matisoft
- ***Security Operation Centre:*** Lucideus Technologies
- ***Vulnerability Assessment and Penetration Testing (VAPT):*** Start-ups- Kratikal Tech, Security Brigade
- ***Threat Hunting and Threat Intelligence:*** Start-up- Threatsys
- ***Managed Detection and Response (MDR)***: Paladion Networks, Sequretek
- ***Authentication and Access Management:*** Primeauth, InCights, Block Armour, Sectona are functioning in this domain.
- ***Application Security:*** HaltDos
- ***Surveillance:*** Uncanny vision & DeepSight AI labs (GoDeep)
- ***Cloud-based cyber security solutions that use AI:*** Instasafe, CloudSEK & Seconize
- ***Identity Management***: Uniken & Uniphoreb are the start-ups in Identity Management, which is one of the earlier domains to adopt Artificial Intelligence.

Annexure VI: Representative list of Publications by researchers in Indian Institutions

Research papers on AI for cyber security collected from Citeseer, IEEE Xplore digital library and ACM Digital Library (The works were collected by advanced search methodology with search term: Cyber Security in All Metadata, Artificial Intelligence in All Metadata in IEEE Xplore digital library and with Search term: Anywhere – "Cyber Security" AND Anywhere – "Artificial Intelligence" OR Anywhere- "deep learning" OR Anywhere- "machine learning" in ACM Digital Library (2017-2020)), received through questionnaire on CybSec4AI, google search, and through visiting websites of various institutions are listed below. This list provided is not exhaustive and may not include all the research papers published from India in this domain.

1. Nath, K., Roy, S., Nandi, S. (2020). An Incremental Approach for Hierarchical Community Mining in Evolving Social Graphs. International Journal of Intelligent Enterprise, 7(1).
2. Aggarwal, P., Dutt, V. (2020). Role of information about opponent's actions and intrusion-detection alerts on cyber-decisions in Cyber Security games. Cyber Security: A Peer-Reviewed Journal 4, in press.
3. Maqbool, Z., Aggarwal, P., Pammi, V.S.C., Dutt, V. (2020). Cyber Security: Effects of Penalizing Defenders in Cyber-Security Games via Experimentation and Computational Modeling. Frontiers in Psychology 11
4. Z. Yue, S. Ding, L. Zhao, Y. Zhang, Z. Cao, M. Tanveer, A. Jolfaei, X.Zheng, Privacy-preserving time series medical images analysis using a hybrid deep learning frame-work, *ACM Transactions on Internet Technology (TOIT),*2021.
5. Saha, S., Alam, M., Bag, A., Mukhopadhyay, D., & Dasgupta, P. (2020). Leakage Assessment in Fault Attacks: A Deep Learning Perspective. *IACR Cryptol. ePrint Arch., 2020,*
6. Kumar, N., & Awate, S. P. (2020). Semi-Supervised Robust Mixture Models in RKHS for Abnormality Detection in Medical Images. *IEEE Transactions on Image Processing*, *29*, 4772-4787.
7. Koti, N., Pancholi, M., Patra, A., & Suresh, A. (2020). SWIFT: Super-fast and Robust Privacy-Preserving Machine Learning. *arXiv preprint arXiv:2005.10296*.
8. Singh, R., Agarwal, A., Singh, M., Nagpal, S., & Vatsa, M. (2020). On the robustness of face recognition algorithms against attacks and bias. *arXiv preprint arXiv:2002.02942*.
9. Goel, A., Agarwal, A., Vatsa, M., Singh, R., & Ratha, N. K. (2020). DND Net: Reconfiguring CNN for Adversarial Robustness. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 22-23).
10. Sen, S., Ravindran, B., & Raghunathan, A. (2020). Empir: Ensembles of mixed precision deep networks for increased robustness against adversarial attacks. *arXiv preprint arXiv:2004.10162*.
11. Jindal, S., Sood, R., Singh, R., Vatsa, M., & Chakraborty, T. (2020). NewsBag: A Benchmark Multimodal Dataset for Fake News Detection. In *SafeAI@ AAAI* (pp. 138-145).
12. Agarwal, A., Vatsa, M., Singh, R., & Ratha, N. K. (2020). Noise is Inside Me! Generating Adversarial Perturbations with Noise Derived from Natural Filters. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 774-775).

13. Naskar, D., Singh, S. R., Kumar, D., Nandi, S., & Rivaherrera, E. O. D. L. (2020). Emotion Dynamics of Public Opinions on Twitter. *ACM Transactions on Information Systems (TOIS)*, *38*(2), 1-24.
14. Singh, L. G., Anil, A., & Singh, S. R. (2020). SHE: Sentiment Hashtag Embedding Through Multitask Learning. *IEEE Transactions on Computational Social Systems*, *7*(2), 417-424.
15. Shrivastava, P., Jamal, M. S., & Kataoka, K. (2020). EvilScout: Detection and mitigation of evil twin attack in SDN enabled WiFi. *IEEE Transactions on Network and Service Management*, *17*(1), 89-102.
16. Sengupta, J., Ruj, S., & Bit, S. D. (2020). A Comprehensive survey on attacks, security issues and blockchain solutions for IoT and IIoT. *Journal of Network and Computer Applications*, *149*, 102481.
17. Nath, K., Roy, S., & Nandi, S. (2020). InOvIn: A fuzzy-rough approach for detecting overlapping communities with intrinsic structures in evolving networks. *Applied Soft Computing*, *89*, 106096.
18. Agrawal, P., Sharma, T., & Verma, N. K. (2020). Supervised approach for object identification using speeded up robust features. *International Journal of Advanced Intelligence Paradigms*, *15*(2), 165-182.
19. Kumar, A., Gupta, M., Kumar, G., Handa, A., Kumar, N., Shukla, S.K. (2020). A Review: Malware Analysis Work at IIT Kanpur. In: Shukla S., Agrawal M. (eds) Cyber Security in India. IITK Directions, vol 4. Springer, Singapore.
20. Singh, A., Handa, A., Kumar, N., Shukla, S.K. (2020). Malware Analysis Using Image Classification Techniques. In: Shukla S., Agrawal M. (eds) Cyber Security in India.IITK Directions, vol 4. Springer, Singapore.
21. Negi, R., Dutta, A., Handa, A., Ayyangar, U., Shukla, S.K. (2020). Intrusion detection & prevention in Programmable Logic Controllers: A Model-driven Approach. IEEE International Conference on Industrial Cyber-Physical Systems (ICPS 2020), Finland.
22. Kumar, N., Singh, A., Handa, A., Shukla, S. K (2020). Detecting Malicious Accounts on the Ethereum Blockchain with Supervised Learning. Accepted for the 4th International Symposium on Cyber Security Cryptology and Machine Learning (CSCML 2020), Ben Gurion University, Be'er Sheva, Israel.
23. Devi, S. S., Singh, N. H., & Laskar, R. H. (2020). Fuzzy C-Means Clustering with Histogram based Cluster Selection for Skin Lesion Segmentation using Non-Dermoscopic Images. *International Journal of Interactive Multimedia & Artificial Intelligence*, *6*(1).
24. Rawat, R., Singh, B., Sur, A., & Mitra, P. (2020). Steganalysis for clustering modification directions steganography. *Multimedia Tools and Applications*, *79*(3), 1971-1986.

2019

1. Bakker, C., Bhattacharya, A., Chatterjee, S., & Vrabie, D. L. (2019). Learning and Information Manipulation: Repeated Hypergames for Cyber-Physical Security. *IEEE Control Systems Letters*, *4*(2), 295-300.
2. Tariang, D. B., Chakraborty, R. S., & Naskar, R. (2019). A Robust Residual Dense Neural Network for Countering Antiforensic Attack on Median Filtered Images. *IEEE Signal Processing Letters*, *26*(8), 1132-1136.

3. Santikellur, P., Bhattacharyay, A., & Chakraborty, R. S. (2019). Deep Learning based Model Building Attacks on Arbiter PUF Compositions. *IACR Cryptol. ePrint Arch.*, *2019*, 566.
4. Byali, M., Chaudhari, H., Patra, A., & Suresh, A. (2020). FLASH: fast and robust framework for privacy-preserving machine learning. *Proceedings on Privacy Enhancing Technologies*, *2020*(2), 459-480.
5. Gautam, C., Balaji, R., Sudharsan, K., Tiwari, A., & Ahuja, K. (2019). Localized multiple kernel learning for anomaly detection: one-class classification. *Knowledge-Based Systems*, *165*, 241-252.
6. Maqbool, Z., Banerjee, A., Pammi, V.S.C., Dutt, V. (2019). Behavioral Cyber Security: Investigating the influence of patching vulnerabilities on cyber decision-making via cognitive modeling in IJCSA 4(1).
7. Mukherjee, S., Asnani, H., Lin, E., & Kannan, S. (2019). Clustergan: Latent space clustering in generative adversarial networks. In Proceedings of the AAAI Conference on Artificial Intelligence (Vol. 33, pp. 4610-4617).
8. Uma, E., Sirija, M.,Mehala, E. (2019). Feature based Spam Detection Framework to Identify Fake Reviews in Online Social Media. Journal of Computational Information Systems, published by Binary Information Press. Vol. 15, Issue 1, pp. 162-175 .
9. Patel, O. P., Bharill, N., Tiwari, A., Patel, V., Gupta, O., Cao, J., Li, J. & Prasad, M. (2019). Advanced Quantum Based Neural Network Classifier and Its Application for Objectionable Web Content Filtering. *IEEE Access*, *7*, 98069-98082.
10. Dwivedi, R., & Dey, S. (2019). A novel hybrid score level and decision level fusion scheme for cancelable multi-biometric verification. *Applied Intelligence*, *49*(3), 1016-1035.
11. Vijaykeerthy, D., Suri, A., Mehta, S., & Kumaraguru, P. (2019). Hardening deep neural networks via adversarial model cascades. In *2019 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-8).
12. Agarwal, A., Sehwag, A., Vatsa, M., & Singh, R. (2019). Deceiving the protector: Fooling face presentation attack detection algorithms. In *2019 International Conference on Biometrics (ICB)* (pp. 1-6).
13. Goswami, G., Agarwal, A., Ratha, N., Singh, R., & Vatsa, M. (2019). Detecting and mitigating adversarial perturbations for robust face recognition. *International Journal of Computer Vision*, *127*(6-7), 719-742.
14. Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in Cyber Security: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4), e1306.
15. Singhal, S., Shah, R. R., Chakraborty, T., Kumaraguru, P., & Satoh, S. I. (2019). SpotFake: A Multi-modal Framework for Fake News Detection. In *2019 IEEE Fifth International Conference on Multimedia Big Data (BigMM)* (pp. 39-47).
16. Kaur, S., Kumar, P., & Kumaraguru, P. (2019). Automating fake news detection system using multi-level voting model. *Soft Computing*, 1-21.
17. Mukherjee, S., Asnani, H., & Kannan, S. (2019). CCMI: Classifier based conditional mutual information estimation. *arXiv preprint arXiv:1906.01824*.

18. Kanimozhi, V., & Jacob, T. P. (2019). Artificial Intelligence based Network Intrusion Detection with hyper-parameter optimization tuning on the realistic cyber dataset CSE-CIC-IDS2018 using cloud computing. In *2019 International Conference on Communication and Signal Processing (ICCSP)* (pp. 0033-0036).
19. Sagar, B. S., Niranjan, S., Kashyap, N., & Sachin, D. N. (2019). Providing Cyber Security using Artificial Intelligence–A survey. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 717-720).
20. Goyal, Y., & Sharma, A. (2019). A Semantic Machine Learning Approach for Cyber Security Monitoring. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 439-442).
21. Vijayakumar, S., Gowtham, K. S., Nigam, N., & Singh, R. V. R. (2019). An Novel Approach in Designing a Security Workbench with Deep Learning Capabilities and Process Automation. In *TENCON 2019-2019 IEEE Region 10 Conference (TENCON)* (pp. 263-268).
22. Kumar, M. S., Ben-Othman, J., Srinivasagan, K. G., & Krishnan, G. U. (2019). Artificial Intelligence Managed Network Defense System against Port Scanning Outbreaks. In *2019 International Conference on Vision Towards Emerging Trends in Communication and Networking (ViTECoN)* (pp. 1-5).
23. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., & Venkatraman, S. (2019). Robust intelligent malware detection using deep learning. *IEEE Access*, *7*, 46717-46738.
24. Soni, S., & Bhushan, B. (2019). Use of Machine Learning algorithms for designing efficient Cyber Security solutions. In *2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT)* (Vol. 1, pp. 1496-1501).
25. Vinayakumar, R., Alazab, M., Soman, K. P., Poornachandran, P., Al-Nemrat, A., & Venkatraman, S. (2019). Deep learning approach for intelligent intrusion detection system. *IEEE Access*, *7*, 41525-41550.
26. Akarsh, S., Sriram, S., Poornachandran, P., Menon, V. K., & Soman, K. P. (2019). Deep Learning Framework for Domain Generation Algorithms Prediction Using Long Short-term Memory. In *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)* (pp. 666-671).
27. Vijayanand, R., Devaraj, D., & Kannapiran, B. (2019). A Novel Deep Learning Based Intrusion Detection System for Smart Meter Communication Network. In *2019 IEEE International Conference on Intelligent Techniques in Control, Optimization and Signal Processing (INCOS)* (pp. 1-3).
28. Vinayakumar, R., Alazab, M., Jolfaei, A., Soman, K. P., & Poornachandran, P. (2019). Ransomware triage using deep learning: twitter as a case study. In *2019 Cyber Security and Cyberforensics Conference (CCC)* (pp. 67-73).
29. Nagisetty, A., & Gupta, G. P. (2019). Framework for Detection of Malicious Activities in IoT Networks using Keras Deep Learning Library. In *2019 3rd International Conference on Computing Methodologies and Communication (ICCMC)* (pp. 633-637).
30. Sharma, M., Elmiligi, H., Gebali, F., & Verma, A. (2019). Simulating Attacks for RPL and Generating Multi-class Dataset for Supervised Machine Learning. In *2019 IEEE 10th Annual Information Technology, Electronics and Mobile Communication Conference (IEMCON)* (pp. 0020-0026).

31. Roseline, S. A., Sasisri, A. D., Geetha, S., & Balasubramanian, C. (2019). Towards Efficient Malware Detection and Classification using Multilayered Random Forest Ensemble Technique. In *2019 International Carnahan Conference on Security Technology (ICCST)* (pp. 1-6).
32. Mathai, K. J. (2019). Performance Comparison of Intrusion Detection System between Deep Belief Network (DBN) Algorithm and State Preserving Extreme Learning Machine (SPELM) Algorithm. In *2019 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)* (pp. 1-7).
33. Akarsh, S., Simran, K., Poornachandran, P., Menon, V. K., & Soman, K. P. (2019). Deep Learning Framework and Visualization for Malware Classification. In *2019 5th International Conference on Advanced Computing & Communication Systems (ICACCS)* (pp. 1059-1063).
34. Oak, R., Du, M., Yan, D., Takawale, H., & Amit, I. (2019). Malware Detection on Highly Imbalanced Data through Sequence Modeling. In *Proceedings of the 12th ACM Workshop on Artificial Intelligence and Security* (pp. 37-48).
35. Alam, M., & Mukhopadhyay, D. (2019). How secure are deep learning algorithms from side-channel based reverse engineering?. In *Proceedings of the 56th Annual Design Automation Conference 2019* (pp. 1-2).
36. Ravikiran, M., & Madgula, K. (2019). Fusing Deep Quick Response Code Representations Improves Malware Text Classification. In *Proceedings of the ACM Workshop on Crossmodal Learning and Application* (pp. 11-18).
37. Pandya, D. (2019). Spam Detection Using Clustering-Based SVM. In *Proceedings of the 2019 2nd International Conference on Machine Learning and Machine Intelligence* (pp. 12-15).
38. Maheshwari, A., Goyal, A., Hanawal, M. K., & Ramakrishnan, G. (2019).DynGAN: Generative Adversarial Networks for Dynamic Network Embedding.33rd Conference on Neural Information Processing Systems (NeurIPS 2019), Vancouver, Canada.
39. Jami, S. K., Chalamala, S. R., & Jindal, A. K. (2019). Biometric Template Protection Through Adversarial Learning. In *2019 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-6).
40. Singh, A. K., Mishra, A., Shekhar, S., & Chakraborty, A. (2019). From Strings to Things: Knowledge-enabled VQA Model that can Read and Reason. In *Proceedings of the IEEE International Conference on Computer Vision* (pp. 4602-4612).
41. Jindal, A. K., Chalamala, S. R., & Jami, S. K. (2019). Securing Face Templates using Deep Convolutional Neural Network and Random Projection. In *2019 IEEE International Conference on Consumer Electronics (ICCE)* (pp. 1-6).
42. Rosni, K. V., Shukla, M., Banahatti, V., & Lodha, S. (2019). Consent Recommender System: A Case Study on LinkedIn Settings. In the Proceedings of the Privacy-Enhancing Artificial Intelligence and Language Technologies (PAL), Palo Alto, USA, pp. 53-60.
43. Tupsamudre, H., Singh, A. K., & Lodha, S. (2019). Everything Is in the Name–A URL Based Approach for Phishing Detection. In *International Symposium on Cyber Security Cryptography and Machine Learning* (pp. 231-248). Springer, Cham.

44. Pradeep, R., Reddy, M. K., & Rao, K. S. (2019). LSTM-Based Robust Voicing Decision Applied to DNN-Based Speech Synthesis. *Automatic Control and Computer Sciences*, *53*(4), 328-332.
45. Singh, M., Singh, R., Vatsa, M., Ratha, N. K., & Chellappa, R. (2019). Recognizing disguised faces in the wild. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *1*(2), 97-108.
46. Gupta, V., Wadbude, R., Natarajan, N., Karnick, H., Jain, P., & Rai, P. (2019). Distributional Semantics Meets Multi-Label Learning. In *Proceedings of the AAAI Conference on Artificial Intelligence* (Vol. 33, pp. 3747-3754).
47. Bharill, N., Patel, O. P., Tiwari, A., Mu, L., Li, D. L., Mohanty, M., Kaiwartya, O., & Prasad, M. (2019). A generalized enhanced quantum fuzzy approach for efficient data clustering. *IEEE Access*, *7*, 50347-50361.
48. Basavaraju, S., Gaj, S., & Sur, A. (2019). Object memorability prediction using deep learning: location and size bias. *Journal of Visual Communication and Image Representation*, *59*, 117-127.
49. George, C. P., Xia, W., & Michailidis, G. (2019). Analyses of multi-collection corpora via compound topic modeling. In *International Conference on Machine Learning, Optimization, and Data Science* (pp. 205-218). Springer, Cham.
50. Saluja, R., Maheshwari, A., Ramakrishnan, G., Chaudhuri, P., & Carman, M. (2019). OCR On-the-Go: Robust End-to-end Systems for Reading License Plates & Street Signs. In *15th IAPR International Conference on Document Analysis and Recognition (ICDAR)*.
51. Maheshwari, A., Goyal, A., Kumar, A., Hanawal, M. K., & Ramakrishnan, G. (2019). Representation Learning on Graphs by Integrating Content and Structure Information. In *2019 11th International Conference on Communication Systems & Networks (COMSNETS)* (pp. 88-94).
52. Jagadish, B., Mishra, P. K., Kiran, M. P. R. S., & Rajalakshmi, P. (2019). A real-time health 4.0 framework with novel feature extraction and classification for brain-controlled iot-enabled environments. *Neural Computation*, *31*(10), 1915-1944.
53. Tanveer, M., Tiwari, A., Choudhary, R., & Jalan, S. (2019). Sparse pinball twin support vector machines. *Applied Soft Computing*, *78*, 164-175.
54. Dalal, M., Tanveer, M., & Pachori, R. B. (2019). Automated identification system for focal EEG signals using fractal dimension of FAWT-based sub-bands signals. In *Machine Intelligence and Signal Analysis* (pp. 583-596). Springer, Singapore.
55. Law, A., Evani, B., & Ghosh, A. (2019). Optimized Functional Link Artificial Neural Network for Multi-label Classification. *Australian Journal of Intelligent Information Processing Systems, Special Issue: Neural Information Processing, 26th International Conference on Neural Information Processing (ICONIP 2019), Sydney, Australia*, vol.16, issue 4, pp. 56-63, 2019.
56. Subudhi, B. N., Veerakumar, T., Esakkirajan, S., & Ghosh, A. (2019). Kernelized Fuzzy Modal Variation for Local Change Detection From Video Scenes. *IEEE Transactions on Multimedia*, *22*(4), 912-920.
57. Law, A., & Ghosh, A. (2019). Multi-label classification using a cascade of stacked autoencoder and extreme learning machines. *Neurocomputing*, *358*, 222-234.

58. Subudhi, B. N., Rout, D. K., & Ghosh, A. (2019). Big data analytics for video surveillance. *Multimedia Tools and Applications*, *78*(18), 26129-26162.
59. Chakraborty, D., Narayanan, V., & Ghosh, A. (2019). Integration of deep feature extraction and ensemble learning for outlier detection. *Pattern Recognition*, *89*, 161-171.
60. Dauda, K. A., Pradhan, B., Shankar, B. U., & Mitra, S. (2019). Decision tree for modeling survival data with competing risks. *Biocybernetics and Biomedical Engineering*, *39*(3), 697-708.
61. Diwan, C., Srinivasa, S., & Ram, P. (2019). Automatic Generation of Coherent Learning Pathways for Open Educational Resources. In *European Conference on Technology Enhanced Learning* (pp. 321-334). Springer, Cham.
62. Jain, A., Mishra, A., Shukla, A., & Tiwari, R. (2019). A novel genetically optimized convolutional neural network for traffic sign recognition: A new benchmark on Belgium and Chinese traffic sign datasets. *Neural Processing Letters*, *50*(3), 3019-3043.
63. Thokchom, S., & Saikia, D. K. (2019). Privacy Preserving and Public Auditable Integrity Checking on Dynamic Cloud Data. *IJ Network Security*, *21*(2), 221-229.
64. *Editors:* Nandi, S., Jinwala, D., Singh, V., Laxmi, V., Gaur, M.S., Faruki, P. (2019). Security and Privacy (ISBN 978-981-13-7561-3). The Proceedings of Second ISEA International Conference, ISEA-ISAP 2018, *Publishers: Springer Nature Singapore Pte Ltd. Vol. CCIS-939, 2019.*
65. Trivedi, T., Parihar, V., Khatua, M., & Mehtre, B. M. (2019). Threat Intelligence Analysis of Onion Websites Using Sublinks and Keywords. In Emerging Technologies in Data Mining and Information Security (pp. 567-578). Springer, Singapore.
66. Soni, B., Das, P. K., & Thounaojam, D. M. (2019). Geometric transformation invariant block based copy-move forgery detection using fast and efficient hybrid local features. *Journal of information security and applications*, *45*, 44-51.
67. Bakshi, P., & Nandi, S. (2019). Using Privacy Enhancing and Fine-Grained Access Controlled eKYC to implement Privacy Aware eSign.Advances in Science, Technology and Engineering Systems Journal, Volume 4, Issue 4, Page No 347-358.
68. Gupta, V., Agarwal, M., Arora, M., Chakraborty, T., Singh, R., & Vatsa, M. (2019). Bag-of-lies: A multimodal dataset for deception detection. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*.
69. Wazid, M., Das, A. K., Rodrigues, J. J., Shetty, S., & Park, Y. (2019). IoMT malware detection approaches: Analysis and research challenges. *IEEE Access*, *7*, 182459-182476.
70. Wazid, M., Zeadally, S., & Das, A. K. (2019). Mobile banking: evolution and threats: malware threats and security solutions. *IEEE Consumer Electronics Magazine*, *8*(2), 56-60.
71. Srinivas, J., Das, A. K., & Kumar, N. (2019). Government regulations in Cyber Security: Framework, standards and recommendations. *Future Generation Computer Systems*, *92*, 178-188.
72. Fadadu, F., Handa, A., Kumar, N., & Shukla, S. K. (2019). Evading API Call Sequence Based Malware Classifiers. In *International Conference on Information and Communications Security* (pp. 18-33). Springer, Cham.

73. Bhagwani, H., Negi, R., Dutta, A. K., Handa, A., Kumar, N., & Shukla, S. K. (2019). Automated classification of web-application attacks for intrusion detection. In International Conference on Security, Privacy, and Applied Cryptography Engineering (pp. 123-141). Springer, Cham.
74. Kumar, N., Mukhopadhyay, S., Gupta, M., Handa, A., & Shukla, S. K. (2019). Malware Classification using Early Stage Behavioral Analysis. In *2019 14th Asia Joint Conference on Information Security (AsiaJCIS)* (pp. 16-23).
75. Kumar, A., Kumar, N., Handa, A., & Shukla, S. K. (2019). PeerClear: Peer-to-Peer Bot-net Detection. In *International Symposium on Cyber Security Cryptography and Machine Learning* (pp. 279-295). Springer, Cham.
76. Singh, A., Handa, A., Kumar, N., & Shukla, S. K. (2019). Malware classification using image representation. In *International Symposium on Cyber Security Cryptography and Machine Learning* (pp. 75-92). Springer, Cham.
77. Handa, A., Sharma, A., & Shukla, S. K. (2019). Machine learning in Cyber Security: A review. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(4), e1306.
78. Patel, O. P., Tiwari, A., Chaudhary, R., Nuthalapati, S. V., Bharill, N., Prasad, M., & Hussain, O. K. (2019). Enhanced quantum-based neural network learning and its application to signature verification. *Soft Computing*, *23*(9), 3067-3080.

2018

1. Ulybyshev, D., Palacios, S., Mani, G., Alsalem, A. O., Bhargava, B., & Goyal, P. (2018). On-the-fly Analytics over Encrypted Records in Untrusted V2X Environments.
2. Tripathi, N., & Hubballi, N. (2018). Detecting stealth DHCP starvation attack using machine learning approach. *Journal of Computer Virology and Hacking Techniques*, *14*(3), 233-244.
3. Sohony, I., Pratap, R., & Nambiar, U. (2018). Ensemble learning for credit card fraud detection. In *Proceedings of the ACM India Joint International Conference on Data Science and Management of Data* (pp. 289-294).
4. SMaqbool, Z., Pammi, V. C., & Dutt, V. (2018). Cyber Security: Influence of patching vulnerabilities on the decision-making of hackers and analysts. International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA 2018), At Glasgow UK.
5. Aggarwal, P., Moisan, F., Gonzalez, C., & Dutt, V. (2018). Understanding Cyber Situational Awareness in a Cyber Security Game Involving Recommendations. International Journal of Cyber Situational Awareness. 3(1), 1-29.
6. Chhabra, S., Singh, R., Vatsa, M., & Gupta, G. (2018). Anonymizing k-facial attributes via adversarial perturbations. *arXiv preprint arXiv:1805.09380*.
7. Goel, A., Singh, A., Agarwal, A., Vatsa, M., & Singh, R. (2018). Smartbox: Benchmarking adversarial detection and mitigation algorithms for face recognition. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)* (pp. 1-7).

8. Agarwal, A., Singh, R., Vatsa, M., & Ratha, N. (2018). Are image-agnostic universal adversarial perturbations for face recognition difficult to detect?. In *2018 IEEE 9th International Conference on Biometrics Theory, Applications and Systems (BTAS)* (pp. 1-7).
9. Aggarwal, A., Viswanath, B., Zhang, L., Kumar, S., Shah, A., & Kumaraguru, P. (2018). I spy with my little eye: Analysis and detection of spying browser extensions. In *2018 IEEE European Symposium on Security and Privacy (EuroS&P)* (pp. 47-61).
10. Mishra, P., Varadharajan, V., Tupakula, U., & Pilli, E. S. (2018). A detailed investigation and analysis of using machine learning techniques for intrusion detection. *IEEE Communications Surveys & Tutorials*, *21*(1), 686-728.
11. Verma, K., & Jain, N. (2018). IoT Object Authentication for Cyber Security: Securing Internet with Artificial intelligence. In *2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)* (pp. 1-3).
12. Vigneswaran, K. R., Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2018). Evaluating shallow and deep neural networks for network intrusion detection systems in Cyber Security. In *2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-6).
13. Jayaprakash, S., & Kandasamy, K. (2018). Database Intrusion Detection System Using Octraplet and Machine Learning. In *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT)* (pp. 1413-1416).
14. Shibija, K., & Joseph, R. V. (2018). A Machine Learning Approach to the Detection and Analysis of Android Malicious Apps. In *2018 International Conference on Computer Communication and Informatics (ICCCI)* (pp. 1-4). IEEE.
15. Ranjan, S., Garhwal, P., Bhan, A., Arora, M., & Mehra, A. (2018). Framework for Image Forgery Detection and Classification Using Machine Learning. In *2018 2nd International Conference on Trends in Electronics and Informatics (ICOEI)* (pp. 1-9).
16. Brahanyaa, S., & Anbarasi, L. J. (2018). Classification of SNMP Network Dataset for DDoS attack prevention. In *2018 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)* (pp. 1-5).
17. Subramanyam, D. (2018). Classification of Intrusion Detection Dataset using machine learning Approaches. In *2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS)* (pp. 280-283).
18. Amma, N. G. B., & Subramanian, S. (2018). VCDeepFL: vector convolutional deep feature learning approach for identification of known and unknown denial of service attacks. In *TENCON 2018-2018 IEEE Region 10 Conference* (pp. 0640-0645).
19. Sewak, M., Sahay, S. K., & Rathore, H. (2018). An investigation of a deep learning based malware detection system. In *Proceedings of the 13th International Conference on Availability, Reliability and Security* (pp. 1-5).
20. Kumar Jindal, A., Chalamala, S., & Kumar Jami, S. (2018). Face template protection using deep convolutional neural network. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (pp. 462-470).
21. Mishra, A., & Singh, A. K. (2018). Deep embedding using Bayesian risk minimization with application to sketch recognition. In *Asian Conference on Computer Vision* (pp. 357-370). Springer, Cham.

22. Panchal, G., & Samanta, D. (2018). A novel approach to fingerprint biometric-based cryptographic key generation and its applications to storage security. Computers & Electrical Engineering, 69, 461-478.
23. Aggarwal, P., Gonzalez, C., Dutt, V. (2018). HackIt: A Real-Time Simulation Tool for Studying Real-World Cyber-Attacks in the Laboratory. Handbook of Computer Networks and Cyber Security: Principles and Paradigms Link
24. Choudhury, A., Kaushik, S., & Dutt, V. (2018). Social-network analysis in healthcare: analyzing the effect of weighted influence in physician networks. Network Modeling Analysis in Health Informatics and Bioinformatics, 7(1), 17.
25. Revathi, A. R., & Kumar, D. (2018). Hybridisation of feed forward neural network and self-adaptive PSO with diverse of features for anomaly detection. *International Journal of Biomedical Engineering and Technology*, *26*(2), 111-140.
26. Bhilare, S., Kanhangad, V., & Chaudhari, N. (2018). A study on vulnerability and presentation attack detection in palmprint verification system. *Pattern Analysis and Applications*, *21*(3), 769-782.
27. Ghosh, K., Neogy, S., Das, P. K., & Mehta, M. (2018). Intrusion detection at international borders and large military barracks with multi-sink wireless sensor networks: an energy efficient solution. *Wireless Personal Communications*, *98*(1), 1083-1101.

2017

1. Murthy, K. R., & Ghosh, A. (2017). Norm discriminant eigenspace transform for pattern classification. *IEEE Transactions on Cybernetics*, *49*(1), 273-286.
2. Singh, S. K., Khanna, K., Bose, R., Panigrahi, B. K., & Joshi, A. (2017). Joint-transformation-based detection of false data injection attacks in smart grid. *IEEE Transactions on Industrial Informatics*, *14*(1), 89-97.
3. Khatua, M., Safavi, S. H., & Cheung, N. M. (2017). Detection of internet traffic anomalies using sparse laplacian component analysis. In GLOBECOM 2017-2017 IEEE Global Communications Conference (pp. 1-6).
4. Aggarwal, P., Gonzalez, C., & Dutt, V. (2017). Modeling the effects of amount and timing of deception in simulated network scenarios. In *2017 International Conference on Cyber Situational Awareness, Data Analytics and Assessment (Cyber SA)* (pp. 1-7).
5. Alam, M., Bhattacharya, S., Mukhopadhyay, D., & Bhattacharya, S. (2017). Performance Counters to Rescue: A Machine Learning based safeguard against Micro-architectural Side-Channel-Attacks. *IACR Cryptol. ePrint Arch.*, *2017*, 564.
6. Manjani, I., Tariyal, S., Vatsa, M., Singh, R., & Majumdar, A. (2017). Detecting silicone mask-based presentation attack via deep dictionary learning. *IEEE Transactions on Information Forensics and Security*, *12*(7), 1713-1723.
7. Khanna, K., Panigrahi, B. K., & Joshi, A. (2017). AI-based approach to identify compromised meters in data integrity attacks on smart grid. *IET Generation, Transmission & Distribution*, *12*(5), 1052-1066.
8. Maniath, S., Ashok, A., Poornachandran, P., Sujadevi, V. G., Sankar, A. P., & Jan, S. (2017). Deep learning LSTM based ransomware detection. *Recent Developments in Control, Automation & Power Engineering (RDCAPE)* (pp. 442-446).

9. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Evaluating effectiveness of shallow and deep networks to intrusion detection system. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1282-1289).
10. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Deep android malware detection and classification. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 1677-1683).
11. Vinayakumar, R., Soman, K. P., Velan, K. S., & Ganorkar, S. (2017). Evaluating shallow and deep networks for ransomware detection and classification. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 259-265).
12. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Secure shell (ssh) traffic analysis with flow based features using shallow and deep networks. In *2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2026-2032).
13. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Evaluating shallow and deep networks for secure shell (ssh) traffic analysis. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 266-274).
14. Vinayakumar, R., Soman, K. P., & Poornachandran, P. (2017). Applying deep learning approaches for network traffic prediction. *International Conference on Advances in Computing, Communications and Informatics (ICACCI)* (pp. 2353-2358).
15. Sethi, K., Chaudhary, S. K., Tripathy, B. K., & Bera, P. (2017). A novel malware analysis for malware detection and classification using machine learning algorithms. In *Proceedings of the 10th International Conference on Security of Information and Networks* (pp. 107-113).
16. Barnett, A., Santokhi, J., Simpson, M., Smart, N. P., Stainton-Bygrave, C., Vivek, S., & Waller, A. (2017). Image Classification using non-linear Support Vector Machines on Encrypted Data. *IACR Cryptology ePrint Archive*, *2017*, 857.
17. Samal, M., Saradhi, V. V., & Nandi, S. (2017). Scalability of Correlation Clustering. The Pattern Analysis and Applications.
18. Tanveer, M., & Shubham, K. (2017). A regularization on Lagrangian twin support vector regression. *International Journal of Machine Learning and Cybernetics*, *8*(3), 807-821.
19. Tanveer, M. (2017). Linear programming twin support vector regression. *Filomat*, *31*(7), 2123-2142.
20. Dheenadayalan, K., Srinivasaraghavan, G., & Muralidhara, V. N. (2017). Policy Gradient Reinforcement Learning for I/O Reordering on Storage Servers. In *International Conference on Neural Information Processing* (pp. 849-859). Springer, Cham.
21. Dheenadayalan, K., Srinivasaraghavan, G., & Muralidhara, V. N. (2017). Self-tuning filers—overload prediction and preventive tuning using pruned random forest. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 495-507). Springer, Cham.
22. Varshneya, D., & Srinivasaraghavan, G. (2017). Human trajectory prediction using spatially aware deep attention models. *arXiv preprint arXiv:1705.09436*.

23. Subramanian, A., Garg, A., Poddar, O., & Srinivasa, S. (2017). Towards Semantically Aggregating Indian Open Government Data from data. gov. in. In *International Semantic Web Conference (Posters, Demos & Industry Tracks)*.
24. Vaghmare, A. K., & Rao, C. V. R. (2017). Unsupervised noise removal technique based on constrained NMF. IET Signal Processing, 11(7), 788-795.
25. Powell, B. M., Kalsy, E., Goswami, G., Vatsa, M., Singh, R., & Noore, A. (2017). Attack-resistant aiCAPTCHA using a negative selection artificial immune system. In *2017 IEEE Security and Privacy Workshops (SPW)* (pp. 41-46).
26. Mohamed, M., Gao, S., Sachdeva, N., Saxena, N., Zhang, C., Kumaraguru, P., & Van Oorschot, P. C. (2017). On the security and usability of dynamic cognitive game CAPTCHAs. *Journal of Computer Security*, *25*(3), 205-230.

Annexure VII: List of institutions/departments working in the area of AI and cyber security

List of Indian institutions (Indian Institute of Technology (IIT), National Institute of Technology (NIT), Indian Institute of Information Technology (IIIT)) and R&D labs working in the areas of Artificial Intelligence and cyber security based on web search, in addition to the inputs received through the questionnaire on CybSec4AI, are given below:

- Department of Computer Science and Engineering, IIT Madras
- Computer Science and Engineering Department, IIT Kanpur
- Department of Computer Science & Engineering, IIT Delhi
- Department of Computer Science & Engineering, IIT Guwahati
- Department of Computer Science & Engineering, IIT Kharagpur
- Department of Computer Science & Engineering, IIT Jammu
- Department of Electrical Engineering and Computer Science, IIT Bhilai
- Department of Computer Science & Engineering, IIT Bombay
- Department of Computer Science & Engineering, IIT Jodhpur
- School of Computer Science and Engineering, IIT Goa
- Discipline of Mathematics, IIT Indore
- Department of Computer Science and Engineering, IIT Mandi
- Department of Computer Science and Engineering, IIT Ropar
- WiNet Lab, IIT Hyderabad
- Machine Intelligence Unit, Indian Statistical Institute, Kolkata
- Department of Computer Science, Dr. B.R. Ambedkar NIT Jalandhar
- Department of Electronics and Communication Engineering, Malaviya NIT Jaipur
- Department of Computer Science & Engineering, Maulana Azad NIT Bhopal
- Department of Computer Science & Engineering, Motilal Nehru NIT, Allahabad
- Department of Computer Science & Engineering, NIT Raipur
- Department of Computer Science & Engineering, NIT Delhi
- Department of Computer Science & Engineering, NIT Goa
- Department of Computer Science & Engineering, NIT Hamirpur
- Department of Computer Science & Engineering, NIT Jamshedpur
- Department of Computer Science & Engineering, NIT Manipur
- Department of Computer Science & Mathematics, NIT Srinivasanagar
- Department of Computer Science & Engineering, NIT Govt. Polytechnic, Srinagar
- Department of Computer Science & Engineering, NIT Puducherry
- Department of Mathematical and Computational Sciences, NIT Surathkal
- Department of Computer Science and Department of Electronics and Communication Engineering, NIT Meghalaya
- Department of Computer Science, Indian Institute of Engineering Science and Technology, Shibpur

- Department of Computer Science & Engineering, Department of Electronics & Communication Engineering, IIIT Delhi
- Centre for Security, Theory and Algorithmic Research (CSTAR), IIIT Hyderabad
- Department of Computer Science & Engineering, Dr. SPM IIIT Naya Raipur
- Department of Computer Science & Engineering, IIIT Bangalore
- Department of Computer Science & Engineering, IIIT Kottayam
- Department of Computer Science & Engineering, IIIT Nagpur
- C-DAC Kolkata
- C-DAC Mohali
- C-DAC Mumbai
- C-DAC Bangalore

Other Indian Institutions working in the areas of Artificial Intelligence and cyber security are,

- IIT Hyderabad
- IIT Gandhinagar
- Indian Institute of Science (IISc.), Bangalore
- IIIT Pune
- IIIT Hyderabad
- Tata Institute of Fundamental Research (TIFR)
- Jadavpur University
- College of Engineering, Guindy (CEG)
- Madras Institute of Technology (MIT)
- Anna University K B Chandrashekar (AU-KBC) Research Centre, MIT

Annexure VIII: TCS Research

TCS Research has got a special focus on CybSec4AI:
TCS has 11 Research Areas (RA) as part of our corporate research. The following RAs are most pertinent for our current context:

- Data & decision sciences RA
- Deep learning and Artificial Intelligence RA
- Cyber security and privacy RA

Cyber security and privacy RA has five themes

- Secure, safe and private AI/ML
- Confidentiality and integrity of data in cloud
- Securing IoT
- Privacy by design
- Strengthening the 'Carbon' layer

It also has multiple (10+) research projects including one on 'Secure and Private Learning' under 'Secure, Safe and Private AI/ML' theme.

Professor Yuval Elovici, Director of Cyber Research at Ben-Gurion University, Israel is their research adviser. Yuval works at the intersection of AI and CS.

TCS Innovations, be in Cyber security or otherwise, make judicious use of AI/ML:
Cyber security and privacy RA has delivered multiple R&D based products. The biggest two are:

- TCS MasterCraftDataPlus (for data quality, privacy, governance)
- TCS Consent Management Solution (for consent and data rights management)

Not only these two, but numerous other TCS innovations, products and platform offerings use AI/ML.

Cyber security is one of the top three patent portfolios of TCS.

Cyber security and Privacy RA is driving multiple collaboration in CybSec4AI domain:

- Blavatnik Interdisciplinary Cyber Research Centre at Tel Aviv University, Israel
  - Currently a sponsored research project titled 'Secure Deep Learning for Vision Applications' with Professor Shai Avidan and Professor Benny Applebaum.
- Australian Cyber Security Cooperative Research Centre, Australia
  - Multiple projects involving AI/ML under consideration, e.g., 'Privacy Preserving Machine Learning'.
- Yale Institute of Network Science at Yale University, USA
  - Currently a sponsored research project titled 'Data Analytics with The Right to be Forgotten' with Prof Amin Karbasi.

TCS Business Units are keen on CybSec4AI:

- TCS cyber security practice has marked CybSec4AI as a topic of interest.
  - This unit has 4000+ professionals, 350+ million USD revenue (FY19).
  - It provides services in multiple domains, e.g., IDAM, VAPT, GRC, (i) OT, etc.

  - It is one of the fastest growing business for TCS.
- Interest from TCS Analytics & Insights (A&I) business unit in "Trustworthiness of AI".

TCS customers are interested:

- Several customers are seeking TCS story, point of view (PoV) on CybSec4AI.
- TCS has presented their PoV on CybSec4AI in TCS Innovation Forums held at London and NY this year (2019). These events are attended by senior executives from customer organizations.
- AI and cyber security were the top two picks by the attendees during our US Innovation Forum last year (2018) when they were asked to choose their highest priority interest areas.

Also visible in our community service work:

- International conference on Cyber Security, Cryptography and Machine Learning (CSCML)
  - In collaboration with Prof Shlomi Dolev, Ben-Gurion University, Israel for past three years (2017, 18, 19)
- TCS research scholar program across Indian academia
  - 20+ PhD students working on cyber security and privacy topics
  - Many more working in the AI/ML domain

Additional (related) pointers on Trustworthiness in AI

IBM is offering a platform called AI Fairness 360, an open-source toolkit of metrics to check for unwanted bias in datasets and machine learning models, and state-of-the-art algorithms to mitigate such bias. URL: https://www.research.ibm.com/artificial-intelligence/trusted-ai/

- Accenture white paper
  URL: https://www.accenture.com/_acnmedia/pdf-92/accenture-afs-responsible-ai.pdf
- E&Y states that it has built a trusted AI framework to help enterprises (white paper)
  URL: https://www.ey.com/en_in/digital/how-do-you-teach-ai-the-value-of-trust
- KPMG white paper
  URL:https://assets.kpmg/content/dam/kpmg/ph/pdf/services/TrustInArtificialIntelligence.pdf

Annexure IX: National workshop on Deep Learning for Cyber Security

On 6-7 March 2020, SETS organized a National Workshop on Deep Learning for Cyber Security that focussed on the applications of Artificial Intelligence (AI) for cyber security, industry's take on shaping AI trajectories in security, investigating the robustness and resiliency of AI systems and the rise of Adversarial Learning with its impact. Following is the summary of those discussions, framed around research questions and possible topics for future research directions.

The workshop was inaugurated by Dr. R. Chidambaram, Homi Bhabha Professor, BARC, Former Principal Scientific Adviser to the Government of India and President SETS, Chennai. Shri. R. S. Mani, Head, National knowledge Network, NIC, Chennai presented the special address and Dr.B. Ravindran, Professor, Indian Institute of Technology, Madras delivered the keynote address.

Dr. R. Chidambaram in his Inaugural address stated that information searched on the internet travels through a multitude of servers before it reaches the user. Each server is only adding to the million lines of logs that it maintains and added that the time has come that we look at those logs, gather valuable information that would thwart cyber-attacks and strengthen the cyber security measures of our nation.

Shri. R. S. Mani in his special address stated the end user is more conscious of the speed of the internet; along with the speed is the cost of detectability of attacks. There has always been a trade-off between the speed and attack detection latency. Denial of Service (DOS) attacks last a few minutes, even before the user is aware of the attack, the attack has happened. He added that combining the strengths of AI with cyber security, security professionals will have additional resources to defend vulnerable networks and data from cyber attackers.

Prof. Ravindran started his keynote address with Turing's vision and the journey from biological neural network to the present advanced recurrent and attention networks. He also stressed on how Reinforcement Learning (RL) models learn by a continuous process of receiving rewards and punishments on every action taken and how it is able to train systems to respond to the unforeseen environments.

The important practical/implementation considerations that must be taken into account when training neural networks have been presented by one speaker. Some of the considerations include:

- Do we need to pre-process the training data? If so, how?
- How many hidden units do we need?
- Are some activation functions better than others?
- How do we choose the initial weights from which we start the training?
- Should we have different learning rates for the different layers?
- How do we choose the learning rates?
- Do we change the weights after each training pattern, or after the⁄whole set?
- How do we avoid flat spots in the error function?
- How do we avoid local minima in the error function?
- When do we stop training?

In the session on Artificial Intelligence for cyber security, *DL-based face biometrics was elaborately presented.* DL techniques can be adopted for malware detection using hash functions. In another lecture, *unsupervised learning using auto-encoders and Restricted Boltzmann Machine (RBM) for cyber security was discussed. Another presentation stressed on supervised learning- Recurrent Neural Networks (RNN)s for cyber security.*

In another lecture, IDS Dataset Preparation was considered in detail. The criteria that are necessary for building a reliable benchmark dataset have been highlighted. Another presentation dealt with *next gen malware detection using deep learning techniques.* Other topics were on *mitigation of cyber-attacks through deep learning and similarity learning using CNN. Machine learning based DDoS detection* was described in another presentation*.* Interestingly, evaluation of side channel analysis using deep learning was also discussed.

In the session on cyber security for Artificial Intelligence, the topics dealt with are *privacy techniques in machine learning* and *deep neural networks implementation considerations.* Another topic was *towards trust worthy deep learning* which covered recent findings of adversarial attacks on DL, counter measures and defence methods.

One presentation described the national strategy on AI and cyber security (AI-CS). It was stressed that it is imperative for us to build self-reliance in Technology development in "Cyber security tools using AI/ML and building secured AI/ML systems and applications" for protecting our ICT (Information and Communications Technology) systems including critical infrastructure. A brief survey of initiatives on AI-CS by international institutions, and by SETS on AI-CS Project Management Unit (PMU) was presented.

There was a session on transfer learning which is a machine learning method where a model developed for a task is reused as the starting point for a model on a second task. It was observed that transfer learning saves training time, increasing the performance of neural networks, and does not need a lot of data. It was highlighted that cyber security problems that can be addressed using transfer learning are personalized spam filtering, intrusion detection systems were also discussed.

Conclusion

This workshop witnessed the presence of eminent speakers from R&Ds, industry and academia where each presented their recommendations and suggestion to address threats and strengthening cyber security posture not only as individuals but also a nation. This report presents the SETS perspective of sessions, happened at the National Workshop on Deep Learning for Cyber Security. It was communicated that R&D, Industries and academia work collaboratively to outrun adversarial-AI counterparts.

Acknowledgement

SETS thank Dr. R. Chidambaram, Shri. R. S. Mani, Dr. B. Ravindran and the eminent speakers from various domains of Government R&D labs, industries, academia and SETS are greatly acknowledged.

Annexure X: Major areas/institutions identified for International collaboration

The following research departments working on cyber security and AI in various universities abroad are identified, initially, for possible international collaborations.

1) Department of Computer Science, University of Texas at Dallas

Research Areas: Research in data security and data analytics for security applications.

2) Department of Computer Science and Engineering, Arizona State University

Research Areas: Social media analysis & information integration, analysis of human-behaviour on social media platforms, adaptive techniques for query optimization and execution in information integration, source discovery and source meta-data learning.

3) Saarland University, Germany

Research Areas: Trusted cloud computing, studies in social computing systems attempt to understand, predict and control the behaviours of constituent human users and computer systems, tackling the challenge associated with assessing the credibility of information shared by anonymous online crowds, understanding and controlling privacy risks for users sharing data on online forums, understanding, predicting and influencing viral information diffusion on social media sites, and enhancing fairness and transparency of data driven (algorithmic) decision making increasingly being used to model and replace human decision making in social computing systems.

4) Electrical and Computer Engineering, Carnegie Mellon University

Research Areas: The goal is to ensure that data-driven systems that employ AI and machine learning are not inscrutable black-boxes; rather their operation is explained in a form that enables trust in their operation and protection of societal values, including privacy and fairness, developing the areas of privacy through accountability and compositional security.

5) Department of Electrical Engineering & Computer Sciences, University of California at Berkeley

Research Areas: Development of theory and tools to aid the construction of provably dependable and secure systems; Seeking to advance the state-of-the-art in automated formal methods through the following thrust:

- Computational engines: Development of efficient algorithms and tools for core computational problems such as satisfiability modulo theories (SMT) solving, model counting, and syntax-guided synthesis.
- Algorithmic strategies for verification and synthesis: Development of new formal verification and synthesis techniques based on a combination of inductive inference and deductive reasoning, an approach outlined with illustrative applications.
- New application frontiers: Pursuing new applications of formal methods.
- Cyber-physical systems (CPS): Verification and control of Human-CPS, cyber-physical systems that operate in concert with humans, such as semi-autonomous vehicles
- Computer security: Focus on systems that leverage trusted hardware-software platforms.

6) Ben-Gurion University of the Nege, Department of Software and Information Systems Engineering

Research Areas: Detection of malicious code using machine learning techniques, social network security, privacy and anonymity in the electronic society, complex network.

7) Department of Computer Science, Virginia Tech, USA

Research Areas: Data mining and applied machine learning with emphasis on solving big-data problems in networks and sequences; understanding and managing efficiently, dynamical mechanisms (like propagation) on networks, occurring across natural, social and technological systems; these include problems motivated from public health, cyber security, urban computing and social media. Theoretical analysis of models, developing efficient algorithms and empirical studies on large scale data.

8) Computer Science and Engineering, University of California, San Diego

Research Areas: Trustworthy machine learning, which includes problems such as learning from sensitive data while preserving privacy, learning under sampling bias, and in the presence of an adversary.

9) Computer Science Department, Cyber Security, Technology, and Society (ISTS), Dartmouth College

Research Areas: Behavioral analysis of insider threat, optimal defense of enterprise systems, identifying bad actors in social media, global cyber-vulnerability report, automatically learning models of high-level activities (e.g., cyber-attacks) from observations and logs, developing a game-theoretic method to alert cyber analysts about which machines to monitor within an enterprise.

10) School of Computer Science, Carnegie Mellon University

Research Areas: Privacy risk in machine learning, feature wise bias amplification, individual fairness revisited: transferring techniques from adversarial robustness.

11) Berkeley Artificial Intelligence Lab and of the EECS department, UC Berkeley

Research Areas: Ensuring safety of AI systems.

12) Department of Electrical and Electronic Engineering, University of Cagliari, Italy Pluribus One, Cagliari, Italy

Research Areas: Adversarial machine learning, evasion attacks, poisoning attacks, adversarial examples, secure learning, and deep learning.

13) School of Computer Science, Georgia Institute of Technology and School of Computing and Information, University of Pittsburgh

Research Areas: Privacy-preserving federated learning.

14) University of California Berkeley, Department of Computer and Information Science, Linkoping University

Research Areas: Malware detection.

15) Osaka University

Research Areas: Machine learning for security.

16) Department of Computer Science, Aalto University, EECS at UC Berkeley, Department of Electrical Engineering, Princeton University and Department of computer science in Purdue University

Research Areas: Adversarial machine learning and adversarial examples.

Annexure XI: Glossary of Terms and List of Abbreviations

5G	Fifth Generation of wireless communications technologies supporting cellular data networks
AI-CS	Artificial Intelligence- Cyber Security
AI-CS-CBP	AI-CS Capacity Building Program
Botnet	A number of Internet-connected devices, each of which runs one or more bot
CybSec4AI	Cyber Security for Artificial Intelligence/Machine Learning system, solutions and products
DFG	Deutsche Forschungs Gemeinschaft is the self-governing organisation for science and research in Germany
e-Abhedya	A Public Key Infrastructure solution by SETS
Honeypot	A network-attached system set up as a decoy to lure cyber attackers and to detect, deflect or study hacking attempts in order to gain unauthorized access to information systems
NITI Aayog	National Institution for Transforming India is a policy think tank of the Government of India
T10kT	'Train 1000 Teachers' programme initiated by IIT Bombay

List of Abbreviations

ACM	Association for Computing Machinery
ACTS	Advanced Computing Training School
ADFA-WD	Australian Defence Force Academy-Windows Dataset
AI	Artificial Intelligence
AICSET	AICS Education and Training
ATD	Algorithms for Threat Detection
B5G	Beyond the 5th Generation
B.Tech.	Bachelor of Technology
BEL	Bharat Electronics Limited
BFSI	Banking and Finance Services and Insurance
BIS	Bureau of Indian Standards
B-PPDR	Broadband Wireless Communications System for Public Protection and Disaster Recovery
C3I	Centre for Cyber security and Cyber defense of Critical Infrastructures
CAIDA	Centre for Applied Internet Data Analysis
CBSE	Central Board of Secondary Education
CCoE	Cyber security Centre of Excellence
CCS	Computer and Communications Security
C-DAC	Centre for Development of Advanced Computing
CDX	Cyber Defense eXercise
CERT-In	Indian Computer Emergency Response Team

CHESS	Computers and Humans Exploring Software Security
CI	Critical Infrastructure
CIA	Confidentiality- Integrity- Availability
CICI	Cyber security Innovation for Cyber Infrastructure
CMU	Carnegie Mellon University
COA	Course of Action
CODS	Conference on Data Sciences
CORE	Centre of Research Excellence
CPS	Cyber Physical Systems
CPU	Central Processing Unit
CS	Cyber Security
CSCRC	Cyber Security Cooperative Research Centre
CSIR-4PI	Council of Scientific and Industrial Research- Fourth Paradigm Institute
CTF	Catch the Flag
CTI	Cyber Threat Intelligence
CVE	Common Vulnerabilities and Exposures
CWMP	CPE (Customer Premises Equipment) WAN (Wide Area Network) Management Protocol
DARPA	Defense Advanced Research Projects Agency
DDOS	Distributed Denial of Service
DEFCON	DEFense readiness CONdition
DHS	Department of Homeland Security
DL	Deep Learning
DNS	Domain Name System
DoD	Department of Defense
DOT	Department of Transportation
DP	Differential Privacy
DQND	Deep-Q-Network Detection
DRL	Deep Reinforcement Learning
DSCI	Data Security Council of India
E&Y	Ernst and Young
EMBER	Endgame Malware BEnchmark for Research
FDP	Faculty Development Program
FHWA	Federal Highway Administration
FOSS	Free and Open Source Software
FTP	File Transfer Protocol
GAN	Generative Adversarial Network
GDPR	General Data Protection Regulation
GIS	Geographic Information System
GOI	Government of India
GSES	Grid Security Expert System
HIDS	Host based Intrusion Detection System
HQN	High Quality Node

HTTP	Hyper Text Transfer protocol
HTTPS	Hypertext Transfer Protocol Secure
IBM	International Business Machines
ICICS	International Conference on Information and Communication Systems
ICS	Industrial Control System
ICSP	Industrial Control System Protection
ICT	Information and Communication Technology
ICTAI	International Centres of Transformational AI
IDS	Intrusion Detection System
IEEE	Institute of Electrical and Electronics Engineers
IIIT	Indian Institute of Information Technology
IIoT	Industrial Internet of Things
IISc.	Indian Institute of Science
IIT	Indian Institute of Technology
IMT2020	International Mobile Telecommunications 2020
IoT	Internet of Things
IP	Internet Protocol
IPR	Intellectual Property Rights
ISEA	Information Security Education and Awareness
IS	Information Systems
ISO	International Organization for Standardization
ISP	Internet Service Providers
IT	Information Technology
ITES	Information Technology Enabled Services
ITMA	Integrated Threat Management Appliance
KDD	Knowledge Discovery in Databases
KDE	Key Distillation Engine
LPN	Learning Parity with Noise
LWE	Learning With Errors
M.Tech.	Master of Technology
Mal-API	Malware-Application Program Interface
MBA	Master of Business Administration
MCA	Master of Computer Application
MeitY	Ministry of Electronics and Information Technology
MIT	Massachusetts Institute of Technology
ML	Machine Learning
MOOC	Massive Open Online Course
MoU	Memorandum of Understanding
MPC	Multi Party Computation
MQTT	Message Queuing Telemetry Transport
MS	Master of Science
NASSCOM	National Association of Software and Service Companies
NBIoT	Narrowband Internet of Things

NCATS	National Centre for Advancing Translational Sciences
NCCC	National Cyber Coordination Centre
NCERT	National Council of Educational Research and Training
NCETIS	National Centre of Excellence in Technology for Internal Security
NCI	National Cancer Institute
NCIIPC	National Critical Information Infrastructure Protection Centre
NFV	Network functions virtualization
NGIDS	Next Generation Intrusion Detection System
NIBIB	National Institute of Biomedical Imaging and Bioengineering
NIFA	National Institute of Food and Agriculture
NIH	National Institutes of Health
NIPS	Neural Information Processing Systems
NIST	National Institute of Standards and Technology
NIT	National Institute of Technology
NKN	National Knowledge Network
NLP	Natural Language Processing
NMEICT	National Mission on Education through Information and Communication Technology
NPRS	National Privacy Research Strategy
NSF	National Science Foundation
NSM	National Supercomputing Mission
NSTC	National Science and Technology Council's
NTRO	National Technical Research Organisation
OSM	Online Social Media
PCAP	Packet CAPture
pdAD	physical domain Attack Detection
PDF	Portable Document Format
PE	Portable Executables
PG	Post-Graduate
PG-DITISS	Post-Graduate Diploma in IT Infrastructure Systems and Security
PhD	Doctor of Philosophy
PLC	Programmable Logic Controller
PMG	Project Management Group
POP3	Post Office Protocol version 3
PPP	Public Private Partnership
PSA	Principal Scientific Adviser
PSU	Public Sector Undertakings
PUF	Physically Unclonable Functions
QKD	Quantum Key Distribution
R&D	Research and Development
RAPPOR	Randomized Aggregatable Privacy Preserving Ordinal Response
RC	Resource Centre
RDP	Research Data Protection

RDSP	Research and Development Strategic Plan
RMB	Ren Min Bi – Chinese Currency
RSA algorithm	Rivest, Shamir, Adleman algorithm
S&T	Science and Technology Directorate
SaTC	Secure and Trustworthy Cyberspace
SCADA	Supervisory Control And Data Acquisition
SDC	State Data Centres
SDLC	Secure Software Development Lifecycle
SDN	Software Defined Networking
SEO	Search Engine Optimization
SERB	Science and Engineering Research Board
SETS	Society for Electronic Transactions & Security
SIEM	Security Information and Event Management
sCPS	smart Cyber Physical Systems
SMB	Server Message Block
SMT	Satisfiability Modulo Theories
SMTP	Simple Mail Transfer Protocol
SNAP	Stanford Network Analysis Project
SNMP	Simple Network Management Protocol
SOC	Security Operation Centre
SPACE	Security, Privacy and Applied Cryptography Engineering
SQLDB	Structured Query Language Data Base
SSC	Secure Scientific Cyberinfrastructure
SSGAN	Secure Steganography Based on Generative Adversarial Networks
SSH	Secure Socket Shell
TCS	TATA Consultancy Services
TF	Task Force
TFR	Task Force Report
ToR	The onion Router
UG	Under-Graduate
URL	Uniform Resource Locator
US DoD	United States Department of Defense
USDA	U.S. Department of Agriculture
VA	Veterans Affairs
VAPT	Vulnerability Assessment and Penetration Testing
VoIP	Voice over Internet Protocol
VPN	Virtual Private Network
WEF	World Economic Forum
WSN	Wireless Sensor Networks