

استنباط خواص آماری جریان‌ها از روی بسته‌های نمونه‌برداری شده از شبکه با استفاده از شبکه‌های عصبی

مهدی شادور و مسعود صبائی

ایران، تهران، دانشگاه صنعتی امیر کبیر، دانشکده مهندسی کامپیوتر و فناوری اطلاعات

E-mail: {shadvar, sabaei}@aut.ac.ir

چکیده: نظارت و سنجش رفتار ترافیک در شبکه پیش‌نیاز طیف وسیعی از فعالیت‌های مهندسی و مدیریت شبکه است. این کار مستلزم ثبت تک تک بسته‌های عبوری از یک لینک است. از سوی دیگر رفتار ترافیک در سطح بسته دارای جهش‌های ناگهانی و زیاد است، در حالی که ترافیک در سطح جریان رفتار قابل پیش‌بینی‌تری را از خود به نمایش می‌گذارد. در شبکه‌های امروزی با وجود خطوط پرسرعت گیگابیت و ترابیت گرفتن آمار تمام بسته‌های عبوری سبب اختلال در عملکرد مسیریاب و تحمیل بار اضافی به آن خواهد شد. نمونه‌برداری از بسته به عنوان یک راه حل تعمیم‌پذیر برای حل این مشکل پیشنهاد شده است. برای مثال کار گروه‌های IPFIX و PSAMP در IETF هر دو نمونه‌برداری از بسته را توصیه می‌کنند و استاندارد را برای این کار تدوین کرده‌اند. اساسی‌ترین و مهم‌ترین سؤال در ارتباط با نمونه‌برداری از بسته این است که این اندازه‌گیری تا چه اندازه دقیق است، به عبارت دیگر از چه روشی می‌بایست استفاده گردد تا آمارهای بدست آمده از نمونه‌برداری با آمارهای واقعی بیشترین تطابق را داشته باشد. در این تحقیق روشی بر مبنای شبکه‌های عصبی برای بازبازی آمارهای جریان ترافیک در بازه‌های زمانی کوتاه پیشنهاد شده است. در این روش علاوه بر تعداد متوسط و میانگین طول جریان‌ها، توزیع اندازه این آمارها نیز با دقت قابل قبولی بدست می‌آید. همچنین کارایی این روش با روش‌های قبلی مقایسه شده که بهبود قابل ملاحظه‌ای در نتایج مشاهده می‌شود.

کلمات کلیدی: توزیع generalized pareto، جریان‌های ترافیکی، شبکه‌های عصبی MLP، نمونه‌برداری بسته‌ها

۱-مقدمه

بسته‌ها در لینک بیشتر از زمانی است که برای پردازش بسته‌ها و انتقال آن به جمع‌آوری کننده مورد نیاز است. برای گرفتن آمار دقیق جریان‌های عبوری، می‌بایست برای هر یک از جریان‌ها یک مدخل جداگانه در جدول جریان مسیریاب ایجاد شود و هر بسته که از لینک عبور می‌کند مشخصاتش با مشخصات جریان‌های موجود در جدول تطابق داده شود. چنانچه این جریان قبلاً جزو جریان‌های فعال در جدول موجود بود، اطلاعات بسته جدید به آن اضافه می‌شود و در غیر این صورت یک مدخل جدید در جدول مسیریاب برای آن ایجاد می‌شود. محدودیت اصلی منابع برای جمع‌آوری اطلاعات جریان‌ها در حافظه جریان مسیریاب است. به منظور انجام مراجعه به فیلدهای کلیدی بسته و

اندازه‌گیری و نظارت ترافیک پایه و اساس محدوده وسیعی از فعالیت‌های شبکه‌های IP شامل مدیریت شبکه، فعالیت‌های مهندسی شبکه مانند رفع اشکال، حسابرسی، سابقه‌نما مصرف، پیکربندی وزن‌های مسیریابی، تعدیل بار، تخمین ظرفیت مورد نیاز و ... است. سابق بر این، این کار بوسیله ثبت تک تک بسته‌هایی که از یک رابط یا لینک عبور می‌کردند انجام می‌شد. امروزه با وجود لینک‌های پرسرعت (گیگا بیت یا ترابیت) و بار زیادی که بر روی لینک‌های هسته شبکه حمل می‌شود، امکان گرفتن آمار تک تک بسته‌ها و در نتیجه آمارهای ترافیک با دقت زیاد وجود ندارد. در واقع در این مسیریاب‌ها سرعت حرکت