Seguridad

Cómo aplicamos el aprendizaje automatizado en soluciones para amenazas avanzadas de Cisco

Últimamente, hay mucho de que hablar sobre machine learning cuando se trata de la ciberseguridad. Parece que no se puede hablar de un tema sin mencionar el otro. Muchas de las organizaciones con las que he hablado en los últimos meses están interesadas en saber más, pero a menudo acaban más confundidas después de comenzar a investigar. Eche un vistazo a Desmitificando el machine learning en el punto de acceso.

En Cisco, hemos estado utilizando ML durante décadas, por lo que este tema no es algo nuevo. Solo en seguridad contamos con varios equipos y más de 20 doctores en machine learning. Nuestros equipos utilizan el machine learning como un método para detectar y analizar las amenazas. Es un método, no un resultado. Es una distinción importante en seguridad. En los últimos años, hemos visto que muchas empresas promocionan su machine learning, pero nunca están dispuestas a explicar lo que significa realmente.

¿De qué manera es diferente Cisco?

En 2013, adquirimos Cognitive Security, una compañía completamente dedicada al machine learning. Hemos integrado rápidamente su tecnología (ahora se llama inteligencia cognitiva) con nuestras soluciones de seguridad web para aumentar las detecciones (ver blogs). Se trata de un enfoque pasivo en la detección. Los registros se envían desde el proxy a la inteligencia cognitiva para su análisis. Analizamos los atributos de los registros, sin tener que examinar nunca el contenido, para descubrir la actividad anómala en el curso de la normalidad. El resultado es simple: la inteligencia cognitiva sólo alerta sobre los hosts de los que se puede decir que están definitivamente comprometidos. Ya que solo le avisa de las infecciones confirmadas, los analistas no pierden el tiempo y se establecen estrictamente en buscar un remedio y en limpiar.

Esto fue solo el principio de la incorporación del machine learning en nuestra cartera de seguridad. Rápidamente nos dimos cuenta del valor de esta tecnología y comenzamos a utilizar sus grandes capacidades de análisis en otras partes del stack de seguridad. Hemos incorporado los algoritmos para correlacionar grandes cantidades de datos y proporcionar inteligencia más allá de lo que se podría ver desde un único vector. Por ejemplo, puede correlacionar datos de tráfico en la red con la comunicación proxy saliente para identificar un host en riesgo con privilegios de administrador y movimientos laterales, que sería imposible de detectar utilizando una única tecnología. Sin embargo, podría hacerlo si conectase varias piezas juntas. Y ahí nos dimos cuenta del valor real de lo que teníamos.

Machine learning aplicado a la telemetría de red

Obviamente, a Cisco se le conoce por ser líder en switches y routers. Básicamente, creamos la red troncal de Internet y la infraestructura de la mayoría de las organizaciones. Esta infraestructura de red existente es una rica fuente de datos. Por ejemplo, Stealthwatch recopila y analiza la telemetría de red con el fin de detectar las amenazas que podrían esconderse dentro. Asimismo, integra el motor de machine learning de inteligencia cognitiva, que correlaciona comportamientos de las amenazas vistos localmente dentro de la empresa con los que se han visto a nivel mundial. Es capaz de detectar anomalías y es también lo suficientemente inteligente para supervisar después las partes individuales reales de la " actividad de las amenazas" (porque lo que es anómalo no tiene por qué ser malicioso necesariamente), lo que lleva a las alertas críticas y de alta fidelidad. También es la tecnología principal que está detrás del Análisis de tráfico cifrado (ETA, por sus siglas en inglés), que puede detectar malware en tráfico cifrado *sin descifrar*; una primicia en el sector.

Machine learning en el punto de acceso

Cuando se habla de seguridad en el punto de acceso, normalmente se acepta que la detección basada en firmas (por ejemplo, haches de archivo) forma parte de la solución, pero no es LA solución. La complejidad de los cambios en los valores de hash de archivo o rangos de direcciones IP es trivial, lo que significa que los adversarios pueden generar nuevos hashes SHA256 para cada infección. Mientras que un valor hash puede ser suficiente para identificar un único archivo malicioso, no ayuda a identificar otros relacionados con infecciones de malware polimórfico que pueden estar asociadas con el mismo ataque o incluso el mismo atacante. Básicamente, el mismo hash nunca se verá dos veces.

Cuando aplicamos el machine learning a estos archivos, podemos diseccionar cada uno de ellos que analizamos en partes. Es parecido a mirar las piezas individuales que forman el coche frente al coche completo. Sí, los coches tienen neumáticos, motor, parabrisas, ventanas, un chasis y demás. Pero, obviamente, no todos los coches son iguales. Lo mismo pasa con el malware. Podemos descomponer cada amenaza individual en mínimos detalles (más de 400 diferentes características) Estos atributos se utilizan como clasificadores específicos en el modelo del machine learning, el nivel aumentado de los detalles tiene como resultado un algoritmo más inteligente y mejor formado, así como unos resultados con mayor fidelidad. Esto significa que nuestro machine learning es mejor a la hora de buscar esas nuevas y rediseñadas amenazas. Los agentes de amenazas a menudo vuelven a empaquetar sus vulnerabilidades en diferentes formatos, como la vulnerabilidad de Flash de CVE-2018-4878 que se utilizó en varias vulnerabilidades, como ROKRAT, y es la campaña de seguimiento. El machine learning es una de las 14 diferentes técnicas que AMP para terminales utiliza para detectar y protegerse contra las amenazas.

Unir las piezas

Una de las formas en las que estamos subiendo el nivel en Cisco es mediante la definición de modelos de agente de amenazas mediante el motor de análisis y machine learning, la inteligencia cognitiva. Al correlacionar la telemetría de los registros de proxy web (Cisco y terceros), la telemetría de red (de Stealthwatch), los valores SHA256 y el comportamiento de archivos de AMP, identifica cómo funcionan los atacantes, lo que hacen e incluso quiénes son. Cuando enviemos esta cantidad de datos a nuestros algoritmos de machine learning, obtendrá un nivel de detecciones sin precedentes y, lo que es más importante, bloquearemos más amenazas antes de que se conviertan en un problema. Exploraremos varios clasificadores en profundidad en futuros blogs.

Puede probar AMP para terminales con una prueba gratuita aquí: www.cisco.com/go/tryamp.

Para ver más en profundidad cómo utilizamos y aplicamos el machine learning dentro de las soluciones de seguridad de Cisco, vea este vídeo técnico.

Etiquetas:

- Protección frente a malware avanzado
- soluciones de amenazas avanzadas
- seguridad para terminales
- machine learning

En un esfuerzo por mantener actualizadas sus conversaciones, los blogs de Cisco cierran los comentarios después de 60 días. Visite la página de Blogs de Cisco para ver el contenido más reciente.

- Suscríbase a la seguridad
- Póngase en contacto con la seguridad

- Otras lecturas
  - Análisis de vulnerabilidad CERT
  - Defensa e investigación de seguridad de Microsoft
  - SANS Internet Storm Center
  - Schneier en Seguridad
- Enlaces relacionados
  - Asesoramiento y respuestas
  - Informes de riesgo cibernético
  - Mejores prácticas de seguridad
  - Security Intelligence Operations
  - Productos de seguridad