iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** II **Month of publication:** February 2017

DOI: http://doi.org/10.22214/ijraset.2017.2043

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

Volume 5 Issue II, February 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Analysis on Security Evaluation of Pattern Classifiers under Attack

Rupali B. Navalkar[1], Prof. Rajeshri R. Shelke[2]
[1]ME Second Yr (CSE) H.V.P.M's COET, Amravati
[2]Associate Professor, ME (CSE) H.V.P.M's COET, Amravat

Abstract: ***Analysis on security evaluation of pattern classifiers under attack describes pattern classification systems that are security evaluation problems because of different attacks. Pattern Classification commonly used in adversarial applications, like biometric authentication, network intrusion detection, and spam filtering. In these applications data can be purposely manipulated by humans to undermine their operation. This adversarial scenario's exploitation may sometimes affect their performance, systems may exhibit vulnerabilities and limit their practical utility. This adversarial scenario is not taken into account by classical design methods. These Applications have an intrinsic adversarial nature since the input data can be purposely manipulated by an intelligent and adaptive adversary to undermine classifier operation. This often gives rise to an arms race between the adversary and the classifier designer. The system evaluates at design phase the security of pattern classifiers, namely, the performance degradation under potential attacks they may incur during operation. A generalize framework is used for evaluation of classifier security that formalizes and generalizes the training and testing datasets, to discriminate between a "legitimate" and a "malicious" pattern class Training and Testing sets have been obtained from distribution using a classical reassembling technique like bootstrapping or cross validation. Security evaluation can be carried out by averaging the performance of the trained and tested data.***

Keywords: Pattern classification, adversarial classification, performance evaluation, security evaluation, Biometric Threats, Spoofing.

I. INTRODUCTION

Analysis on security evaluation of pattern classifiers under attack describes pattern classification systems that are security evaluation problems because of different attacks. Pattern classification systems based on machine learning algorithms are commonly used in security-related applications like biometric authentication, network intrusion detection, and spam filtering, to discriminate between a "legitimate" and a "malicious" pattern classes (e.g., legitimate and spam emails). Contrary to traditional ones, these Applications have an intrinsic adversarial nature since the input data can be purposely manipulated by an intelligent and adaptive adversary to undermine classifier operation. This often gives rise to an arms race between the adversary and the classifier designer. examples of attacks against pattern classifiers are: submitting a fake biometric trait to a biometric authentication system (spoofing attack); modifying network packets belonging to intrusive traffic to evade intrusion detection systems (IDSs); manipulating the content of spam emails to get them past spam filters (e.g., by misspelling common spam words to avoid their detection) . Biometric security threats: identified several different levels of attacks that can be launched against a biometric system (i) a fake biometric trait such as an artificial finger may be presented at the sensor, (ii) illegally intercepted data may be resubmitted to the system.

It is now acknowledged that, since pattern classification systems based on classical theory and design methods do not take into account adversarial settings, they exhibit vulnerabilities to several potential attacks, allowing adversaries to undermine their effectiveness. A systematic and unified treatment of this issue is thus needed to allow the trusted adoption of pattern classifiers in adversarial environments. A framework is used for evaluation of classifier security that formalizes and generalizes the training and testing datasets. Results show that security evaluation can provide a more complete understanding of the classifier's behaviour in adversarial environments, and lead to better design choices. Adversarial machine learning is a research field that lies at the intersection of machine learning and computer security.

II. LITERATURE REVIEW

Here we review previous, Security Evaluation of Pattern classification systems based on machine learning algorithms are commonly used in security-related applications like biometric authentication, network intrusion detection, and spam filtering, to discriminate between a "legitimate" and a "malicious" pattern class[1][2]. Pattern classification systems based on classical theory and design

©IJRASET: All Rights are Reserved

methods do not take into account adversarial settings. They exhibit vulnerabilities to several potential attacks, allowing adversaries to undermine their effectiveness [3]. Biometric systems have been found to be useful tools for person identification and verification. A biometric characteristic is any physiologic al of behavioural trait of a person that can be used to distinguish that person from other people [2][6]. Spoof attacks consist in submitting fake biometric traits to biometric systems [2][4], and this is a major threat in security.

The presence of an intelligent and adaptive adversary makes the classification problem highly non-stationary[1], and makes it difficult to predict how many and which kinds of attacks a classifier will be subject to during operation, that is, how the data distribution will change. In particular, the testing data processed by the trained classifier can be affected by both exploratory and causative attacks, while the training data can only be affected by causative attacks, In both cases, during operation, testing data may follow a different distribution than that of training data, when the classifier is under attack. Therefore, security evaluation cannot be carried out according to the classical paradigm of performance evaluation [1][2][3].

Security problems often lead to a "reactive" arms race between the adversary and the classifier designer [2]. At each step, the adversary analyzes the classifier defences, and develops an attack strategy to overcome them [1]. Many authors implicitly performed security evaluation as a what-if analysis, based on empirical simulation methods; they mainly focused on a specific application, classifier and attack, their goal was either to point out a previously unknown vulnerability, or to evaluate security against a known attack.

III. EXISTING SYSTEM

Security Evaluation of pattern classification systems based on classical theory and design methods do not take into account adversarial settings; they exhibit vulnerabilities to several potential attacks, allowing adversaries to undermine their effectiveness. A systematic and unified treatment of this issue is thus needed to allow the trusted adoption of pattern classifiers in adversarial environments, starting from the theoretical foundations up to novel design methods, extending the classical design cycle of. In particular, three main open issues can be identified: (i) analyze the vulnerabilities of classification algorithms, and the corresponding attacks. (ii) Developing novel methods to assess classifier security against these attacks, which are not possible using classical performance evaluation methods. (iii) Developing novel design methods to guarantee classifier security in adversarial environments.

IV. PREVIOUS REVIEW ON SECURITY EVALUATION

Previous review in the adversarial learning system can be categorized according to the two main steps, the pro-active arms race and the re-active arms race. It pivoted on identifying vulnerabilities of the adversarial learning algorithms and assessing impact of corresponding attacks on the targeted classifier.

A. Arms-Race Problem

The Arms-Race is linking of the classifier designer and the adversary which is modeled. For example: Fake biometric traits in Biometric Authentication. It analyzes the classifier defenses and develops the Attack strategy to overcome them. The designer reacts by analyzing the novel attack samples, and, if required, updates the classifier; typically, by retraining it on the new collected samples, and/or adding features that can detect the novel attacks. Examples of this arms race can be observed in spam filtering and malware detection, where it has led to a considerable increase in the variability and sophistication of attacks and countermeasures. To secure a system, a common approach used in engineering and cryptography is security by obscurity that relies on keeping secret some of the system details to the adversary.

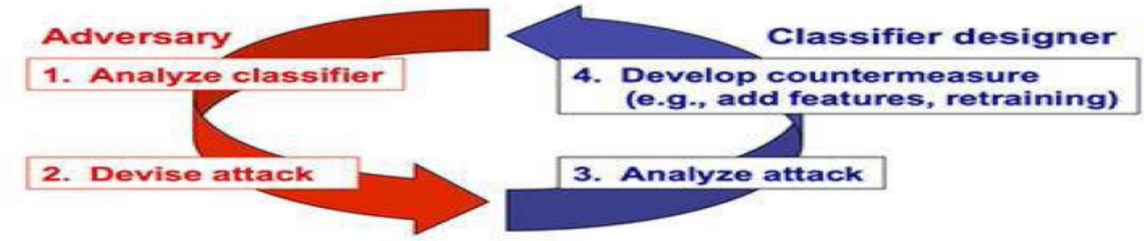


Fig-1: Classical Reactive Arms Race

B. 'Re-active' Arms Race

In this type, the classifier designer reacts to the attack by analyzing its effects and grows the countermeasures. system designer

©IJRASET: All Rights are Reserved

Volume 5 Issue II, February 2017
ISSN: 2321-9653
International Journal for Research in Applied Science & Engineering Technology (IJRASET)

should anticipate the adversary by simulating a "proactive" arms race to (i) figure out the most relevant threats and attacks, and (ii) devise proper countermeasures, before deploying the classifier. This paradigm typically improves security by delaying each step of the "reactive" arms race, as it requires then adversary to spend a greater effort (time, skills, and resources) to find and exploit vulnerabilities. System security should thus be guaranteed for a longer time, with less frequent supervision or human intervention. The goal of security evaluation is to address issue (i) above, i.e., to simulate a number of realistic attack scenarios that may be incurred during operation, and to assess the impact of the corresponding attacks on the targeted classifier to highlight the most critical vulnerabilities.

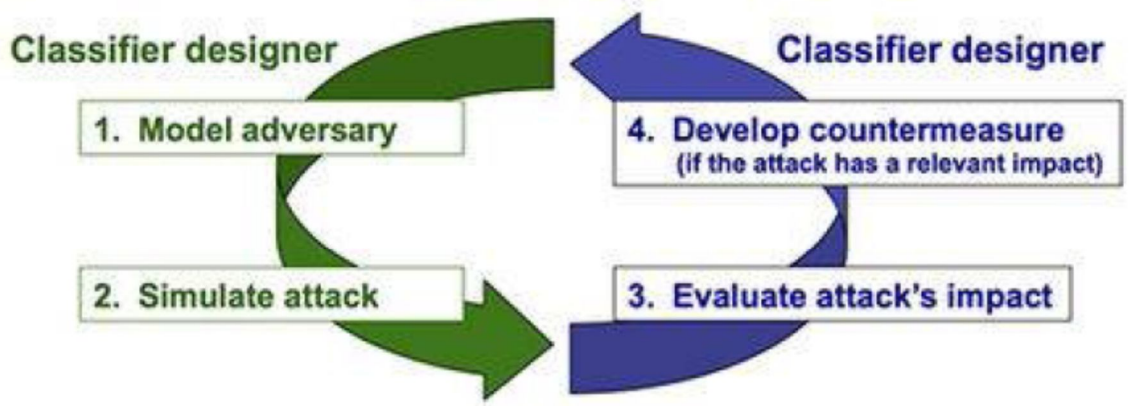


Fig-2: Classical Proactive Arms Race

C. Modules

Attack Scenario and Model of the Adversary ,Pattern Classification

Adversarial classification Security modules

The definition of attack scenarios is ultimately an application-specific issue, it is possible to give general guidelines that can help the designer of a pattern recognition system. Pattern classification is the scientific discipline whose goal is to classify the objects into a number of classes or categories. Depending on the type of application, these objects may be any type of measurements, images or signal waveforms that need to be classified. In pattern classification, typically a set of patterns (the raw data), whose class is unknown, is given. In addition, proper actions can be taken based on the outcome of the pattern classification. a classifier is designed by training it on a set of patterns (samples or feature vectors) whose true class is known referred also as training set or design set, to find a classification function.

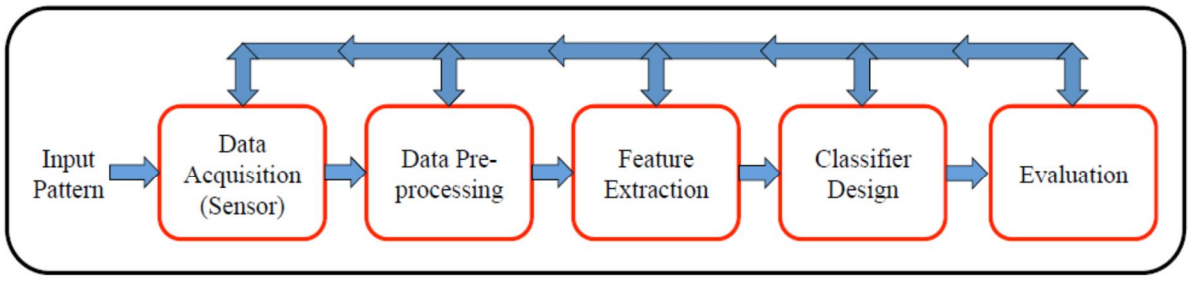


Figure 1.1: The basic stages involved in the design of a pattern classification system.

Figure 1.1 shows the various stages followed for the design of a pattern classification system. The first step is to collect a pre-

©IJRASET: All Rights are Reserved

processed set of training samples. The role of data pre-processing module is therefore to segment the pattern of interest from the background, remove noise and any other operation which will contribute in defining a compact representation of the pattern. Features are then extracted from each training sample. In practice, a larger than necessary number of feature candidates is generated and then the best of them are adopted. The classifier, which is chosen among different algorithms, is trained on appropriate features. Finally, once the classifier has been designed (trained), one can evaluate the performance of the designed classifier.

D. Objectives

Our future work will be devoted to develop techniques for simulating attacks for different applications.
Prevents developing novel methods to assess classifier security against these attacks.
The presence of an intelligent and adaptive adversary makes the classification problem highly non-stationary.
We also propose an algorithm for the generation of training and testing sets to be used for security evaluation, which can naturally accommodate application-specific and heuristic techniques for simulating attacks.

V. APPLICATION EXAMPLES

A. Biometric authentication

In this firstly, our system take the pattern from the user at the time of registration. Then this pattern is saved into the database .Later on when the user trying to use his account, he has to give only this pattern which is entered at the time of registration. Then and then only permission is granted to system particular user. If the attacker try to attack on the account of user, our system firstly match this pattern to the pattern saved in database. If these matches then and then only access is granted otherwise it is consider as attack and access is denied. The admin take appropriate action. We are take pattern as face recognition and trying to improve the performance.

B. Spam filtering

A spam filter is a program is used to detect unwanted email and prevent those messages from getting to a user's inbox. Some methods are not especially effective, too often omitting perfectly legitimate messages and letting actual spam through. Regarding to the Spam filtering the list of bad words are provided. When the mail is occur, if there are most number of links provided or a single word is repeated again and again, then our system is consider that this is a spam mail. Then the system compares those bad words to the bad word list.

If some words are matches then we definitely tell that it is really a spam mail. our system replaces those bad words by good words so that it is become the legitimate mail and we send it into inbox. Then this mail submitted to the legal mails i.e. in inbox. If those mails are not useful then it send to the illegal mails i.e. spam.

B. Nids

NIDS is a network security system used to focus on the attacks that come from the inside of the network (authorized users). Network Intrusion Detection Systems (NIDS) are placed within the network to monitor traffic to and from all devices on the network. It analysis of passing traffic on the entire subnet and match the traffic that is passed on the subnets to the known attacks. Once an attack is identified, or abnormal behavior is sensed, the alert sends to the administrator. In this article, our system prevented to uploading the .exe file which is attach to the mail by attacker. So that the network intrusion does not occur. And if the attack happens, it is handled by our system and Administrator.

VI. CONCLUSION

In this Paper our main contribution is a framework for empirical security evaluation that formalizes and generalizes ideas from previous work, and can be applied to different classifiers, learning algorithms, and classification tasks. It is grounded on a formal model of the adversary, and on a model of data distribution that can represent all the attacks considered in previous work, provides a systematic method for the generation of training and testing sets that enables security evaluation; and can accommodate application-specific techniques for attack simulation. This is a clear advancement with respect to previous work, since without a general framework most of the proposed techniques could not be directly applied to other problems. By combining multiple sources of information, these systems improve matching performance, also the paper focused on innovative security evaluation of pattern classifiers that deployed in adversarial environments and analyzing on security evaluation of pattern classification under attack applying various methods.

©IJRASET: All Rights are Reserved

Volume 5 Issue II, February 2017
ISSN: 2321-9653

REFERENCES

[1] Battista Biggio, Member, IEEE , Giorgio Fumera, Member, IEEE, and Fabio Roli, Fellow, IEEE ," Security Evaluation of Pattern Classifiers under Attack", IEEE TRANSACTIONS ON KNOWLEDGE AND DATA ENGINEERING, VOL. 26, NO. 4, APRIL 2014.

[2] S.P.Mohana Priya[1], S.Pothumani , "Identifying Security Evaluation of Pattern Classifiers Under attack", International Journal of Innovative Research in Science, Engineering and Technology (An ISO 3297: 2007 Certified Organization)Vol. 4, Issue 3, March 2015.

[3] Kale Tai. , Prof. Bere S. S., "A Survey on: Security Evaluation of Pattern Classifiers under Attack", International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 3, Issue 11, November 2015.

[4] Yadigar Imamverdiyev , Lala Karimova, Vugar Musayev , James Wayman, " TESTING BIOMETRIC SYSTEMS AGAINST SPOOFING ATTACKS" , The Second International Conference "Problems of Cybernetics and Informatics" September 10-12, 2008, Baku, Azerbaijan.

[5] Tanisha Aggarwal, Dr. ChanderKant Verma, "Spoofing Technique for Fingerprint Biometric system" ,IJSRD - International Journal for Scientific Research & Development| Vol. 2, Issue 03, 2014 .

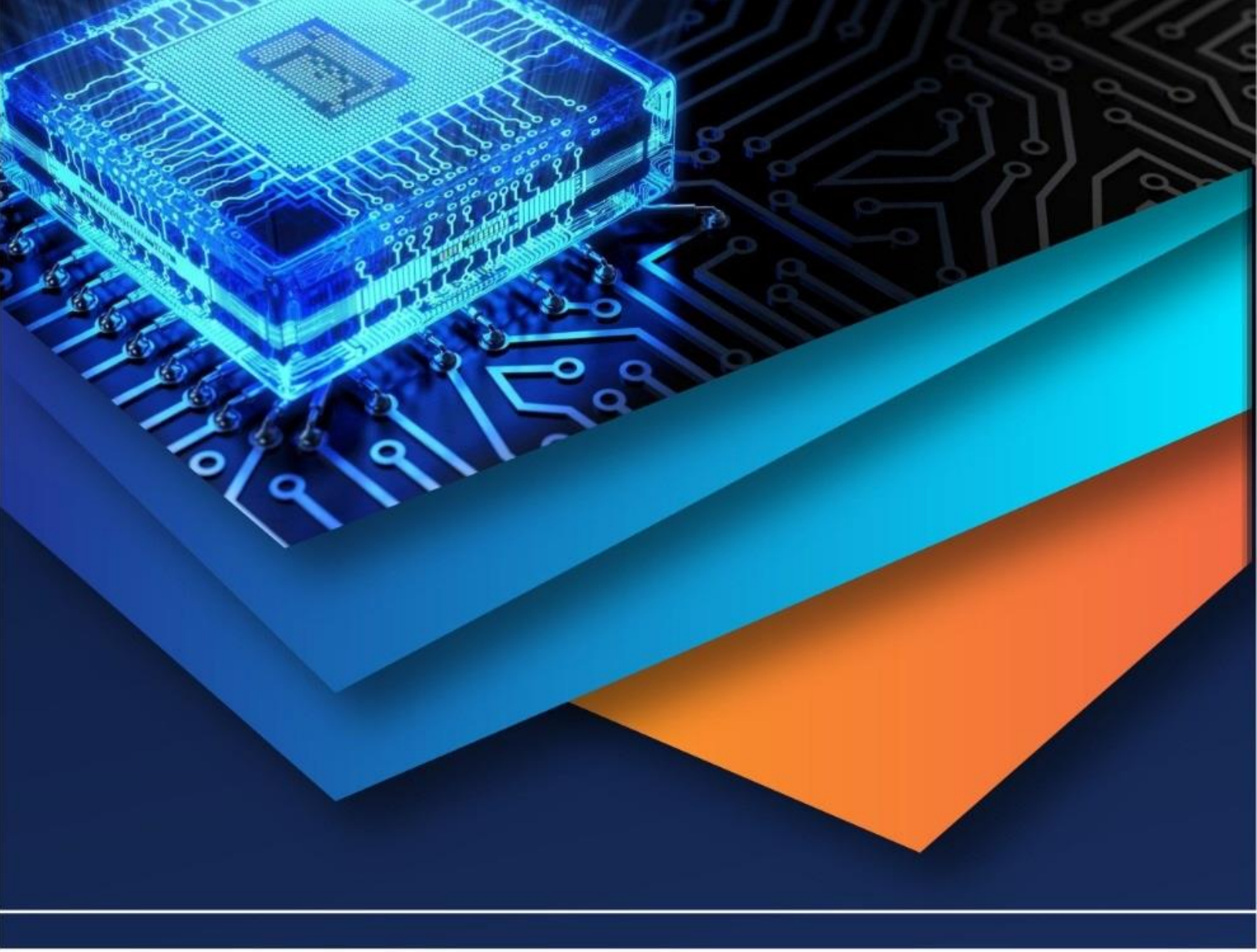
[6] Arun Ross and Anil K. Jain,"MULTIMODAL BIOMETRICS: AN OVERVIEW" , Appeared in Proc. of 12th European Signal Processing Conference (EUSIPCO), (Vienna, Austria), pp. 1221-1224, September 2004.

[7] R.N. Rodrigues, L.L. Ling, and V. Govindaraju, "Robustness of Multimodal Biometric Fusion Methods against Spoof Attacks," J. Visual Languages and Computing, vol. 20, no. 3, pp. 169-179, 2009.

[8] R. R. Shelke ,Dr. V. M. Thakare, Dr. R . V. Dharaskar,"Study of Data Mining Approach for Mobile Computing Environment",International Journal on Computer Science and Engineering (IJCSE) , ISSN : 0975-3397, Vol. 4, 12 Dec 2012 ,pp.1920-1923

[9] Dr. Amitabh wahi, C.Prabhakaran "A Literature Survey on Security Evaluation of Pattern Classifiers under Attack" International Journal of Advance Research in Computer Science and Management Studies Volume 2, Issue 10, October 2014.

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)