Science and Application of Cryptology and Its Implication on Internet of Things

BIMAL ROY

To speak on a subject like this without going into technical things is very challenging for me.

First, I'll try to give you a clear idea about Cryptology, which is our main topic today. After that I'll be discussing Internet of Things (IoT), and finally, the connection between the two.

Before entering into the subject I would like to share with you my experience I had in Surat. About a month back I had gone to Surat, on invitation of the Vice-Chancellor, to be the chief guest at the golden jubilee celebration of the Computer Science Department of the South Gujarat University. After the programme, the Municipal Commissioner of Surat, who was present there, requested me to visit his office, for he wanted to show me something. I accompanied him to his office, where I could see quite a big hall and a number of large screens, similar to TV screens, all around. In front of each screen there was a person, carefully watching the events. Suddenly a sound triggered on the screen, and the person sitting there immediately said—'This vehicle is speeding'. There was no policeman on that part of the road; just sitting in front of the screen, speeding of the car was detected. Another car was seen not following the traffic signal. These reports are being generated and the fines are deposited in the Police Department as usual. All these are controlled from this single room.

The Municipal Commissioner informed me that if any person is detected roaming suspiciously at night, immediately that message is generated and is sent to the nearest police station. Police then interrogate the person. Similarly, if smoke is seen on the screen, the news is immediately passed on to the nearest police station. The entire system is operating very smoothly.

In Surat, 6 cameras in each of the 600 selected areas, making a total of 3600 cameras, have been installed. The photographs taken by the cameras are displayed on the screen, and then, with the help of their software, they interpret the meaning of these photographs. The action to be taken based on these interpretations is generated using another piece of software. If there is any discrepancy in the otherwise normal photograph, the system is beautifully implemented to explain the cause. To my mind, this is the first instance of *Internet of Things*. Its correctness is beyond doubt, and apart from the installation costs, there is no appreciable cost involved. IoT means a lot. The state government is doing this project on a trial basis in a small town in Gujarat—this is the first of its kind in India. I wish all the cities in India, big and small, get this facility.

I have lot many things to discuss about

IoT and the related problems. Whenever you start doing something innovative, there will always be oppositions trying to defy your attempts. I'll discuss all these along with the remedies.

Meaning of cryptology

Now I'll come to *Cryptology*—why is cryptology so important, what the possible threats are, and how to overcome those threats. In order to understand the meaning of cryptology we need to look back. In the word cryptology 'crypt' means to hide and 'logy' means systematic study. So the literal meaning of cryptology is the systematic study of hiding things. It is a scientific process. Why do we need to hide things?

History tells us that hiding things started in the Greek regime. In those days there were many City States in Greece—each State had a king, with around 10000 subjects. The kings of these City States wanted to expand their territories and hence war was a regular feature. Sometimes two kings of such small kingdoms used to collude to defeat another king in the war. It was very common before the medieval period, and these days also it is not rare. In such collusive acts when the kings used to plan for the war, they had to communicate with each other. They used to communicate their strategies through messengers. The third king, against whom the conspiracy was being hatched, on suspecting some foul play from the movement of the messenger, could snatch the paper containing the message.

In order to eliminate such possibility, one king came out with a plan. The message used to be tattooed on the cleanly shaven head of the messenger and the person was sent for his destination only after certain amount of hair had grown. On reaching the proper destination, the messenger was made to shave his head for reading the message. This is our very first evidence of secret communication. But it has a problem. In case the rival king decides to check the messenger in search of any hidden information and by any chance notices some tattoos on the scalp, the message remains no more hidden. So this method of communication is not very secure.

The recent technique is to transform the information which is known technically as *encryption*. To 'encrypt' means sending the message in such a way that common people can understand nothing. Only the person who is worth doing this is entrusted upon the responsibility. In this connection, an example may be given which Julius Caesar adapted. He used to meet his generals, maybe once a month, to tell them the 'key' for that month. The key could be a number in between 1 and 25, say for example, 10. This means in all communications among themselves, the letters were made to shift 10 positions to the right. Thus the letter 'A' would become 'K' and 'X' changed to 'H' (YZABCDEFGH). The recipient would shift each letter 10 positions to the left while reading the message. This is literally known as 'Caesar Cipher'.

We generally believe that any method is a mathematical development—so it cannot be kept hidden. What can be hidden is the number, 10 in this case. If the number of keys is less, it is quite easy to try and read the message. For example, if it is known that the number is in between 1 and 25, maximum 25 attempts will suffice. The possible solution to maintain secrecy is to increase the number of keys so that attempting all the keys may not be quite easy.

During my school days in Narendrapur our English teacher used to advise us to read English newspaper and English books. My first English book-reading started with

Sherlock Holmes, where I read a story—about the dancing figures. That was the story of a culprit who used to draw sequence of different dancing postures at the crime spot after committing each crime. Police and Scotland Yard could find no clues. As the Scotland Yard could not solve it, Sherlock Holmes was given the case. Sherlock Holmes first of all counted the number of distinct dancing postures—these were found to be 26. From this number, he got the idea that probably each figure is depicting one English alphabet. Accordingly there were 26 possibilities for A, 25 possibilities for B, 24 for C and so on. Thus the total correspondence is 26 factorial (26x25x24……x3x2x1). This is a huge number, unlike the case of Julius Caesar where trial was between 1 to 25.

Sherlock Holmes, who was a genius, thought that there must be some way out. He noticed that the dancing postures were not repeated equal number of times and in English language also, the alphabets are not equally frequent. Some of the dancing postures were repeated fewer number of times, some others were seen more frequently. In English, 'E' comes maximum number of times; hence the figure with the maximum frequency must be 'E'. Similarly the figure with least frequency is 'X' or 'Z'. In general, the vowels are more frequent than the consonants. Among the consonants 'T' has the maximum frequency. In terms of frequency in vowels 'A' comes after 'E'. Thus he solved the case by applying his knowledge of English literature. This story of Sherlock Holmes is my first exposure to Cryptology.

My second exposure

My second exposure is a game planned by our mathematics teacher in school. Let me first explain the game—there lies the basic idea. In our Narendrapur hostel, we were not allowed to keep money with us. Since it was prohibited, we were tempted to keep some money with us, and our mathematics sir knew well that we would never disclose the amount on interrogation. With the intention of knowing the amount of money we had on an average, he played a game.

He told each of us to think of a number from 0 and10. We all thought of a number and were then asked to find the average which was nearly 5. He then wanted each of us to add to this number, the amount of money we individually had and tell the sum. We were kind of convinced that from this sum, he would not be able to 'guess' the amount of money each one of us had. So, we cooperated and each one of us told the sum. Once he got all the numbers, the next task was to find the average. This average is the average of the total money we had plus the average of all the numbers we thought of (5 in this case). So from this overall average, when 5 was subtracted, he could obtain the average of the total money we had. This is now known as '*Privacy Preserving Computation*'. This means, to compute some private information that no one wants to disclose. In our game also, sir could know only the average and not the money that we had individually. In those days this was an exposure, but little did I know then that this is Cryptology.

Now the medium of storage and the medium of communication are both considered to be highly insecure; hence proper safeguard is required. The method used for safeguarding is not a secret one. For example, if I install an encryption in my mobile so that any conversation between two persons be encrypted, no one else can make out anything on tapping. But this is not an individual attempt. Here, first of all, a lot of mathematics is required, and then some algorithm is to be developed from that mathematics. The complexity is to be studied; the time needed and how successfully it can be done are also to be studied. There is a software engineering aspect to have the best implementation. After implementation it goes to the electronic engineers who prepare a small chip to be installed in the machine. Thus so many people are involved, hence it is not possible to keep it secret.

It is taken for granted that the entire methodology will be known to all and it is safeguarded against the worst possibility—the channel of storage and the channel of communication are all unsafe. But there is something called a 'key'. If we consider Caesar Cipher, everyone knew that the numbers are to be shifted but the secret thing was by how many positions and that is the key. This key will remain confidential between the sender and the recipient. All the sense of security is related to this key. If I want to communicate something using some methodology, it is assumed that the methodology is open to all. The channel of communication may be phone, e-mail or fax. This channel of communication is completely unsafe. Anyone can see and read what is being sent. The entire mathematical part, even the design of the chip, is not secure. What is not known to others is the key. It is my responsibility to prove that the key is highly safe and secure. I am to convince others that even after knowing everything else, detecting the key is extremely difficult, if not impossible.

We earlier used to play a game called 'knapsack'. In the game there is a set of weights and a particular object whose weight is intended to be determined exactly using those given weights. If the weights are fewer in number and of smaller magnitude, much trial and error is not required. For larger number of weights, say 100 weights, the total number of subsets is 2^{100}, and thus it is almost impossible to try all of them. If by any magic someone can discover the combination of the weights, the person is said to have solved one of the extremely difficult problems, namely, the knapsack problem.

In school days I never knew that factorization is really very difficult. In fact, a very large number cannot be factorized efficiently. Suppose I have built a system assuming that factorization is not an easy job. When I present a full-fledged scheme, everything but the key will be known to all. Using all other information if anyone can discover the key, the person is said to have factorized a very large number. Since it is not possible, people will believe that the system I have developed is a safe way of doing things. One standard of security is shown this way. This is known as—*reduction to a hard mathematical problem*. Many such mathematical problems are there, one example being logarithm. It cannot be found out exactly; few of the type $\log_{10}100$ are very easy ($\log_{10}100 = \log_{10}10^2 = 2$) but majority (for example $\log_{10}117$) cannot be found out exactly. We learnt about Taylor series in college. While expanding the infinite series, correctness depends on the number of terms taken.

Logarithm computation

Actually logarithm computation is especially difficult if the numbers are not real numbers, but discrete. We usually deal, in Cryptology, with the finite numbers that form finite groups. Logarithm computation in this context is still a challenge.

Before starting our work we first make a list of the mathematical problems that cannot be solved. Our development of any system, be it of storage or of communication, ultimately up to the level of ASIC chip, the method is based on that mathematics. We then prove that if anyone can find out the key magically, a computationally difficult problem is essentially solved. This is one notion of security.

Another way of finding whether a method is secure is the notion of information developed by the great American scientist, Claude Shannon. According to Shannon, *if the mutual information between the message to be sent or stored actually and the one that is to be sent or stored be zero, the system is said to be perfectly secure*. That means the cipher does not contain any information about the actual message. Normally mutual information can never be exactly zero; it is abstract. We call it equivalent to tossing unbiased coin independently as many times as we like. The probability of finding a head and a tail in tossing an unbiased coin are equal. The fact that the coin is being tossed independently means one is not influenced by the other. This is an abstract thing; we know that mutual information can never be exactly zero. If it is very close to zero, it is said to be secure.

Thus the security of a system can be studied by two different means—one is through mathematical proof and another is by Shannon's notion. We are carrying out lots of research work as to how to make a system secure.

Internet of things

Let us come back to internet of things. Considering the Surat case, suppose a group is attempting robbery in a house. All the movements are being captured in the camera and if the captured movements are suspicious, report is immediately generated and is sent to the nearest police station. A notorious dacoit having knowledge of modern technology may, however, try to corrupt the channel by introducing some noise into the channel. As a result, whatever message is being captured, the noise will dominate the message.

First, we are to assume that everyone will not cooperate with such a nice system developed by us. There is always some opposition, who will try to corrupt it for their own benefit. In fact, we need not think of securing the system had there been no enemy.

Once I watched a slightly offbeat Hollywood movie—*How Lie Was Born*. In the first part of the film it is shown that no one knows what is lie, everyone is honest and always speaks the truth and never thinks of any evil. The second half shows that it is an apparent contradiction and is not possible in real life. There will be no existence of society under such idealistic situation. A complete truthful society also faces a number of undesirable problems and thus lie was gradually born.

We cannot have any system where everything will go smoothly. There must be aberration and our task is to detect those aberrations. The people who try to damage our innovation might mingle some noise with the signal to spoil it. But in order to safeguard the message we are to take certain steps. Considering the Surat case, without sending the photograph as it is to the office of the municipal commissioner, it can be encrypted. If encrypted, mingling with noise is not possible. Once the encrypted message reaches the destination, there is provision for decryption. We must make sure that our method of encryption can withstand any threat. In fact, this is to be done experimentally by employing any one of the two methods discussed before.

We all are familiar with the Google which maintains a server. Does it mean that all the employees in Google can read my mail? They should not have access and that's why we login to Gmail. Google encrypts the mails in such a way that only their trusted people get the key. Suppose 'A' is sending a mail to 'B'. By the time it reaches the server of the Gmail, it has already been encrypted. Again it gets decrypted before 'B' receives the mail from the server. The communication in between can be considered very vulnerable, experts can open it. Gmail is reasonably, but not absolutely, secure. It is the responsibility of the Gmail authority to prove that their method is safe. They normally use Shannon's notion and are quite sure that no ordinary person can break it.

Another IoT is ATM. It is a serious threat. The moment we insert the card in the ATM machine, communication starts between the ATM and the main server. The server then wants to know the PIN and the transaction is allowed after verification of the PIN. This communication is through a channel. If an expert hacker manages to know the card number, the PIN and the name of the card holder—in no time he can produce an identical card. So those who are installing these ATMs in banks must make sure that the communication between the ATM and the main server of the bank through internet is highly secure. In order to install ATM, permission is taken from the RBI that has a cell which on evaluation certifies the security of the system. Without this certificate, ATM cannot be installed. In fact the type of encryption that is going to be actually implemented is most important. Tapping a channel is nothing difficult for a hacker, especially if the encryption is just ordinary. All our communications through internet are encrypted by any one method.

Suppose I want an air ticket through internet. First of all I'll start a communication of my machine with the Air India site. Before making payment they ask for the OTP (one time PIN) which is being generated and sent to my registered mobile number. Communication of money transfer or payment will be complete only after entering the OTP. Since my credit card has no physical security, OTP is essential. OTP ensures that I am the right person since the OTP is sent in my mobile.

In case of shopping through internet, the communication is quite insecure; there is encryption no doubt, but the encryption is not very strong. They must verify that I am the person who is actually the owner of the credit card that is to be used for payment; otherwise anyone can use my credit card after stealing it. OTP is a means of double-checking. This type of checks and balance are a must for internet use.

I'll give another example—the patient's medical data. Suppose a person gets admitted to a hospital; all the medical data of the patient are recorded in the server of the computer in the hospital. The attending doctors need not be present physically to check the data; instead they can login to their computer at home because these are connected through internet. In this case encryption is done in such a way that only the attributes will have access to it—hence it is known as attribute-based encryption. If only the doctors are the attributes, only they will have access. If the Insurance Company is an attribute as well, the company also will have access. For a single attribute, only one key will do, but if the attributes are more than one, more keys are required depending on the number of attributes.

Summary data

Of course anyone has the right to know the summary data. I personally feel that I should have access to the public information like the success rate in heart or lung operation in a particular hospital. This has not yet been legal, but attempt is on. If a patient is going to get operated by a particular doctor, the patient has every right to know that doctor's success rate in that type of operation, say open heart surgery. Information regarding the success or failure of the individual patients should not be disclosed but the gross information without knowing the name of the patient must be accessible to public.

Many of us don't know that the success rate in IVF (In-Vitro Fertilization) is only 14%; that is 1 in 7. When someone seeks the help of a concerned doctor for IVF, normally the doctor gives a list of all those who could become mothers but does give no information about them who could not. These all are internet of things.

In a hospital the relatives of a patient are to wait for hours to meet the attending doctor who just informs the condition of the patient. I think normally the relatives should have access to these hospital records. For this internet of things, a proper administrative system is to be designed taking into account of those who have access and these are to be ensured. What everyone should be accessed to is the summary of information.

Yesterday I had gone to the Reserve Bank where I talked, over phone, to the would be Governor, Dr Urjit Patel, with whom I shared my views on bank loan. I told him that as a common citizen I have every right to know the bad debt amount of any bank. I also have the right to know the number of defaulters; I need not know their names. Let the banks store the encrypted data. I should get access to know the total default amount and the percentage of the total defaulters who have the maximum default amount. If it is found that 90% of the total default amount is confined to only 3-4 persons and the amount is also quite huge, then disbursing loan like this is itself a retardant. This will get financial transparency. Technically the method of Data Obfuscation can be employed for such purpose. At ISI, we are doing lot of research on this area.

Illegal money transfer

I also informed him about my concern on bitcoin or blockchain; it is high-tech 'Hawala' using crypto. Generally all black money transactions used to take place through the illegal transaction 'Hawala'. Now it is through internet. Huge money is transferred with bitcoin using blockchain technology in such a way that it is virtually impossible to trace. It is operated mainly from Hong Kong. Suppose I purchase 10 bitcoins @ $ 850 each. This purchase process is also so secure, using technology, that it is not possible to hack; only a money transfer can be seen but nothing can be known about the sender and the recipient. This type of anonymity is ensured by those dealing with the bitcoins since their cryptology is very strong. Normally regulatory permission from the RBI is required for any type of money transfer. Bitcoin transfer is also money transfer but without regulation; the RBI has no policy issues. In my opinion, all the illegal payments like the money transaction of the terrorists, black money transfer etc. are being carried out now-a-days through bitcoins. I wanted to know if the RBI has any plan to stop this illegal money transfer.

It is often said that America and Israel sell us only those machineries which they can break. In this connection I have a question regarding the safeguard. Suppose a missile is being purchased from Israel and is installed at Amritsar, targeted at Islamabad.

Is there any guarantee that the missile will reach its destination when triggered? How can I be sure that the missile will not reach Kolkata instead of Islamabad? The methods of formal verification must be employed.

Missiles also have technological developments. Certain tests are to be performed at some steps; otherwise such possibilities as mentioned above cannot be ruled out. The missile is operated from the Sena Bhawan in Delhi immaterial of its place of installation. There also, permissions at different levels are required; the ultimate one is given by the Prime Minister or the President. These all are internet connections —no one can see anything physically. All these activities are internet of things.

During the tenure of Mr Buddhadev Bhattacharyya as the Chief Minister, there was an unsuccessful mine blast near Lalgarh. Actually in such cases some sensors are placed covering most of the journey route of such important citizens as the Chief Minister. The sensors carry out the sabotage test by sensing the presence of explosive if any. The signal obtained is then sent to the police control room (Lal Bazar in case of West Bengal) through internet. The information sensed by the sensor is ultimately reaching the control room by using any satellite or any other network. Without being present physically at the spot, it can be known from the control room itself, where possibly an explosive is being kept. In Lalgarh case the mine blast could not be avoided because of one of the two possibilities. Either our network was not secure or the opponent was scientifically so strong that some noise was mingled with the information sent by the sensor. I personally feel that the second possibility is remote. Thus in our daily life wherever possible we have been heavily dependent on internet; ultimately there is an internet approval. This is internet of things.

What is then our role? Simple information like 'I am fine' need not be secure. But we must secure any secret information. Method of technology has been developed for this purpose. What we have to do is basically a mathematician's job. We are to prove the meaning of 'break' in case anyone manages to do it.

Earlier I mentioned two things. One of these is to show that to 'break' means solving a very difficult mathematical problem and another is to show using Shannon's notion that the mutual information between the actual message and the one being sent is almost zero. If the security is not ensured, I won't purchase that product. Government law-enforcing agencies ensure security before purchase, but I am not quite sure whether our government takes this into consideration in all cases.

Once I was given the responsibility from the Indian Government of certifying the security of the encryption in the walkie-talkies used by a central agency. The encryption was done by Motorola in Chicago. I performed Factory Acceptance Test (FAT) on that apparatus using tests for randomness, robustness, and diffusion. I was present for 7 days at their manufacturing unit where I had to check the security at every stage by applying our methodology. I don't know whether Kolkata Police or other police agencies in the country have the security for the wireless they use. ■

* This article is based on a public lecture delivered in Bengali by Professor Bimal Roy of the Indian Statistical Institute on 23 August 2016 at the Institute. The lecture was organized by the **Vivekananda Science Circle** and is transcripted and translated by **Jayanta Kumar Bandyopadhyay** and **Alaka Banerjee** of the Vivekananda Science Circle.