**Enterprise data breach: causes, challenges, prevention, and future directions**

**Data breach:** The deliberate or unintentional release of private data to unapproved parties is known as a data breach.

What the paper work on?

* data leak threats
* recent data leak incidents
* various state-of-the-art prevention and detection techniques
* new challenges
* and promising solutions and exciting opportunities.

Leaked information can encompass various types such as:

* employee/customer data
* intellectual property
* and medical records.

This review paper discusses enterprise data leak threats, identifies key detection and prevention techniques, and highlights limitations in current approaches. It examines recent incidents and lessons learned, focuses on big data challenges, and introduces a privacy-preserving detection system as a case study. It also encourages future research in this area.

Consequences of it:

Sensitive information loss can cause significant damage to an organization's reputation, financial stability, and long-term stability.

IBM statistics:

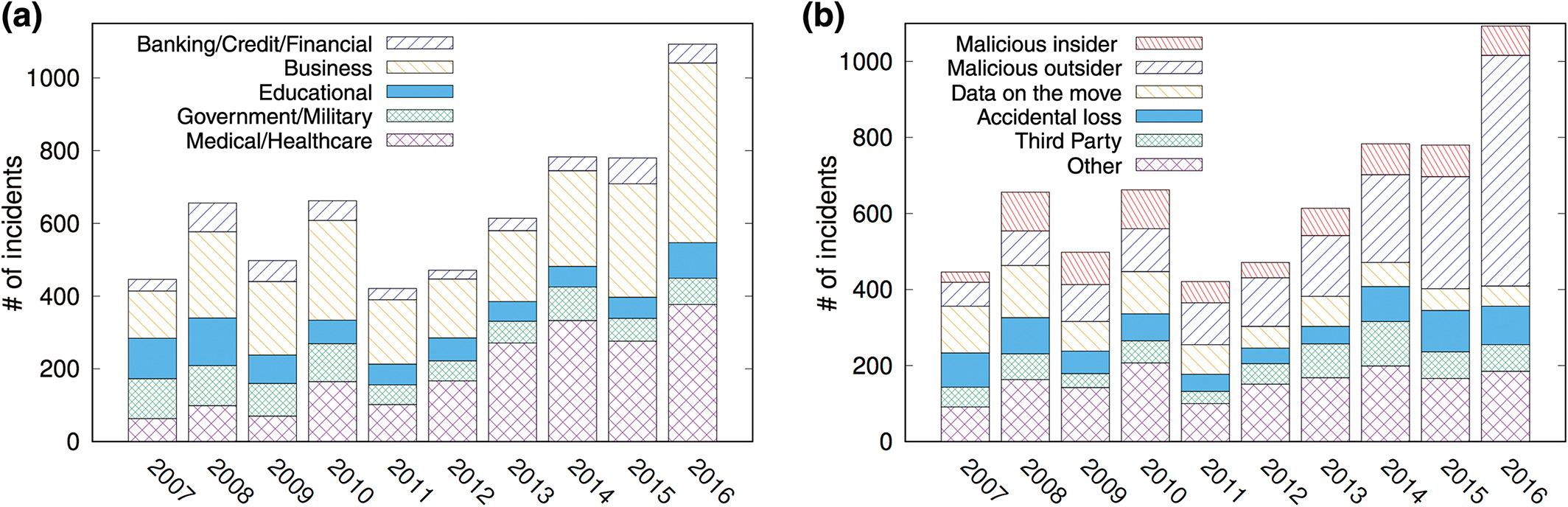
According to IBM's 2016 Cost of Data Breach Study,[1](https://wires.onlinelibrary.wiley.com/doi/full/10.1002/widm.1211#widm1211-bib-0001) the average consolidated cost of a data breach has reached $4 million

Events:

Target Corporation experienced $248 million in losses due to a cybercrime in 2013, stealing 40 million payment cards and 70 million customer information.

In 2014, Yahoo reported a data breach that allegedly involved the theft of at least 500 million accounts.

Data leaks have been increasing in the past five years, with business and medical/healthcare leaks taking the majority of incidents. In 2016, business data breaches accounted for 45.2% of overall breaches, followed by medical/healthcare at 34.5%. Malicious outsiders caused around 55% of breaches in 2016. The Identity Theft Resource Center (ITRC) reports that insider threats are the leading cause of enterprise data leak threats, with over 40% of breaches occurring within a company.



Reason:

Data leakage can occur through intentional or accidental breaches, such as data theft or sabotage by insiders or accidental disclosure of sensitive information by employees and partners.

Intel Security's study reveals that internal employees cause 43% of corporate data leakage, with half accidental. Motivations include corporate espionage, employer grievance, or financial reward.

Accidental leaks mainly result from unintentional activities due to poor business process such as failure to apply appropriate preventative technologies and security policies, or employee oversight.

The purposes of data leak prevention and detection (DLPD) systems:

Data leak prevention and detection systems identify, monitor, and prevent unintentional or deliberate exposure of sensitive information in an enterprise environment using various technical approaches.

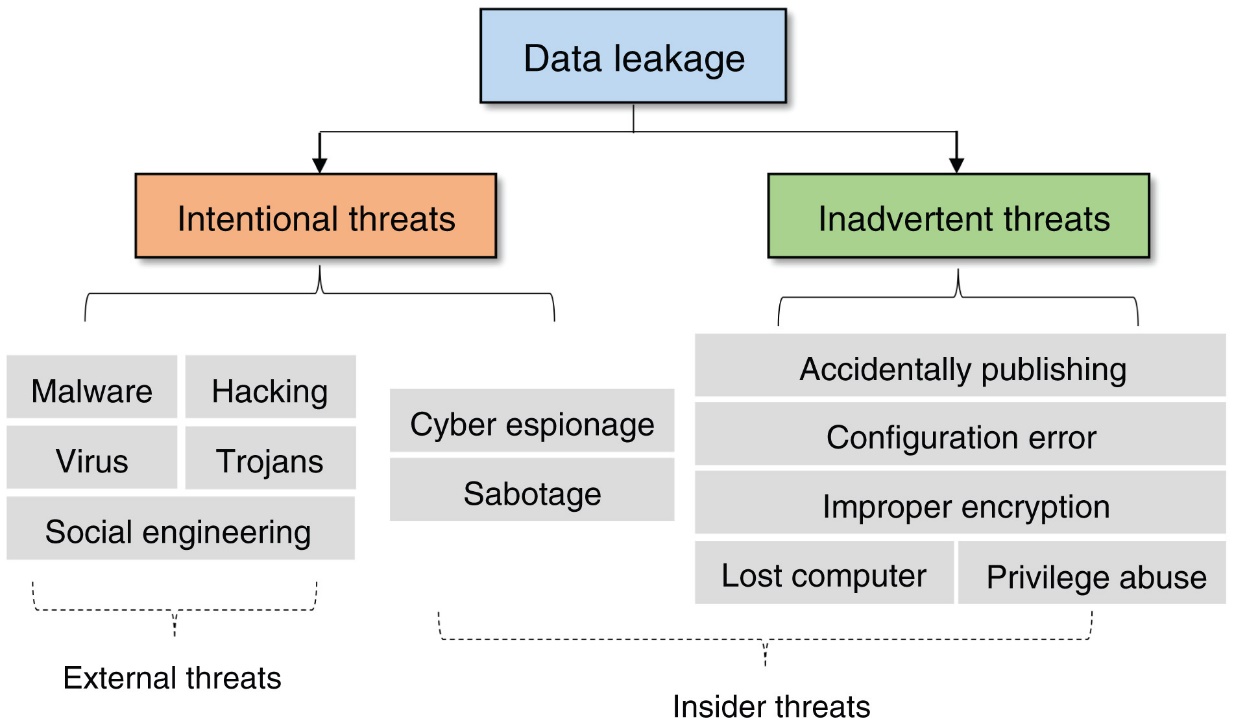
Such as Pioneering works model database access behaviors to identify intruders and detect data breaches, while basic security measures like data use policies and traffic inspection safeguard sensitive information.

challenges:

In the era of big data, companies face significant security challenges due to the increasing risk of data leakage. As data becomes a critical component of an enterprise, managing and analyzing large amounts of data provides a competitive advantage. However, this also puts sensitive data at risk of loss or theft. Modern communication channels, such as cloud file sharing, email, instant messaging, FTP, and laptop theft, increase the potential for data leakage.

Threat:

Data leak threats can be classified based on their causes, either intentionally or inadvertently leaking sensitive information. External data breaches are typically caused by hacker break-ins, malware, viruses, and social engineering. Internal data leakage can be caused by deliberate actions or inadvertent mistakes. Hauer's comprehensive criteria for analyzing 1259 data leakage incidents revealed that over 60% were caused by insiders, emphasizing the importance of both technological and nontechnological measures in preventing data breaches.



Conclusion:

This paper reviews data leak threats and key techniques for Deep Learning-based Data Leak Prevention (DLPD), highlighting challenges in the big data era. It highlights promising research directions for reducing data breach risks in enterprise environments, including cloud service data leak detection and deep learning-based anomaly detection for insider threats.