Artificial Intelligence-Powered Intellectual Property Detection in Cisco Secure Access

Contents

Introduction

Generative Artificial Intelligence (GenAI) tools, such as publicly available Large Language Models (LLMs), have become popular in enterprises worldwide. Similarly, the use of cloud storage in enterprises has seen persistent growth. The use of public LLMs and cloud storage poses problems for enterprises with respect to Intellectual Property (IP). For example, software engineers may use LLMs to debug, test, or optimize code that they have written, in the process leaking intellectual property. Similarly, patent lawyers may store sensitive patent application information in the cloud without installing proper protections.

How can enterprises detect intellectual property movement to public LLMs or inappropriate exposure via cloud storage? Cisco® Secure Access, which is Cisco's complete Security Service Edge (SSE) solution, has a multimode, cloud-delivered Data Loss Prevention (DLP) engine that has been enhanced with Artificial Intelligence (AI) techniques to detect intellectual property in documents. Thus, Cisco Secure Access can monitor intellectual property movement to external resources such as public LLMs and identify unprotected intellectual property in cloud storage.

More on the problem

Some enterprises want their engineers to refrain from using public LLMs for code development. Policies of this nature seek to prevent intellectual property leakage to LLMs and beyond. Publicly available news stories indicate that the risks of sensitive material leaked via LLMs are not just possible in theory but have been observed in real life.[i]

Other enterprises use cloud storage pervasively in their business processes to house intellectual property, such as patent applications and partnership agreements. Sometimes, such documents are in improperly marked and inadequately protected locations. For example, a closely guarded patent application may have been inadvertently dropped into a publicly accessible cloud folder, exposing the contents to miscreants and competitors.

Cisco Secure Access' role and DLP

Cisco Secure Access[ii] is a SSE solution that includes capabilities such as Zero Trust Network Access (ZTNA), Secure Web Gateway (SWG), Cloud Access Security Broker (CASB), Firewall-as-a-Service (FWaaS), Domain Name System (DNS) security, and DLP.

Cisco Secure Access' AI-powered DLP can detect code in many programming languages, including C, C#, C++, Cobol, CSS, Dart, Go, Lang, Java, JavaScript, Kotlin, NoSQL (MongoDB, DynamoDB, Redis), Perl, PHP, PL/SQL, Python, R, Ruby, Rust, Scala, Swift, SQL, and TypeScript.

The AI-powered DLP can also detect many other types of intellectual property documents, such as patent applications, partnership agreements, merger and acquisition agreements, non-disclosure agreements, and employment agreements.

Intellectual property detection with AI

Intellectual property detection in Cisco Secure Access DLP uses LLMs to classify the document into one of the document types discussed above.

Historically, such detection in DLP engines was done using regular expressions and keyword matching. These traditional methods pose specific challenges:

1. Difficulty in using context: DLP engines should use a significant portion of the context of a document. In particular, they should use the words and sentences surrounding a tell-tale keyword to determine if the document meets the configured matching criteria (say, for intellectual property). The use of such context is problematic when the only available technologies are regular expressions and keyword matching, as these technologies were not created for systematic exploration of context around a keyword.
2. Prevalence of false positives: Regular expressions and keyword matching often lead to false positives. For example, matching patent numbers may return documents with math problems and transaction identifiers.
3. Elongated process: To contain the false positives, the regular expression and keyword match creators must endure long test cycles for their filters to ensure an acceptable amount of accuracy without too many false alarms.
4. Supporting new intellectual property types: New intellectual property types, such as a previously unsupported programming language, require creating a fresh set of regular expressions and keyword matches. As a result, detecting new types of intellectual property lags behind business priorities.
5. Supporting non-English natural languages: Many DLP engines support documents in the English language with acceptable performance. However, their support of non-English documents is severely curtailed due to the challenges mentioned above. Consequently, detection again lags behind business priorities.

An AI-powered DLP—especially one that uses natural language processing technology such as LLMs—neatly circumvents the problems with traditional DLP techniques. LLMs use the context around keywords by design. As such, they produce more accurate verdicts with fewer false positives. With adequately trained LLMs, creating new DLP filters and policies is quick, as the LLMs are already adept at understanding most of the standard programming and natural language patterns. Finally, support for new intellectual property types and additional natural languages is either already available in LLMs or can be added in one shot with a new training run.

Notwithstanding the advantages of LLMs, technical teams incorporating them into a DLP engine need to overcome some unique challenges:

1. Cost management: LLMs, depending on their origin, vintage, and network placement, can be expensive to operate. If the cost is too high, either the DLP engine will be priced beyond the reach of most users, or the product will become unviable.
2. Technology management: LLMs are an active area of research and development. As a result, better/cheaper/faster LLMs are always coming to the market. Switching to the most effective LLMs available is critical for viable products, but doing so requires the technical team to build the proper infrastructure ahead of time.
3. Latency management: LLMs can produce a verdict more slowly than traditional techniques. This is a direct result of the context that the LLMs use in processing documents. Thus, deft engineering is needed to ensure that the LLMs in use produce a timely response.
4. Policy management: New capabilities in the DLP means more choice in constructing DLP filters. Sometimes this choice can lead to complex DLP filters and policies. As a result, some mechanism to manage policy complexity is needed.

The Cisco Secure Access technical team has successfully surmounted the challenges surrounding the use of LLMs in DLP. As a result, a cost-effective system that can detect intellectual property has already been deployed and is in widespread use. The detection accuracy is very high, with a very low false positive rate. Further, the product's architecture enables rapid experimentation with and absorption of new LLMs. Also, specific techniques in the product enable the DLP engine to produce verdicts quickly, comparable with the response time from traditional techniques. Finally, the DLP engine comes with ready-to-use common filters, including those applicable to intellectual property, reducing the administrator's policy management burden.

Realtime (Inline) DLP mode

Components of Cisco Secure Access, including DLP, sit between enterprise users' devices and the public internet. As a result, for customers deploying Cisco Secure Access, its DLP component is already in the path of (inline to) user traffic going to destinations outside the enterprise, such as public LLMs. Thus, Cisco Secure Access DLP can detect intellectual property movement. Further, based on administrator configuration, the DLP engine can block the movement of intellectual property, such as proprietary source code, to public LLMs (see Figure 1).

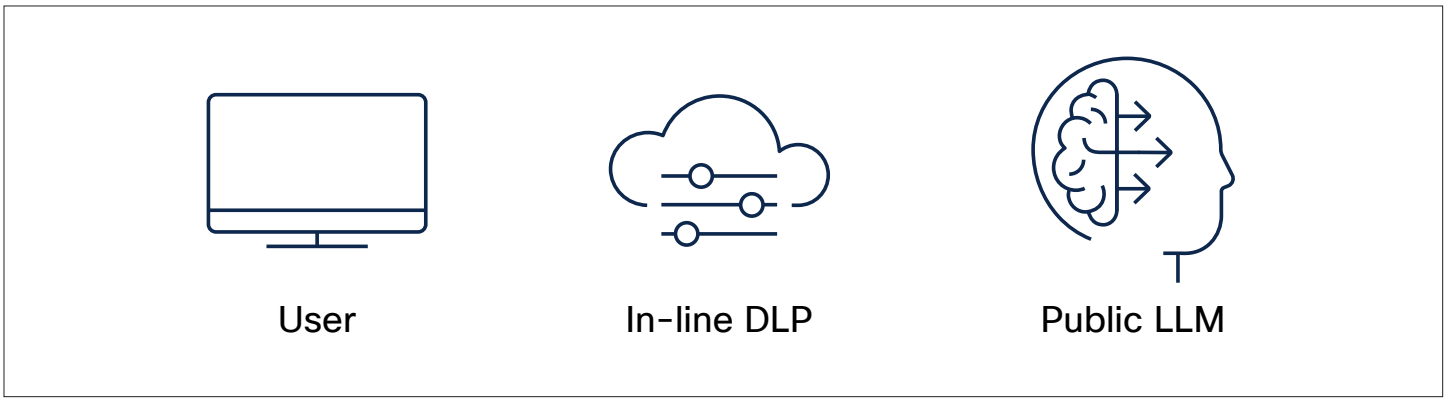

Figure 1: Inline DLP between the user and Public LLM

When Cisco Secure Access DLP sees a document, it doesn't initially know whether it contains intellectual property. However, DLP does know that it has intercepted communication between a user and a public LLM (assuming that the user was, in fact, communicating with an LLM).

The DLP sends the intercepted document to a subcomponent called a scanner. The scanner is responsible for classifying the document and, in the process, determining if the document contains intellectual property such as source code[iii]. Based on the security administrator's configuration settings, the scanner may also return a verdict: block the document's transfer if intellectual property is found or allow the transmission if intellectual property is not found.

The DLP subcomponent that originally intercepted the unclassified document between a user and a public LLM now acts on the verdict. It either allows the document through or blocks its transmission while providing an explanation to the user. Of course, administrator alerts may be raised as well based on configuration settings.

SaaS (Out-of-band) DLP mode

Cisco Secure Access DLP also has a mode where it can scan cloud storage out-of-band to detect intellectual property in stored documents.

In the out-of-band mode, the DLP engine scans the cloud storage using the cloud providers' Representational State Transfer Application Programming Interface (REST API). Each document obtained through the API is processed to determine if it contains intellectual property such as source code or patent application.

When intellectual property documents are found, the DLP component checks for protections applied to those documents. If the protections are inadequate—for example, a patent application resides in a publicly visible folder—the component issues alerts. Cisco Secure Access DLP can also automatically quarantine such documents or revoke public access, buying time for the security administrators to triage the alerts.

Conclusion

Cisco has been working on network security technologies for over thirty years. In recent years, it has invested heavily in AI for Security. Cisco uses AI to assist security administrators, augment human ability to detect incoming threats, and automate mundane and repetitive security tasks.

Cisco Secure Access DLP's intellectual property detection is an example of using AI techniques to augment human ability. Here, AI enables DLP to detect and, where required, block potential or attempted intellectual property movement across trust boundaries.

i Whoops, Samsung workers accidentally leaked trade secrets via ChatGPT, April 6, 2023.

iii Cisco Secure Access, retrieved August 1, 2024.

ii In addition to AI-powered techniques, the DLP engine retains traditional techniques such as Exact Data Matching (EDM), Indexed Document Matching (IDM), Optical Character Recognition (OCR), and more. Mention of these techniques is omitted from the paper for ease of exposition

C11-4684591-00 09/24