Efficient DDoS flood attack detection using dynamic thresholding on flow-based network traffic

Computer and Security 2019

64 citations

این مقاله می‌خواهد با استفاده از یک روش آماری و بر اساس اطلاعاتی که از ترافیک شبکه به دست می‌اورد و همچنین مقدارآستانه تشخیص پویا (چون ترافیک و رفتار شبکه دایماً در حال تغییر می‌باشد)، به شناسایی حملات منع خدمت کمک کند. ویژگی‌های[1] ترافیک را استخراج می‌کند و چهار خصیصه[2] ترافیک شبکه(تعداد بسته‌ها،آدرس مبدا، آدرس مقصد، پروتکل) را به این خاطر که اساس شاخصه‌های[3] حملات منع خدمت می‌باشند، محاسبه می‌کند و زمانی که این خصیصه‌ها در یک بازه‌زمانی[4] از یک مقدار آستانه‌ای[5] بیشتر شد، به معنای شناسایی حمله می‌باشد.

اجزای روش ارایه شده:

برای ساخت شبکه تست‌مان، از virtual box استفاده کرده‌ایم. دو تا هاست، یک سرور و یک مهاجم داریم که از ابزارهای hping و hzenne و همچنین Metasploit استفاده می‌کند.

تولید ترافیک: حمله کننده از hping3 و hyenne برای تولید ترافیک حمله‌های متفاوت (syn flood,IP spoof,...) استفاده می‌کند. دو هاست نیز به وبسایت آپاچی سرور دسترسی خواهند یافت. سپس از وایرشارک نیز برای ضبط بسته‌ها استفاده می‌شود. البته ما ارزیابی را بر روی DARPA نیز انجام می‌دهیم.

استخراج ویژگی: از بسیاری از ویژگی‌های به خصوص ترافیک می‌توان به وقوع حمله پی برد: حجم زیادی از آدرس‌های مبدا متمایز(در صورت استفاده از ابزارها)، آدرس مقصد یکسان، پورت‌های مقصد تصادفی، پروتکل یکسان و سایز بسته در حدود ۴۰-۶۰ بایت. چهار تا خصیصه (تعداد بسته‌ها، آدرس‌های مبدا یکتا، آدرس‌های مقصد یکتا و تعداد پروتکل‌های یکتا) را از ویژگی‌های جمع‌آوری شده[6] ترافیک در هر n ثانیه از ترافیک محاسبه می‌کند. برای شناسایی حملات منع خدمت توزیع شده، خصیصه سوم را به صورت خصیصه دوم/ خصیصه سوم و برای حملات منع خدمت خصیصه اول/ خصیصه سوم و به همین ترتیب برای دو حمله خصیصه چهارم را برابر خصیصه چهارم/ اول درنظر می‌گیریم. سپس این ۴ پارامتر را نرمال می‌کنیم(تقسیم بر بیشترین مقدار).

تشخیص با استفاده از الگوریتم آستانه پویا: مقادیر چهار خصیصه را در هر بازه زمانی محاسبه کرده و واریانس و میانگین این مقادیر را محاسبه کرده و به کمک این مقادیر، آستانه را به صورت پویا برای هر خصیصه تنظیم می‌کند:

$$Th = (\mu + \alpha)*\beta, \beta = 1.5$$

مقدارβ را نیز بر اساس مقایسه مقادیر خصیصه‌ها با بازه‌زمانی قبلی، به صورت پویا تنظیم می‌کند. و اگر مقادیر تمام خصیصه‌ها از آستانه شان بیشتر شد، به عنوان حمله در نظر می‌گیرد.

---

[1] feature
[2] attribute
[3] characteristic
[4] Time interval
[5] threshold
[6] aggregated

ارزیابی و نتایج آن: از دیتاست DARPA در دو نسخه ۹۸ که طول زمانی حمله آن ۵۴ دقیقه و نسخه ۲۰۰۰ که طول زمانی آن ۵ ثانیه استفاده می‌کنیم. بسته ها را در بازه زمانی ۳ ثانیه و ۱ دقیقه جمع کرده‌ایم. و همچنین از دیتاست تولید شده توسط خودش برای تایید و بررسی کارایی مدل استفاده می‌کند.از ابزارهای وایرشارک و متلب نیز در این آزمایشات کمک می‌گیرد. در مقایسه با یکی از روش‌های آماری اخیراً ارایه شده، زمان محاسباتی کمتر(به دلیل اینکه از هیچ مدل مبتنی بر زمان استفاده نمی‌کند)، دقت بالاتر و نرخ شناسایی بالاتر را ارایه می‌دهد.