Especialistas en ciberseguridad han identificado lo que describen como el primer complemento malicioso de Microsoft Outlook detectado activamente en entornos reales.

En este atípico ataque a la cadena de suministro, analizado por Koi Security, un actor desconocido se apoderó del dominio vinculado a un complemento legítimo que había sido abandonado y lo utilizó para alojar una página falsa de inicio de sesión de Microsoft, logrando sustraer más de 4,000 credenciales. La firma de seguridad denominó la operación como AgreeToSteal.

El complemento afectado, llamado AgreeTo, se promocionaba como una herramienta para centralizar distintos calendarios en un único lugar y compartir la disponibilidad del usuario por correo electrónico. Su última actualización se registró en diciembre de 2022.

Idan Dardikman, cofundador y CTO de Koi, explicó a The Hacker News que el incidente evidencia una ampliación de los vectores de ataque dirigidos a la cadena de suministro.

«Se trata del mismo tipo de ataque que hemos observado en extensiones de navegador, paquetes npm y complementos de IDE: un canal de distribución confiable cuyo contenido puede modificarse después de haber sido aprobado», afirmó Dardikman. *«Lo que hace que los complementos de Office resulten especialmente preocupantes es la combinación de factores: se ejecutan dentro de Outlook, donde los usuarios gestionan comunicaciones altamente sensibles, pueden solicitar permisos para leer y modificar correos electrónicos, y se distribuyen a través de la tienda oficial de Microsoft, lo que implica un nivel de confianza inherente.»*

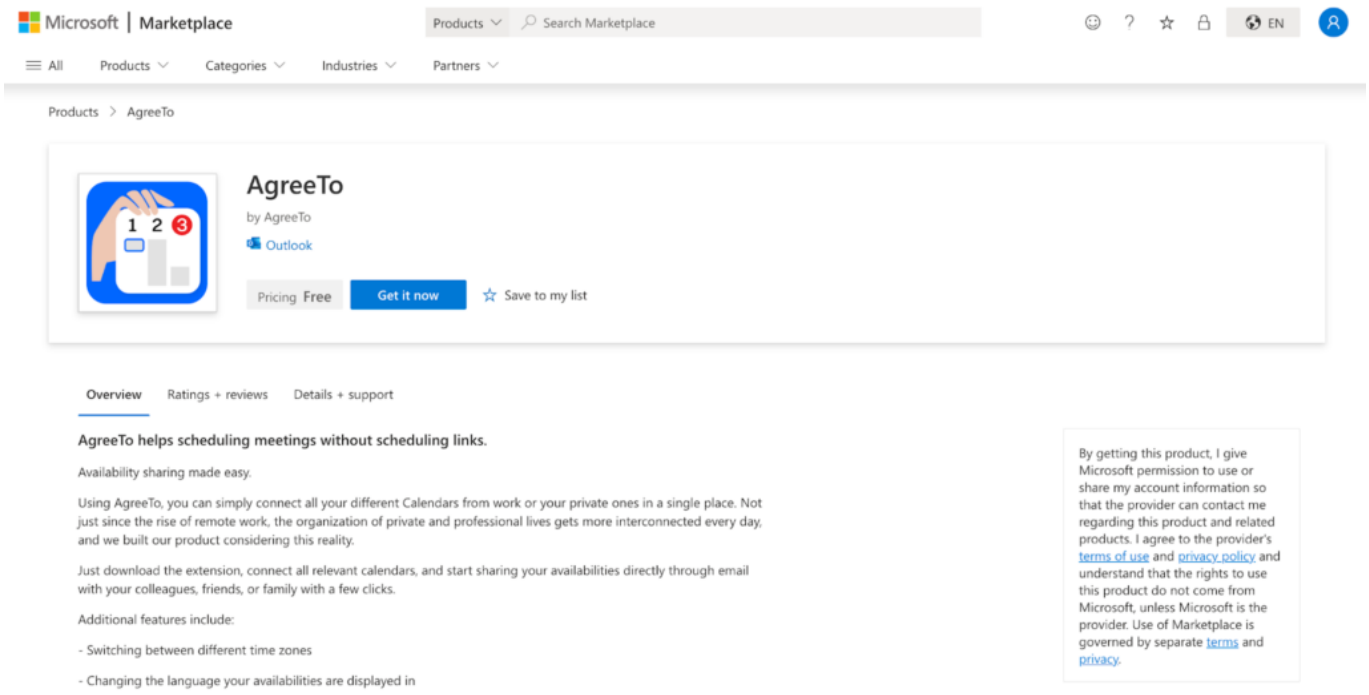
«El caso de AgreeTo añade un elemento adicional: el desarrollador original no cometió ninguna irregularidad. Creó un producto legítimo y posteriormente lo abandonó. El ataque aprovechó el intervalo entre el abandono del proyecto y el momento en que la plataforma detecta esa situación. Cualquier marketplace que dependa de dependencias dinámicas remotas es vulnerable a este escenario.»

En esencia, el ataque se aprovecha de la forma en que operan los complementos de Office y

de la ausencia de un monitoreo periódico del contenido publicado en el Marketplace. Según la documentación de Microsoft, los desarrolladores deben crear una cuenta y enviar su solución al Partner Center, donde pasa por un proceso de revisión antes de su aprobación.

Además, los complementos de Office emplean un archivo de manifiesto que especifica una URL. El contenido alojado en esa dirección se carga y se muestra en tiempo real desde el servidor del desarrollador cada vez que el complemento se abre dentro de un elemento iframe en la aplicación. No obstante, no existe un mecanismo que impida que un tercero reclame un dominio que haya expirado.

En el caso de AgreeTo, el manifiesto apuntaba a una URL alojada en Vercel («outlook-one.vercel[.]app»). Cuando el desarrollador eliminó su implementación —tras convertirse el proyecto en abandonado alrededor de 2023— la dirección quedó disponible para ser reclamada. Al momento de redactar el informe, la infraestructura sigue activa.



El atacante explotó esta circunstancia para desplegar en esa URL un kit de phishing que mostraba una página falsa de autenticación de Microsoft. Las contraseñas introducidas eran capturadas y enviadas mediante la API de Telegram Bot, tras lo cual la víctima era redirigida al portal legítimo de Microsoft.

Sin embargo, Koi advierte que el impacto pudo haber sido más grave. Dado que el complemento contaba con permisos *«ReadWriteItem»* —que permiten leer y modificar los correos electrónicos del usuario— un actor malicioso podría haber utilizado este punto ciego para inyectar código JavaScript capaz de extraer de forma encubierta el contenido del buzón de la víctima.

Los hallazgos refuerzan, una vez más, la necesidad de realizar escaneos periódicos de los paquetes y herramientas publicados en marketplaces y repositorios para identificar actividades maliciosas o sospechosas.

Dardikman señaló que, si bien Microsoft revisa el manifiesto durante la fase inicial de envío, no existe supervisión sobre el contenido dinámico que se obtiene en tiempo real desde el servidor del desarrollador cada vez que el complemento se ejecuta, una vez firmado y aprobado. En consecuencia, la falta de monitoreo continuo sobre lo que entrega la URL abre la puerta a riesgos de seguridad no previstos.

«Los complementos de Office son esencialmente distintos al software tradicional», agregó Dardikman. *«No distribuyen un paquete de código estático. El manifiesto simplemente declara una URL, y cualquier contenido que esa dirección proporcione en un momento determinado es lo que se ejecuta dentro de Outlook. En el caso de AgreeTo, Microsoft firmó el manifiesto en diciembre de 2022 apuntando a outlook-one.vercel.app. Esa misma URL ahora aloja un kit de phishing, y el complemento continúa listado en la tienda.»*

Para mitigar los riesgos derivados de esta amenaza, Koi propone varias medidas que Microsoft podría implementar:

- Activar una nueva revisión cuando la URL de un complemento comience a entregar

contenido distinto al aprobado originalmente.
- Verificar la titularidad del dominio para asegurar que sigue bajo control del desarrollador legítimo y marcar aquellos complementos cuya infraestructura haya cambiado de propietario.
- Implementar un mecanismo para retirar o señalar complementos que no hayan recibido actualizaciones en un periodo prolongado.
- Mostrar el número de instalaciones como referencia para dimensionar el posible impacto.

The Hacker News se ha puesto en contacto con Microsoft para solicitar comentarios y actualizará la información si recibe respuesta.

Cabe señalar que este problema no se limita exclusivamente a Microsoft Marketplace u Office Store. El mes pasado, Open VSX anunció planes para aplicar controles de seguridad antes de permitir la publicación de extensiones de Microsoft Visual Studio Code (VS Code) en su repositorio de código abierto. De forma similar, el VS Code Marketplace de Microsoft realiza escaneos masivos periódicos de todos los paquetes alojados en su registro.

«El problema estructural es el mismo en todos los marketplaces que alojan dependencias dinámicas remotas: se aprueba una vez y se confía para siempre», concluyó Dardikman. *«Los detalles pueden variar según la plataforma, pero la brecha fundamental que hizo posible el caso AgreeTo existe en cualquier entorno donde se revisa un manifiesto en el momento del envío sin supervisar posteriormente el contenido real de las URL referenciadas.»*