NEW HORIZON COLLEGE OF ENGINEERING

DEPARTMENT OF INFORMATION SCIENCE & ENGINEERING

2016 - 17

INFOTECH PATRIKA - A Half- Year Publication

VOLUME 2 ISSUE 1

Foreword

Dr. Jitendranath Mungara, Professor & Head - Information science and Engineering

Welcome. This is my first opportunity to speak with you through the departmental technical magazine InfoTech Patrika. It has been an interesting and busy semester for members of the Department. It gives me great opportunity to present this issue of Technical Magazine exclusively by Information Science & Engineering department.

This magazine is one of the ways in which we can disseminate current trends in technology, research & developments.

I would like to request for your active collaboration over the coming months in the development of a shared vision for the department.

I would like to thank all my colleagues for their tireless efforts to help the department progress at a very steady pace

Dr. Jitendranath Mungara
Prof. & Head – ISE

About the Department:

Information science and Engineering department focuses on current Information Technology trends, and Domain Specific Applications. The program facilitates the evolution of skills in students to help them attain a higher degree of knowledge, global competency and excellence, for the betterment of the society. The Department of Information science and Engineering at NHCE was established in the year of 2001 and offers graduate and PhD programs. The four year B.E degree equip the students to meet day- to- day Technological advancements of the ever dynamic IT field through adept training on various subjects of curriculum of Information Science and engineering and beyond. The department offers B.E program through autonomous scheme from the year 2015. The department has a total intake of over 380 students with a very good team of highly qualified and talented faculty members including Professors, Associate Professors and Assistant Professors.

Information Science and Engineering course at New Horizon College of Engineering is designed to meet industry standard and cope up with the emerging technology. There is a great emphasis on holistic learning to help the students to make significant contributions at all levels and to meet the expectations of stakeholders. The department is well known for its research excellence in various competitive areas of Information Science. Students are made to involve vigorously in research activities. The department provides industry collaborated courses for the students.

LG Taunts the World with Rollable Displays

There's an alternate universe in which every computer screen and smartphone can be rolled like a newspaper, and TVs come in long cardboard tubes. Unfortunately, that's not our reality, despite a new prototype by LG.

The company announced today an ultra-slim, flexible 18-inch OLED display that rolls into a cylinder without being damaged. However, it's just a prototype, and probably not going on sale any time soon.

Coming in sizes from 65-inches to 139-inches, LG is gunning for both the high-end OLED TV consumer market, as well as commercial signage.

The largest monitor in LG's CES 2016 array is an 89-inch display with a 58:9 aspect ratio, which they suggest for use in airports. Secondary, non-sanctioned uses may include visualizing the entire linear Super Mario World.

- Ms. Sai Nayana Sahishna

5 Artificial intelligence trends that will dominate 2018

2017 saw an explosion of machine learning in production use, with even deep learning and artificial intelligence (AI) being leveraged for practical applications.

"Basic analytics are out; machine learning (and beyond) are in," says Kenneth Sanford, U.S. lead analytics architect for collaborative data science platform Dataiku, as he looks back on 2017.

Sanford says practical applications of machine learning, deep learning, and AI are "everywhere and out in the open these days," pointing to the "super billboards" in London's Piccadilly Circus that leverage hidden cameras gathering data on foot and road traffic (including the make and model of passing cars) to deliver targeted advertisements.

So where will these frameworks and tools take us in 2018? We spoke with a number of IT leaders and industry experts about what to expect in the coming year.

Enterprises will operationalize AI

AI is already here, whether we recognize it or not.

"Many organizations are using AI already, but they may not refer to it as 'AI,'" says Scott Gnau, CTO of Horton works. "For example, any organization using a chatbot feature to engage with customers is using artificial intelligence."

But many of the deployments leveraging AI technologies and tools have been small-scale. Expect organizations to ramp up in a big way in 2018.

"Enterprises have spent the past few years educating themselves on various AI frameworks and tools," says Nima Negahban, CTO and co-founder of Kinetica, a specialist in GPU-accelerated databases for high-performance analytics. "But as AI goes mainstream, it will move beyond small-scale experiments to being automated and operationalized. As enterprises move forward with operationalizing AI, they will look for products and tools to automate, manage, and streamline the entire machine learning and deep learning life cycle."

Negahban predicts 2018 will see an increase in investments in AI life cycle management, and technologies that house the data and supervise the process will mature.

- Prof. Swathi B

Bias in training data sets will continue to trouble AI

Reltio's Chen isn't alone in his conviction that enterprises need to get their data in order. Tomer Shiran, CEO and co-founder of analytics startup Dremio, a driving force behind the open source Apache Arrow project, believes a debate about data sets will take center stage in 2018.

"Everywhere you turn, companies are adding AI to their products to make them smarter, more efficient, and even autonomous," Shiran says. "In 2017, we heard competing arguments for whether AI would create jobs or eliminate them, with some even proposing the end of the human race. What has started to emerge as a key part of the conversation is how training data sets shape the behavior of these models."

It turns out, Shiran says, that models are only as good as the training data they use, and developing a representative, effective training data set is very challenging.

"As a trivial example, consider the example tweeted by a Facebook engineer of a soap dispenser that works for white people but not those with darker skin," Shiran says. "Humans are hopelessly biased, and the question for AI will become whether we can do better in terms of bias or will we do worse. This debate will center around data ownership — what data we own about ourselves, and the companies like Google, Facebook, Amazon, Uber, etc. — who have amassed enormous data sets that will feed our models."

- Dr.Vishwanath Y

The Future of a Modern Workplace: "HumAIns" are Coming.

According to the latest Accenture report on the future of workplace collaborations, businesses that manage to balance human ingenuity with machine intelligence will be successful. Released just ahead of the World Economic Forum 2018 in Davos, Accenture's AI report takes a positive stance on the need to grow investments into AI and Human-Machine collaboration over the next five years.

Recommended Read: Will Artificial Intelligence Exceed Human Performance in Marketing and Sales by 2025?

So, what can we expect at such a 'modern' workplace?

Let's call the new professionals that you could be hiring and collaborating with, by 2022-

'The HumAIns"

Who are HumAIns?

HumAIns are machine-driven, human-centric workplace assistants that demonstrate the highest ability to deliver "Live" customer experiences. No biases, even with millions of insights to deal with from historical data, the HumAIns will boost three aspects of any business

Increase Revenues

Maximize Profits

Guarantee Human Employment with Respectable Salaries

According to the Accenture report on AI-Human collaboration, the HumAIns could boost revenues by 38% and grow employment opportunities by 10 percent between 2018 and 2022.

The Convergence of AI and the Human Race within Economic Circles

The HumAIns (no more a hypothesis), would propel the adoption of technology across industries. The economic fields of study for AI would expand further into these five categories-

Deep Learning: Synchronous group of machines running on powerful algorithms led by a human expert.

Robotization: Machines take over humans, freeing the creative minds to focus on refining business strategies.

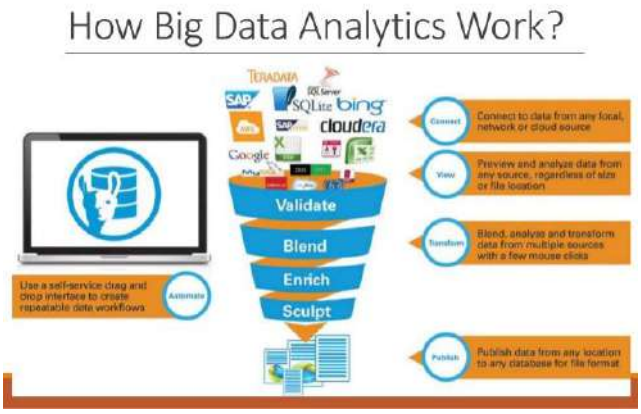
Dematerialization: Voice search, intelligent assistants, automatic streaming tools, and contactless payment solutions.

- **Mr. Ijaz Nijami**

BIG DATA ANALYTICS:

Introduction Change is the new norm for the global healthcare sector. In fact, digitization of health and patient data is undergoing a dramatic and fundamental shift in the clinical, operating and business models and generally in the world of economy for the foreseeable future. Tis shift is being spurred by aging populations and lifestyle changes; the proliferation of software applications and mobile devices; innovative treatments; heightened focus on care quality and value; and evidence-based medicine as opposed to subjective clinical decisions—all of which are leading to offer significant opportunities for supporting clinical decision, improving healthcare delivery, management and policy making, shrivelling disease, monitoring adverse events, and optimizing treatment for diseases affecting multiple organ systems .

As noted above, big data analytics in healthcare carries many benefits, promises and presents great potential for transforming healthcare, yet it raises manifold barriers and challenges. Indeed, the concerns over the big healthcare data security and privacy are increased year-by-year. Additionally, healthcare organizations found that a reactive, bottom-up, technology-centric approach to determining security and privacy requirements is not adequate to protect the organization and its patients .



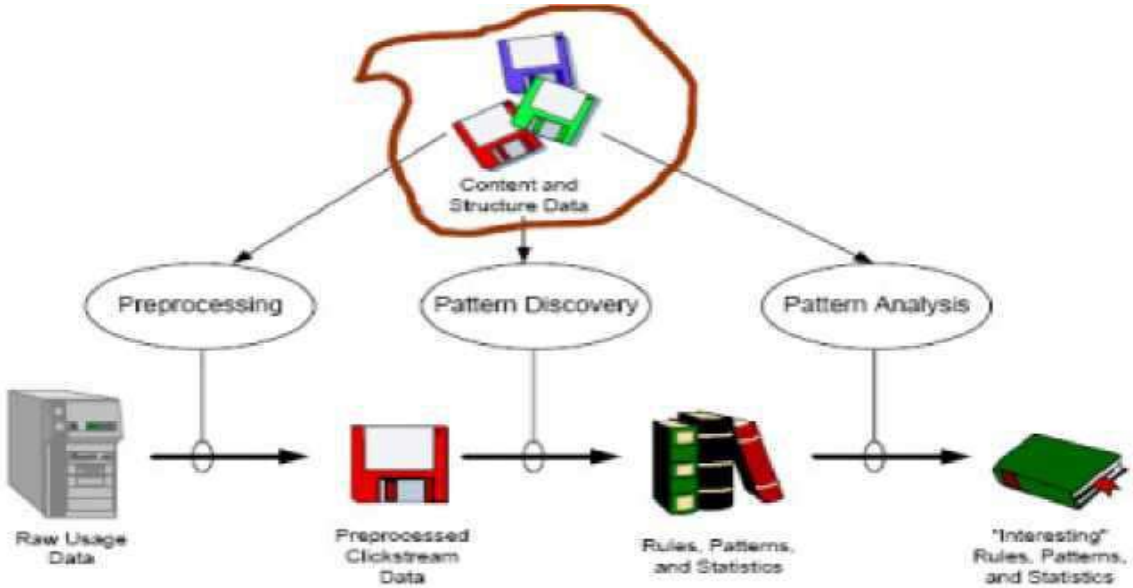
Motivated thus, new information systems and approaches are needed to prevent breaches of sensitive information and other types of security incidents so as to make effective use of the big healthcare data. In this paper, we discuss some interesting related works and present risks to the big health data security as well as some newer technologies to redress these risks.

Web Mining Using Cloud Computing Technology

Introduction: Web mining includes how to extract the useful information from the web and gain knowledge using data mining techniques. Here so many resources and techniques are available i.e. web content mining, web structure mining, web usage mining and access through the web servers. Web mining techniques (specially web usage mining techniques) and applications are much needed in cloud computing. The implementation of these techniques through cloud computing will allow users to retrieve relevant and meaningful data from virtually integrated data warehouse which reduces cost and infrastructure.

Web mining - is the application of data mining techniques to discover patterns from the Web. According to analysis targets, web mining can be divided into three different types, which are Web usage mining, Web content mining and Web structure mining. Web usage mining is the process of extracting useful information from server logs e.g. use Web usage mining is the process of finding out what users are looking for on the Internet using cloud computing . Some users might be looking at only textual data, whereas some others might be interested in multimedia data. Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from Web data in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site.

Web usage mining itself can be classified further depending on the kind of usage data considered. Several data mining methods are used to discover the hidden information in the Web. However, Web mining does not only mean applying data mining techniques to the data



Fig(1): Concept of web usage mining

stored in the Web. The algorithms have to be modified such that they better suit the demands of the Web. The web usage mining generally includes the following several steps: data collection, data pretreatment, and knowledge discovery and pattern analysis.

Data collection: Web usage mining is the process of extracting useful information from server logs e.g. use Web usage mining is the process of finding out what users are

looking for on the Internet. Some users might be looking at only textual data, whereas some others might be interested in multimedia data. Web Usage Mining is the application of data mining techniques to discover interesting usage patterns from Web data in order to understand and better serve the needs of Web-based applications. Usage data captures the identity or origin of Web users along with their browsing behavior at a Web site.

Web usage mining itself can be classified further depending on the kind of usage data considered: Web Server Data: The user logs are collected by the Web server. Typical data includes IP address, page reference and access time. Application Server Data: Commercial application servers have significant features to enable e-commerce applications to be built on top of them with little effort. A key feature is the ability to track various kinds of business events and log them in application server logs. Application Level Data: New kinds of events can be defined in an application, and logging can be turned on for them thus generating histories of these specially defined events. It must be noted, however, that many end applications require a combination of one or more of the techniques applied in the categories above in the figure(1).

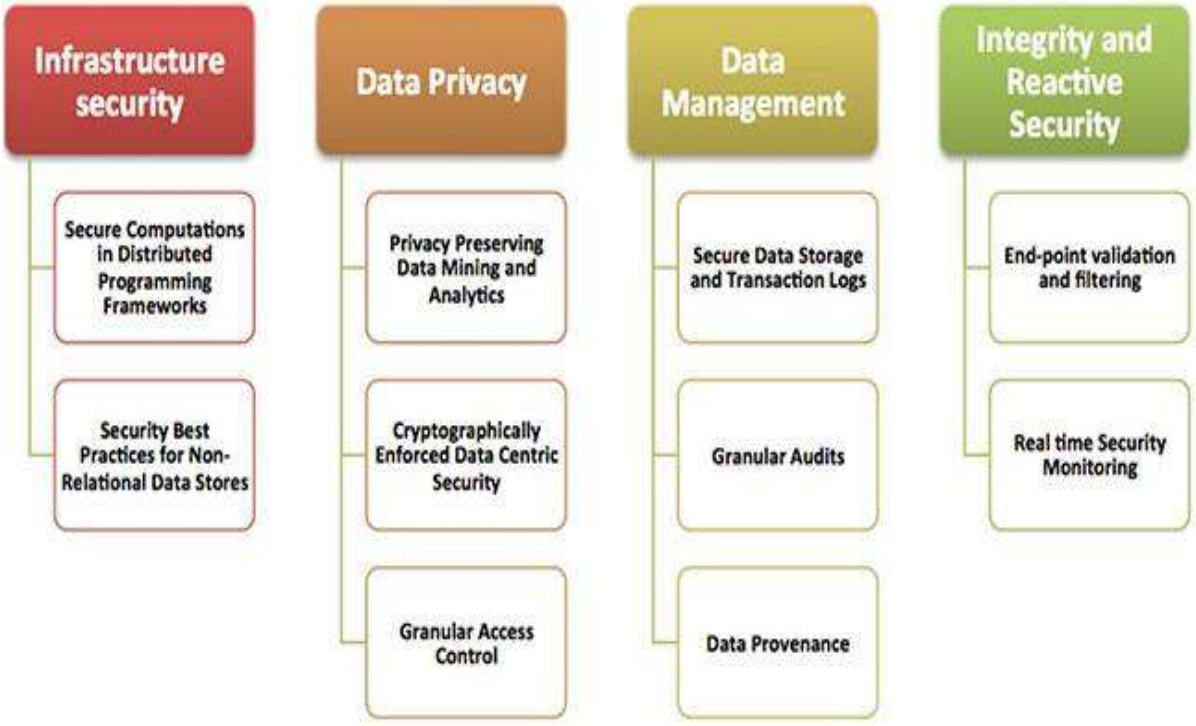
Data preprocessing: Web Usage Mining in cloud computing is one of the categories of data mining technique that identifies usage patterns of the web data, so as to perceive and better serve the requirements of the web applications. The working of WUM involves three steps - preprocessing, pattern discovery and analysis. The first step in WUM - Preprocessing of data is an essential activity which will help to improve the quality of the data and successively the mining results. This research paper studies and presents several data preparation techniques of access stream even before the mining process can be started and these are used to improve the performance of the data preprocessing to identify the unique sessions and unique users in cloud computing . The methods proposed will help to discover meaningful pattern and relationships from the access stream of the user and these are proved to be valid and useful by various research tests.

- Prof. Rajeswari S

Security and Privacy in Cloud Computing

Introduction: The cloud has fundamentally changed the landscape of computing, storage, and communication infrastructures and services. With strong interest and investment from industry and government, the cloud is being increasingly patronized by both organizations and individuals. From the cloud provider's perspective, cloud computing's main benefits include resource consolidation, uniform management, and cost-effective operation; for the cloud user, benefits include on-demand capacity, low cost of ownership, and flexible pricing. However, the features that bring such benefits, such as sharing and consolidation, also introduce potential security and privacy problems. Security and privacy issues resulting from the illegal and unethical use of information, and causing disclosure of confidential information, can significantly hinder user acceptance of cloud-based services. Recent surveys support this observation, indicating that security and privacy concerns prevent many customers from adopting cloud computing services and platforms. In response to such concerns, significant research and development efforts in both industry and academia have

sought to improve the cloud's security and privacy. Here I give a quick (and incomplete) overview of new challenges, opportunities, and solutions in this area, with the purpose of stimulating more in-depth and extensive discussion on related problems in upcoming issues of this magazine. Identifying New Threats and Vulnerabilities: An essential task in cloud security and privacy research is to identify new threats and vulnerabilities that are specific to cloud platforms and services. Several recent reports have explored such vulnerabilities. For example, in 2009, researchers from the University of California, San Diego, and the Massachusetts Institute of Technology demonstrated leakage attacks against Amazon's Elastic Compute Cloud (EC2) virtual machines (VMs). More specifically, the researchers showed that it's possible to probe and infer the overall placement of VMs in the EC2 infrastructure. Furthermore, an attacker can launch a malicious EC2 instance and then determine whether that instance is physically colocated with a targeted (victim) instance. When the attacker's instance is successfully colocated with the victim, it can launch a side-channel attack by monitoring the status of shared physical resources such as level-1 and level-2 caches, and thus infer the victim's computation and I/O activities. Classification of top 10 challenges in Cloud computing in shown below in fig(1).



Fig(1): Classification Of Top 10 Challenges In Cloud Computing.

A follow-up study showed that it's possible to extract private keys via the cross-VM side channel in a lab environment. In another study, researchers from the College of William and Mary reported that side-channel attacks aren't just a potential risk, but a realistic threat. They created a covert channel via another shared resource (the memory bus) that had a level of reliability and throughput of more than 100 bps in both lab and EC2 environments.

These risks represent a small subset of known cloud-specific vulnerabilities and threats. However, they motivate us to think further about new adversary models, trust relations, and risk factors relative to cloud computing stakeholders. In the examples, the

cloud provider isn't trusted because of its resource sharing and VM consolidation practices. Hence, the cloud provider doesn't provide a desirable level of isolation and protection between tenants in the cloud, allowing them to attack each other.

Protecting Virtual Infrastructures: *Virtual infrastructures* are infrastructure-level (virtual) entities, such as VMs and virtual networks, created in the cloud on behalf of users. Side-channel attacks target these virtual infrastructures. Researchers have proposed several solutions to defend against cross-VM side-channel attacks. Duppel, for example, aims to disrupt cache-based side channels. In this self-defensive approach, the target VM's guest operating system injects cache access noise (that is, flushes) so the collocated attack VM can't infer cache access patterns. This solution doesn't require modifying the underlying hypervisor or cloud platform. To defend against memory bus-based side channels, a simple and practical approach is to prevent a VM from locking the memory bus and let the hypervisor emulate the execution of atomic instructions that would otherwise require memory bus locking.

Other attacks against virtual infrastructures include malware attacks against tenant VMs. The cloud presents a new opportunity to defend against these attacks. More specifically, the cloud provides a uniform and tamper-resistant platform to deploy system monitoring and antimalware functions. The uniformity is reflected by the cloud provider's consistent installation, configuration, and update of antimalware services for all hosted tenants. It's tamper resistant because monitoring and detection of malware attacks can be performed from outside the hosted VMs, either by the underlying hypervisor or by the more privileged management domain (for example, Domain 0 of Xen). In CloudAV, a production-quality system that reflects the antivirus-as-a-service idea, a group of in-cloud antivirus engines analyzes suspicious files submitted by agents running in client machines (including VMs) and collectively detects malware in them. VMwatcher, a virtualization- based malware-monitoring and detection system, moves commodity, off-the-shelf antimalware software from the inside to the outside of each tenant VM.7 This way, the antimalware software is out of the malware's reach, preventing the malware from detecting, disabling, or tampering with it. Malware targeting a tenant VM—at either the user or kernel level—can be detected and prevented using such an "out-of-the-box" antimalware service.

A networked virtual infrastructure can consist of multiple VMs connected by a virtual network. With the rapid advances in software-defined networking (SDN), the cloud increasingly supports such networked virtual infrastructures. SDN decouples the control and data-forwarding functions of a physical networked infrastructure, such as a datacenter network. The SDN control plane performs control functions such as routing, naming, and firewall policy enforcement, and the SDN data plane follows the control plane's decisions to forward packets belonging to different flows. Such decoupling makes it easy to optimize the control and data planes without them affecting each other. However, the SDN paradigm raises security issues. Researchers have reported that it's possible to launch attacks against the SDN architecture, incurring excessive workload and resource consumption to both the control and the data plane. Although researchers are developing defenses against such

attacks, we need more generic, scalable solutions that make the SDN architecture secure, robust, and scalable, which would support virtual infrastructure hosting in the cloud.

Protecting Outsourced Computation and Services: Many organizations have been increasingly outsourcing services and computation jobs to the cloud. A client that outsources a computation job must verify the correctness of the result returned from the cloud, without incurring significant overhead at its local infrastructure—the extreme being to execute the job locally, which would nullify the benefit of outsourced job execution. Such verifiability is important to achieving cloud service trustworthiness and hence has become a topic of active research. Encouragingly, researchers have in recent years developed techniques and real systems to bring the vision of a "verifiable cloud service" closer to reality. For example, the Pantry system composes and outsources proof-based verifiable computation with untrusted storage. It achieves theoretically sound verifiability of computation for realistic cloud applications, such as MapReduce jobs and simple MySQL queries.

In addition to computation outsourcing, the cloud can support network service/function outsourcing. Example network functions include traffic filtering, transcoding, firewall policy enforcement, and network-level intrusion detection. Seyed Ka-veh Fayazbakhsh and his colleagues noted that, similar to computation outsourcing, a major challenge is to verify (at end points of network connections) that the "middle boxes" in the cloud correctly execute outsourced network functions with satisfactory performance.10 They also proposed a framework for verifiable network function outsourcing (vNFO) that aims to achieve verifiability, efficiency, and accountability of outsourced network functions. Such a framework will pave the way for deploying trusted network middle boxes, in addition to end points (that is, VMs), in the cloud, enriching the cloud ecosystem.

- Ms. Payal Jain

EDITORIAL BOARD: