VOLUME 4
ISSUE 4

SCMS SCHOOL OF ENGINEERING AND TECHNOLOGY

ARCHIVE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

NOVEMBER 2019

Vision

To achieve academic excellence in creating globally competent professionals and ethically strong global workforce in the field of Computer Science & Engineering, facilitating research activities, catering to the ever changing industrial demands and societal needs.

Mission

1.Creating excellence in Computer Science & Engineering through academic professionalism for the changing needs of the society.

2.Establishing centre of excellence for research and for technical development in the area of Computer Science & Engineering.

3.Developing communication skill, team work and leadership qualities for continuing education among the students, through project based and team based learning.

4.Inculcating ethics and human values for sustainable societal growth and environmental protection.

5.Empowering students for employability, aspiring higher studies and to become entrepreneur.

Program Educational Objectives

1. Apply computer science theory blended with mathematics and engineering to model computing systems.

2.Nurture strong understanding in logical, computing and analytical reasoning among students coupled with problem solving attitude that prepares them to productively engage in research and higher learning.

3.Communicate effectively with team members, engage in applying technologies and lead teams in industry.

4.Engage in lifelong learning, career enhancement and adapt to changing professional and societal needs.

EDITORIAL

Asha .S
Assistant Professor

Jency Rena NM
Assistant Professor

Isham Mohamed
Student

Sara Abraham
Student

Prashant G
Student

Naijith Gopal
Student

Sherin Phillip
Student

Navaneeth Asok
Student

CONTENTS

DETECTION AND REMOVAL OF WEB APPLICATION VULNERABILITIES

USING STATISTICAL ANALYSIS AND DATA MINING

-Rosemary Joy

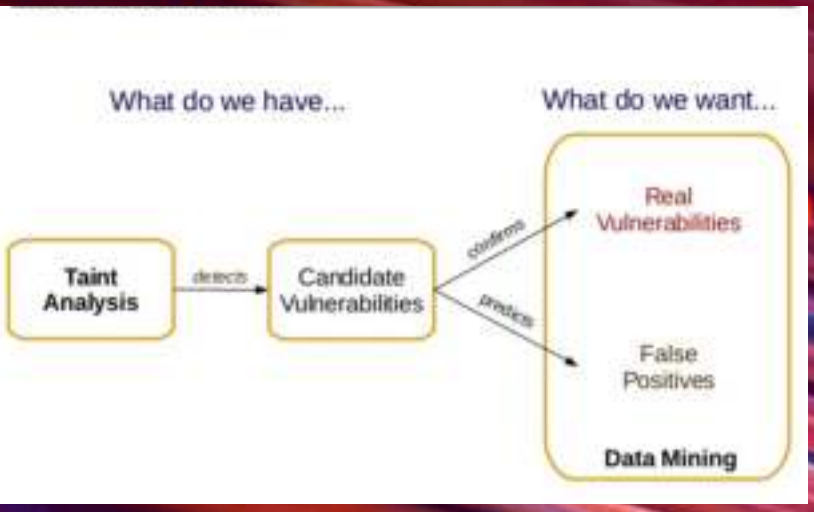
Over the years the world wide web has undergone a complex evolution from a platform that can access text and other media to a framework that can be used for running complex web applications. But all these applications are plagued by security threats. The number of web attacks are facing a upward trend.

The main reason behind this is that programmers lack the appropriate knowledge about secure coding. They create and leave applications with flaws.Even technical giants like Facebook faced web attacks due to their flawed “View As” feature that allows users to view their profiles the way others view it. Thus the types of attacks and the way in which these attacks are executed are becoming more and more diverse.

Most commonly used Web Application Security techniques either keeps the programmer aside and protects the user and the system or they detect the vulnerabilities that exist and leaves the burden of rectifying it to the programmers.

This article attempts to explore a new approach where in the various vulnerabilities in the source code of the web applications will be detected automatically and the fixes of the same will be inserted into the code thus correcting the flaws. The programmer is also kept in loop as he is informed about where the vulnerabilities were detected and also about how these were corrected so that he could learn from his mistakes.

The combination of methods used for this approach is Statistical Analysis with Data mining. Static Analysis can be used to detect the vulnerabilities that exist in the source code of a web application but it reports many false positives (non vulnerabilities) due to its undecidability. In order to prevent this we use a form of statistical analysis called taint analysis to identify the vulnerabilities and data mining to predict the existence of false positives in those detected.



The approach can be implemented as a sequence of steps.

1) Taint analysis: parsing the source code, generating an abstract syntax tree (AST), doing taint analysis based on the AST, and generating trees describing candidate vulnerable control-flow paths from an entry point to a sensitive sink).

2) Data mining: obtaining attributes from the candidate vulnerable control-flow paths, and using the top 3 classifiers to
predict if each candidate vulnerability is a false positive or not. In the presence of a false positive, use induction rules to present the relation between the attributes that classified it.

3) Code correction: given the control-flow path trees of vulnerabilities (predicted not to be false positives), identifying the vulnerabilities, the fixes to insert, and the places where they have to be inserted; assessing the probabilities of the vulnerabilities being false positives; and modifying the source code with the fixes.

4) Feedback: providing feedback to the programmer based on the data collected in the previous steps (vulnerable paths, vulnerabilities, fixes, false positive probability, and the attributes that
classified it as a false positive).

5) Testing: higher assurance can be obtained with two forms of testing specifically program mutation to verify if the fixes
do their function, and regression testing to verify if the behavior of the application remains the same with benign inputs.

This approach was able to find 388 vulnerabilities in 1.4 million lines of code. Its accuracy and precision were approximately 5% better than PhpMinerII's, and 45% better than Pixys.

Reference:https://medium.com

BLOCKCHAIN ENABLED E-VOTING

-Risa Shereen

BLOCKS OR PAPER: WHICH SHOULD GET TO CARRY YOUR VOTE?

The current e-voting system requires more security,privacy,and transparency to become a completely reliable system of voting,it was also observed that the EVM's in use for the election day had modified hardware and/or software in anticipation.

A commercial solution that deals with a token based system built on the blockchain technology .The e-voting process requires features like privacy, security, anonymity,
and verifiability as the core function of this solution.It is important that the choice of the underlying technology is consistent to meet these challenges.

Steps involved in building a blockchain specifically for the electronic voting platform.

1. CREATION OF BLOCK

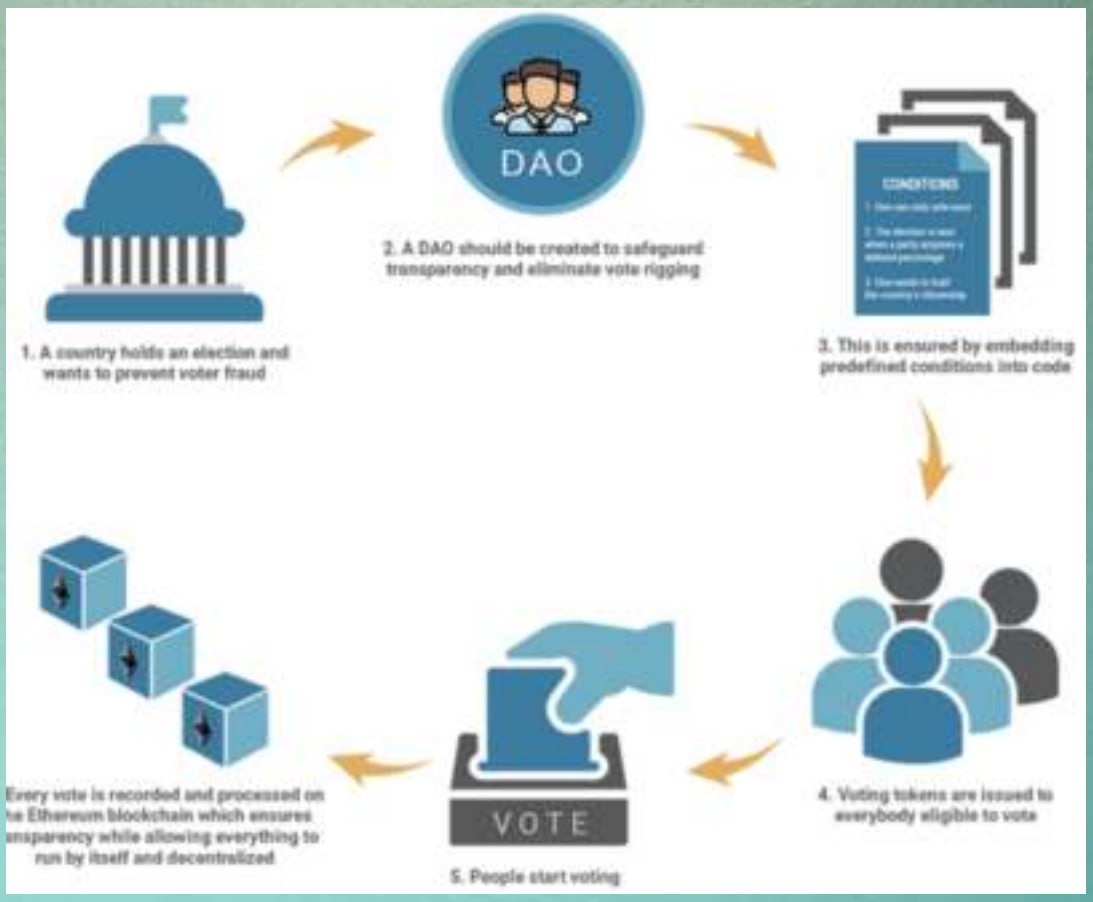
The presiding officer (PO)verifies his unique identity number and his biometric authentication then the biometric are verified and the permission is granted using the SHA-256 hashing algorithm. The next block is generated by the generating a new random number, associating that with the hash of the previous block.

2.DIFFICULTY AND NONCE VALUE

For every proceeding block, the hash of the previous block hash code, unique id of the PO and the random number generated will be derived. e.g. the length of random number for block number 20 will be 20.Each casted vote will serve as a new transaction.

3. SEALING OF THE BLOCKS

Once the polling is complete, the next step is to seal the blocks to ensure that the block come even more secure and adding security is included.The data of the block (i.e. the entire result) will be hashed using the SHA-256 algorithm. This is done by concatenating the results inside the block and hashing them in pairs the block is hashed based on the hashed contents of the block. Another system generated random number can be added in the hashing to make it more secure.The sealing process is use the hashing algorithm called SHA-256.



This is done by concatenating the results inside the block and hashing them in pairs the block is hashed based on the hashed contents of the block. Another system generatedrandom number can be added in the hashing to make it more secure.
The sealing process is use the hashing algorithm called SHA-256.

Reference:https://researchgate.net

DATA DEDUPLICATION

-Vishal M Nair

"Without a systematic way to start and keep data clean, Bad data will happen" -Donato Diorio

Data in general refers to a set of facts which when processed becomes Information. This information is very crucial for analyzing a situation. Despite its importance, data redundancy is very common nowadays. Data Redundancy is a major problem which can lead to data inconsistency and data corruption which in turn makes the database inefficient.

A technique called data deduplication can be implemented to avoid the above situation. This technique ensures that the database consists of unique data, in this way it helps in reducing the transfer and storage of redundant data, which optimizes network bandwidth and storage capacity.

The main techniques in Data deduplication are:

- File-level data deduplication strategy
- Block-level data deduplication technology
- Byte-level data deduplication

Original "blocks" of data

Unique blocks after data reduction

The major application of data deduplication lies in large organizations dealing with highly redundant operations such as constant copying and storing of data for future reference or recovery purpose. This technique can be employed for backup and disaster recovery solution. Another major application involves the reduction of disk space upto 80% thereby increasing the interval of changing disks repeatedly.

With the rapid growth in information and network technology there is a rapid upsurge in the size of the data center, energy consumption in IT spending in the increasing proportion of data deduplication to optimize storage system can greatly curtail the amount of data, thereby reducing energy consumption and reduce heat emissions. Data deduplication can reduce the number of disks used in the operation to reduce disk energy consumption costs. Remove duplicate data for the large data center information technology system backup system a comprehensive, mature, safe and reliable, More green save the backup data storage technology solutions, has a very high value and great academic value, with very high application value and important academic research value.

Reference:https://wikipedia.org

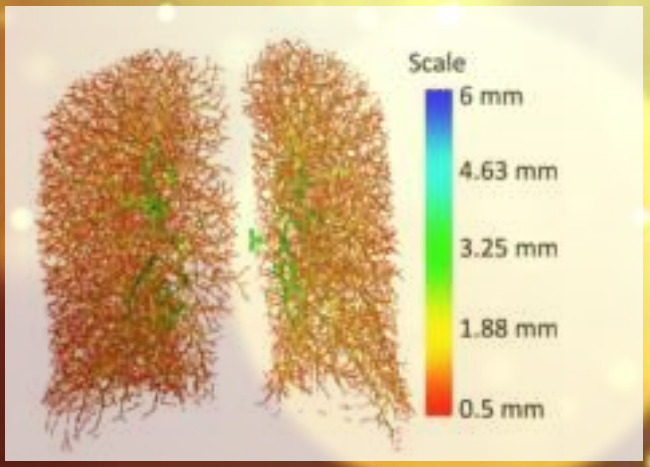
PULMONARY ARTERY VEIN CLASSIFICATION IN CT IMAGES USING DEEP LEARNING

-Vivek C Varghese

"It has become appallingly obvious that our technology has exceeded our humanity"
- Albert Einstien

Recent studies show that pulmonary vascular diseases may specifically affect arteries or veins through different physiologic mechanisms. To detect changes in the two vascular trees, physicians manually analyze the chest computed tomography (CT) image of the patients in search of abnormalities. This process is time consuming, difficult to standardize, and thus not feasible for large clinical studies or useful in real-world clinical decision making. Therefore, automatic separation or classification of arteries and veins in CT images is becoming of great interest, as it may help physicians to accurately diagnose pathological conditions.

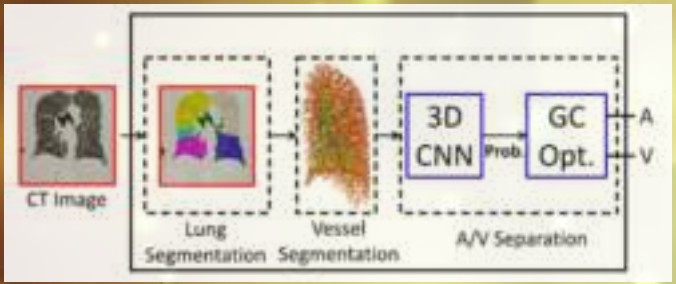
Classification of lung vessels into artery/vein (A/V) may be of great help for physicians to accurately diagnose pulmonary diseases that may affect either the arterial or the venous trees in specific ways.

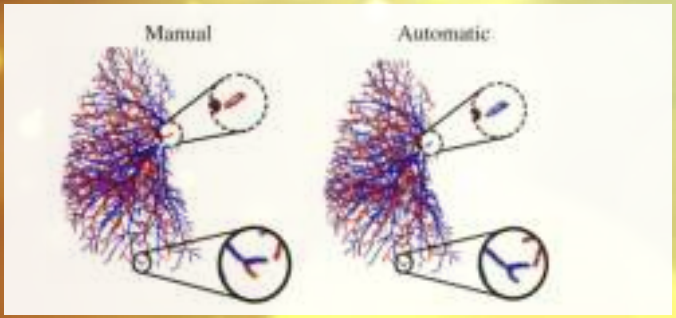


In the last decades, computed tomography (CT) has become the most common imaging technique for diagnosis and treatment of lung diseases. Modern CT scanners combined with modern imaging techniques allow for the use of low radiation doses to automatically identify and extract pulmonary structures,such as vessels and bronchi with high accuracy. Recognition and discrimination of pulmonary arteries and veins represent one of the most challenging problems.

Classification of lung vessels into artery/vein (A/V) may be of great help for physicians to accurately diagnose pulmonary diseases that may affect either the arterial or the venous trees in specific ways.

Throughout the years, several methods have been proposed to either enhance or segment vessels from lung CT images. Although these methods are not able to separate arteries and veins, they are often used as a starting point for most of the A/V segmentation algorithms available. Moreover, most methods try to utilize A/V local information, like seed points automatically defined in the lung region or the proximity of arteries to bronchi , to separate the two vascular trees. The idea of exploiting the proximity of airways to arteries to classify vessels was used in several other methods available. However, the risk of mislabeling vessels using this method increases with decreasing vessel radius, as on CT images small vessels are better visible than bronchi of similar size.





In this method through Deep learning with Image classification, it formulates a fully automatic algorithm that combines a convolution neural network (CNN) approach with a graph-cuts strategy ,to classify vessels into arteries and veins on chest CT images. Small 3D patches are extracted from the CT image around the vessel candidates, defined using the scale-space particles approach described in and, and used to train the neural network.

Referred From :https://ieeexplore.ieee.org/

CHALLENGES OF INDUSTRY 4.0

Are we prepared for it...?

-Vishnu Ramesh

"There was a time when people felt the internet was another world,but now realize it's a tool we use in this world."

- Tim Berners Lee

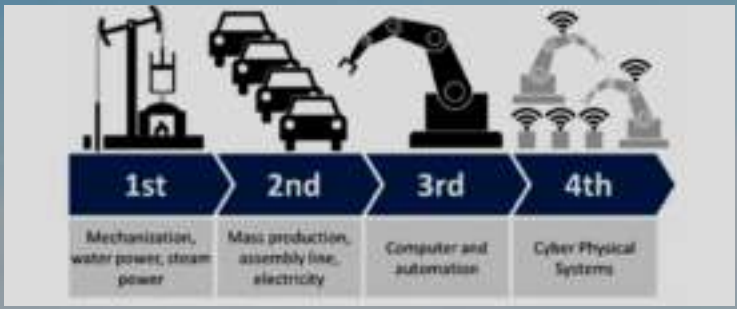
The terms "industry 4.0" and "fourth industrial revolution" are often used interchangeably, "industry 4.0" is simply the working of the industry machinery together with the technology of IoT or otherwise known as IIoT(Industry internet of things).The term Industry 4.0 was first publicly introduced in 2011. The main objective was to reduce the human errors caused and speed up the industrial process with much more efficiency.

This technology consists of many network elements and is divided into device layer, middle layer and application layer. However, security is also a big challenge for IIoT environment. There are numerous cases reported where IoT devices has been attacked , thus entering into the network which opens up a wide variety of opportunities to tamper with other hardware making potential performance loses and confidential information being stolen. This is only a small scenario. A cyber attack on a manufacturing facility would be damaging, but it's not the most prominent threat to a network's operational continuity.

Although rare, a cyber attack could be devastating to your organization's reputation and bottom line. A potential
cyber attack can be as threatening as a ransomware attack where the attacker could withhold the confidential data which jeopardize the whole production process. These incidents can be detected and prevented with the right plan in place. First, your organization must have an up-to-date inventory of its digital assets to develop an understanding of its network to pinpoint any existing vulnerabilities within it. After which the industry must have authentication processes in place to guard your digital and physical assets

Organization needs to have the ability to detect anomalous activity. The most effective way to achieve this is to deploy a continuous monitoring solution. Finally, your organization should have a contingency plan in place to quickly respond and recover from a potential cyberattack.

Other than cyber attacks there can be other challenges that are not directly technical, as Industry 4.0 continues to change the way manufacturers do business. Here are just some of the many challenges they face when looking to tackle 4.0 considering new business models/strategies,reorganizing processes to leverage better outcomes, understanding



Other than cyber attacks there can be other challenges that are not directly technical, as Industry 4.0 continues to change the way manufacturers do business. Here are just some of the many challenges they face when looking to tackle 4.0 considering new business models/strategies,reorganizing processes to leverage better outcomes, understanding business cases, conducting successful pilots, helping others in the company understand where action is needed, changing management (a tough option that is often not considered), taking a look at company culture, connecting departments, recruiting, developing and maintaining new talent.

According to McKinsey, "Industry 4.0 disrupts the value chain and requires companies to rethink the way they do business. They need to drive the digital transformation of their business to succeed in the new environment."

Referred From :https://sciencedirect.com