مجموعه داده‌های شبکه‌سازی مبتنی بر زیرساخت و هدف‌محور برای شبکه‌های خودگردان مبتنی بر هوش مصنوعی نسل پنجم و فراتر از آن

چکیده

در عصر شبکه‌های خودمختار (AN)، هوش مصنوعی (AI) نقش حیاتی در توسعه آنها در شبکه‌های سلولی، به ویژه در شبکه‌های G5 و فراتر از آن، ایفا می‌کند. در دسترس بودن مجموعه داده‌های شبکه‌ای با کیفیت بالا یکی از جنبه‌های ضروری برای ایجاد الگوریتم‌های داده‌محور در وظایف مدیریت و بهینه‌سازی شبکه است. این مجموعه داده‌ها به عنوان پایه و اساس توانمندسازی الگوریتم‌های هوش مصنوعی برای تصمیم‌گیری آگاهانه و بهینه‌سازی کارآمد منابع شبکه عمل می‌کنند. در این کار تحقیقاتی، ما مجموعه داده‌های شبکه IW-IB-5GNET را پیشنهاد می‌کنیم: یک مجموعه داده در سطح زیرساخت و مبتنی بر هدف که برای استفاده در تحقیق و توسعه راه‌حل‌های مدیریت و بهینه‌سازی شبکه در شبکه‌های G5 و فراتر از آن در نظر گرفته شده است. این مجموعه داده در سطح زیرساخت است زیرا شامل اطلاعاتی از تمام لایه‌های شبکه G5 است. همچنین مبتنی بر هدف است زیرا بر اساس اهداف از پیش تعریف شده کاربر آغاز می‌شود. مجموعه داده پیشنهادی در یک شبکه G5 شبیه‌سازی شده با استقرار گسترده حسگرهای شبکه برای ایجاد آن تولید شده است. مجموعه داده IW-IB-5GNET نویدبخش تسهیل توسعه راه‌حل‌های مدیریت شبکه هوشمند و خودمختار است که عملکرد و بهینه‌سازی شبکه را افزایش می‌دهد.

کلمات کلیدی: مجموعه داده شبکه؛ قوانین کنترل شبکه؛ مدیریت شبکه؛ بهینه‌سازی شبکه؛ G5؛ شبکه‌سازی مبتنی بر هدف

1. مقدمه

مدیریت شبکه و بهینه‌سازی شبکه از جنبه‌های کلیدی در توسعه شبکه‌های G5 و فراتر از آن (B5G) و موفقیت و قابلیت اطمینان آنها در پشتیبانی از تقاضاهای آینده خدمات و برنامه‌های ارتباطی هستند. انقلاب AN آماده است تا تأثیر عمیقی بر شبکه‌های B5G داشته باشد و عصر جدیدی از اتصال و ارتباطات هوشمند را آغاز کند. از آنجایی که شبکه‌های سنتی از نظر ظرفیت سیستم، کارایی عملیات و غیره با محدودیت‌هایی مواجه هستند، سیستم‌های خودمختار به عنوان راه حلی برای مقابله با این چالش‌ها و دستیابی به سطوح بی‌سابقه‌ای از عملکرد در حال ظهور هستند. این تکامل فناوری منجر به معرفی سطوح جدیدی از هوش و اتوماسیون در لایه‌های مدیریت و تأمین شبکه G5 می‌شود [1]. با بهره‌گیری از فناوری‌های پیشرفته مانند هوش مصنوعی، یادگیری ماشین (ML) و شبکه‌سازی تعریف‌شده توسط نرم‌افزار (SDN)، ANها می‌توانند توانایی‌های پیشرفته خودمدیریتی، خودبهینه‌سازی و خودترمیمی را به دست آورند. یک AN B5G به زیرساخت شبکه‌ای اشاره دارد که قابلیت‌های فناوری‌های G5 و فراتر از آن و اصول شبکه‌سازی خودمختار را ترکیب می‌کند.

شبکه‌های عصبی خودکار (AN) از اجزای مجازی، عامل‌های خودکار و موتورهای تصمیم‌گیری هوشمند تشکیل شده‌اند که قادر به انجام کنترل‌های حلقه بسته هستند [1]. در مورد شبکه‌های B5G، یک شبکه عصبی خودکار (AN) با ترکیب ویژگی‌های پیشرفته‌ای مانند برش شبکه، مجازی‌سازی، نرم‌افزارسازی و محاسبات لبه‌ای، از تکامل شبکه‌های سلولی سنتی بهره می‌برد. برای اینکه شبکه‌های عصبی خودکار در زیرساخت‌های واقعی ما مستقر شوند، لازم است شبکه‌های مدیریتی داشته باشیم که قادر به انجام همه این ویژگی‌های پیشرفته باشند. یکی از الگوهایی که امروزه معمولاً مورد استفاده قرار می‌گیرد، شبکه مبتنی بر قصد (IBN) است. طبق [2]، یک IBN شبکه‌ای است که می‌توان آن را با استفاده از قصد مدیریت کرد. هدف اصلی یک IBN ایجاد یک شبکه خودمختار با ساده‌سازی مدیریت و عملکرد آن است. برای این منظور، یک IBN ایجاد یک چارچوب کامل شبکه خودمختار را تصور می‌کند [3]، که استحکام شبکه را بهبود می‌بخشد و به عملکرد و نگهداری پویا دست می‌یابد [4]. یک IBN از عملکردهای مدیریتی هدایت‌شده با استفاده از قصد پشتیبانی می‌کند. قصد، توصیف سطح بالایی از مجموعه‌ای از اهداف و نتایج عملیاتی است که یک شبکه باید به ترتیب برآورده کند و قرار است ارائه دهد [2]. این مدل، اهداف و نتایج را به شیوه‌ای کاملاً اعلانی تعریف می‌کند، نه اینکه پیکربندی دقیق شبکه [5] یا نحوه دستیابی به آن را مشخص کند. سپس اهداف به سیاست‌های شبکه تبدیل می‌شوند که جزئیات بسیار دقیق‌تری در مورد پیکربندی‌های شبکه ارائه می‌دهند [3]. چنین سیاست‌های شبکه‌ای منجر به اجرای اقدامات شبکه و دستیابی به نتیجه مطلوب مشخص شده در هدف می‌شوند. نمونه‌هایی از اهداف در یک شبکه G5 به شرح زیر است: (i) "اطمینان حاصل کنید که استقرار قطعه شبکه با توجه به تأخیر، پهنای باند و قابلیت اطمینان، با توافق‌نامه سطح خدمات (SLA) مشخص شده مطابقت دارد". (ii) "بالاترین اولویت را به ترافیک برنامه‌های حیاتی اختصاص دهید، تأخیر کم و پهنای باند تضمین شده را تضمین کنید، در حالی که حداقل سطح خدمات را برای ترافیک استاندارد حفظ کنید". (iii) "هرگونه جریان ترافیکی که رفتار مخرب یا تلاش‌های دسترسی غیرمجاز را نشان می‌دهد، بدون ایجاد اختلال در ترافیک مشروع شبکه، از بین ببرید".

ANها تعاملات مبتنی بر هدف را دنبال می‌کنند و از تعامل انسان و ماشین به تعامل منابع حلقه بسته [1] حرکت می‌کنند. در این زمینه، هوش مصنوعی برای فعال کردن این گذار پدیدار می‌شود. هوش مصنوعی به یک ویژگی کلیدی در مدیریت شبکه و بهینه‌سازی شبکه‌های B5G تبدیل شده است. در همین حال، قبل از آموزش و توسعه هوش مصنوعی، برخی مراحل مرتبط با داده‌ها ضروری هستند. این شامل نه تنها پیاده‌سازی و استقرار حسگرهای شبکه و جمع‌آوری‌کننده‌های داده، بلکه پردازش بیشتر و کفایت داده‌ها نیز می‌شود. چنین رویه‌هایی نه تنها امکان استخراج داده‌ها از هر منبع توپولوژی شبکه را فراهم می‌کنند، بلکه بینش‌های قابل توجهی در مورد فرآیندهای شبکه در زمان واقعی نیز به دست می‌دهند [6]. این امر بهینه‌سازی شبکه را افزایش می‌دهد زیرا اکنون می‌توان از چنین داده‌هایی در مدل‌های یادگیری ماشینی برای انجام تصمیمات سریع‌تر و بهتر استفاده کرد. تصمیماتی که به طور سنتی توسط تعاملات کند انسانی گرفته می‌شوند، اکنون می‌توانند به صورت خودکار توسط الگوریتم‌های یادگیری ماشینی انجام شوند.

با وجود تمام مزایایی که راه‌حل‌های مبتنی بر یادگیری ماشینی می‌توانند برای شبکه‌های B5G به ارمغان بیاورند، پیاده‌سازی عملی آنها دشوار است و چالش‌های متعددی را ایجاد می‌کند. اول از همه، سیستم‌های هوش مصنوعی برای آموزش به طیف گسترده‌ای از داده‌ها نیاز دارند. این به معنای دسترسی به یک شبکه واقع‌گرایانه (واقعی یا شبیه‌سازی شده) و همچنین داشتن امکان و ابزارهای لازم برای دسترسی و استخراج داده‌های در زمان واقعی است. گزینه دیگر استفاده از یک مجموعه داده قابل اعتماد و کافی موجود است. برای رویکرد دوم، کمبود مجموعه داده‌های شبکه عمومی مشهود است. علاوه بر این، اکثر آنها قدیمی و غیرقابل اعتماد هستند. این به دلیل سرعت تغییراتی است که شبکه‌ها، به ویژه شبکه‌های سلولی، در طول سال‌ها تجربه می‌کنند. این تغییر در رفتارها و الگوهای شبکه، نیاز به مجموعه داده‌های پویاتری دارد [6]. چنین مجموعه داده‌های جدیدی نه تنها جریان‌های ترافیک و تفاوت‌ها را منعکس می‌کنند، بلکه انواع مختلف حملات، و همچنین فهرست توپولوژی شبکه‌ای که داده‌ها در آن ثبت می‌شوند. بنابراین، مجموعه داده‌ها در سطح زیرساخت آگاه هستند و تمام سطوح زیرساخت را در نظر می‌گیرند: سطح شبکه، سطح گره، سطح رابط و سطح فناوری. ثبت چنین داده‌هایی، این مجموعه داده‌ها را جامع، قابل تکرار، قابل اصلاح و توسعه‌پذیر می‌کند. چالش‌های فوق انگیزه این تحقیق بوده‌اند. این مقاله یک مجموعه داده جامع جدید برای اهداف مدیریت و بهینه‌سازی شبکه در شبکه‌های B5G مبتنی بر هوش مصنوعی پیشنهاد و ارائه می‌دهد. این مجموعه داده شامل داده‌های در سطح زیرساخت (IW) و مبتنی بر هدف (IB) است که از یک شبکه B5G استخراج شده‌اند. چنین داده‌هایی به صورت بلادرنگ با استفاده از یک چارچوب حلقه بسته استخراج می‌شوند.

بقیه مقاله به شرح زیر سازماندهی شده است. بخش 2 خلاصه‌ای از جدیدترین دستاوردهای مربوط به مجموعه داده‌های مرتبط با شبکه موجود را ارائه می‌دهد. در بخش 3، مواد و روش‌های مورد استفاده برای ایجاد مجموعه داده پیشنهادی ارائه شده است. منطقه مورد مطالعه، یک شبکه چند مستاجری G5، همراه با منابع جمع‌آوری داده و حسگرهای شبکه، ارائه شده است. بخش ۴ شبیه‌سازی سناریو برای تولید مجموعه داده‌ها را شرح می‌دهد. این بخش شامل جزئیات پیاده‌سازی و طراحی و اجرای آزمایش‌ها است. بخش ۵ توضیح مفصلی از ساختار مجموعه داده‌ها ارائه می‌دهد. سپس تجزیه و تحلیل و اعتبارسنجی مجموعه داده‌ها در بخش ۶ ارائه می‌شود. بحث در بخش ۷ گنجانده شده است. در نهایت، نتیجه‌گیری در بخش ۸ ارائه شده است.

۲. کارهای مرتبط

برای اطمینان از اینکه مدل‌های هوش مصنوعی در پرداختن به چالش‌های شبکه B5G در دنیای واقعی مؤثر هستند، دسترسی به مجموعه داده‌های واقعی و کامل ضروری است. چنین مجموعه داده‌هایی باید به طور دقیق ترافیک شبکه، توپولوژی شبکه و فعالیت‌هایی را که مدل برای تجزیه و تحلیل و پیش‌بینی در نظر گرفته شده است، نشان دهند. در این بخش، توضیحی عمیق از مجموعه داده‌های مختلف شبکه ارائه شده است. این مجموعه داده‌ها در جدول ۱ شرح داده شده‌اند که شامل دسته‌های مختلفی است که در ردیف‌ها مشخص شده‌اند. ابتدا، در ردیف «نوع شبکه»، مشخص

می‌شود که چه نوع زیرساخت شبکه‌ای برای تولید مجموعه داده توسعه داده شده است. سپس، برخی ردیف‌ها برای توصیف اطلاعات مجموعه داده تعیین شده‌اند. ردیف‌های «توپولوژی» به سطح جزئیات توپولوژی شبکه گزارش شده در مجموعه داده اشاره دارند. ما به «مجموعه داده‌های زیرساختی» به عنوان مجموعه داده‌ای اشاره می‌کنیم که شامل اطلاعاتی از تمام لایه‌های شبکه مشخص شده در ردیف‌های توپولوژی است. در نهایت، چند ردیف به مشخص کردن اینکه آیا مجموعه داده‌ها دارای فراداده و معیارها مطابق با اجزای مختلف سطح زیرساخت هستند یا خیر، اختصاص داده شده است: میزبان، پورت (یا رابط شبکه)، فناوری‌های سطح داده در هر پورت، جریان‌های داده، صف‌های پورت و قوانین پورت فناوری سطح داده. مجموعه داده‌های شرح داده شده در ستون‌ها قرار گرفته‌اند و به سه نوع مختلف تقسیم شده‌اند. نوع 1 شامل مجموعه داده‌های استخراج شده از معماری‌های شبکه رایج مانند LAN، شبکه نظامی یا ابر است. نوع 2 بر شبکه‌های IoT (اینترنت اشیا) ادغام شده در شبکه‌های G5 تمرکز دارد و نوع 3 مجموعه داده‌های استخراج شده از یک شبکه G5 را جمع‌آوری می‌کند. در زیر خلاصه‌ای از آنها آمده است.

جدول 1. جدول مقایسه مجموعه داده‌های شبکه مختلف (np: ارائه نشده است)

		TYPE 1		TYPE 2		TYPE 3	
		KDDCUP99	HIKARI-2021	Edge-IIoTset	Bot-IoT	5G-NIDD	Ours
Network Type		**Military**	**LAN**	**Edge-IoT**	**IoT**	**5G**	**5G**
Dataset info	Extraction tools	tcpdump	tcpdump, Zeek	Zeek	pcap capturing	pcap capturing	Linux scripts, Python tools
	Purpose	NID	NID	NID	NID	NID	NMC *
	Format	csv	csv	csv	different sets	pcap, csv	csv
	Resolution	n.p	n.p	timestamp	timestamp	timestamp	timestamp
	Simulated	Yes	Partial	No	No	No	No
	Features	42	86	61	n.p	112	101
	Year	1999	2021	2022	2019	2022	2023
	Available	[7]	[8]	[9]	[10]	[11]	[12]
Topology	Host	✘	✔	✘	✘	✘	✔
	Port	✘	✘	✘	✘	✘	✔
	Technology	✘	✘	✘	✘	✘	✔
	Flow	✘	✔	✔	✔	✔	✔
Metadata	Host	✘	✘	✘	✘	✘	✔
	Port	✘	✘	✔	✘	✘	✔
	Technology	✘	✘	✘	✘	✘	✔
	Queue	✘	✘	✘	✘	✘	✔
	Flow	✔	✔	✔	✔	✔	✔

		TYPE 1		TYPE 2		TYPE 3	
		KDDCUP99	HIKARI-2021	Edge-IIoTset	Bot-IoT	5G-NIDD	Ours
Network Type		Military	LAN	Edge-IoT	IoT	5G	5G
Metrics	Host	✔	✔	✘	✘	✘	✔
	Port	✔	✔	✔	✘	✔	✔
	Technology	✘	✘	✘	✘	✘	✔
	Queue	✘	✘	✘	✘	✘	✔
	Flow	✔	✔	✔	✔	✔	✔
	Rule	✘	✘	✘	✘	✘	✔

* Network management and control.

با مجموعه داده‌های موجود در نوع 1 شروع می‌کنیم. مجموعه داده KDDCUP99 به منظور توسعه یک آشکارساز نفوذ شبکه ایجاد شد. این مجموعه داده در سال‌های گذشته به طور گسترده برای مشکلات تشخیص ناهنجاری مورد استفاده قرار گرفته است [13]. بیش از 20 نوع حمله در مجموعه داده شبیه‌سازی شده‌اند که می‌توانند به چهار دسته تقسیم شوند: انکار سرویس (DoS)، از راه دور به محلی (R2L)، کاربر به ریشه (U2R) و کاوش. برای دستیابی به تولید مجموعه داده، طیف گسترده‌ای از دستورالعمل‌ها در یک محیط شبکه نظامی شبیه‌سازی شدند. با این حال، این مجموعه داده دارای مشکلات قابل توجهی است که به شدت بر عملکرد IDSها تأثیر می‌گذارد. برای غلبه بر چنین مشکلاتی، مجموعه داده‌های دیگری از این مجموعه داده تولید شدند. به عنوان مثال، مجموعه داده NSL-KDD که توسط Tavallaee و همکاران در [14] پیشنهاد شده است. در [15]، Ferriyan و همکاران مجموعه داده HIKARI-2021 را پیشنهاد کردند، یک مجموعه داده IDS که ترافیک شبکه رمزگذاری شده را در یک محیط دنیای واقعی ارائه می‌دهد. این مجموعه داده با حملات شبکه مختلف برچسب گذاری شده است. این مجموعه داده تا 86 ویژگی استخراج شده با ابزار Zeek دارد که شامل معیارهای میزبان و جریان است.

با تمرکز بر انواع دیگر مجموعه داده‌های شبکه، نوع 2 را در جدول 1 برجسته کرده‌ایم. در [9]، فرراگ و همکارانش مجموعه داده Edge-IIoT را ارائه می‌دهند، یک مجموعه داده امنیت سایبری از برنامه‌های IoT و IIoT. نویسندگان یک بستر آزمایشی IoT/IIoT را با دستگاه‌های مختلف IoT آماده کردند. در چنین سناریویی، آنها حملات مختلف مربوط به پروتکل‌های اتصال IoT و IIoT را شناسایی و تجزیه و تحلیل کردند. علاوه بر این، کورونیوتیس و همکارانش Bot-IoT [16] را ارائه می‌دهند، یک مجموعه داده ترافیک شبکه که شامل سناریوهای بات‌نت در یک شبکه واقعی IoT است. این یک مجموعه داده تشخیص نفوذ است که مدل‌ها را برای تشخیص حملات مختلف بات‌نت در شبکه‌های IoT آموزش می‌دهد. هر دو کار با مجموعه داده‌های ترافیک شبکه IoT واقع‌بینانه و با کیفیت بالا برای NID ارائه می‌شوند. با این حال، هر دوی آنها فقط بر اساس معیارهای جریان هستند و از زیرساخت آگاه نیستند. علاوه بر این، یکی دیگر از معایب این است که هیچ یک از آنها در معماری شبکه G5 مستقر نشده‌اند. در نهایت، نوع ۳ شامل مجموعه داده G-NIDD5 است

که توسط ساماراکون و همکارانش در [17] ارائه شده است، یک مجموعه داده کاملاً برچسب‌گذاری شده که بر روی یک شبکه آزمایشی G5 کاربردی ساخته شده است. این شبکه سناریوهای حمله مختلف و ترافیک غیر مخرب از کاربران واقعی را در خود جای داده است. این مجموعه داده بر NID متمرکز است و در مجموع 112 ویژگی از جمله فرادادهها و معیارهای جریان را در خود جای داده است.

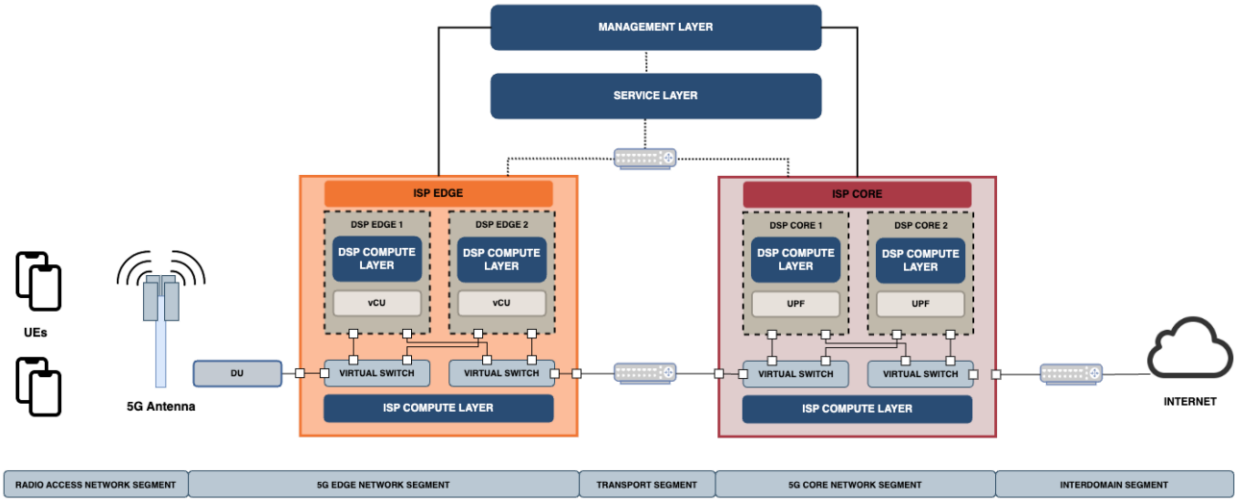
همه این مجموعه دادههای افشا شده با یک هدف ایجاد شدهاند: تشخیص حمله و ناهنجاری. به همین دلیل، همه آنها ویژگی‌های مشابهی در مورد سطح توپولوژی و معیارها و فرادادههای آن در هر دو سطح پورت و جریان دارند. با این حال، نیاز به یک مجموعه داده وجود دارد که توپولوژی شبکه را به همراه دادههای مبتنی بر هدف آن، یعنی دادههایی که فرادادهها و معیارهای مرتبط با سیاست‌های شبکه را که با اهداف شبکه مرتبط هستند، منعکس می‌کنند، با عمق بیشتری ثبت کند. بنابراین، این مجموعه داده نه تنها اطلاعات مربوط به نظارت بر شبکه، بلکه اطلاعات مربوط به کنترل شبکه را نیز ارائه می‌دهد. علاوه بر این، تا آنجا که ما می‌دانیم، هیچ یک از مجموعه دادههای یافت شده شامل دادههای مبتنی بر هدف در معماری شبکه G5 نیستند. یک مجموعه داده با چنین مشخصاتی می‌تواند برای مدیریت، کنترل و بهینه‌سازی شبکه خودمختار مورد استفاده قرار گیرد. این انگیزه اصلی تحقیق حاضر بوده است.

۳. مواد و روش‌ها

در این بخش، حوزه مطالعاتی مجموعه دادهها ارائه شده است. این به یک شبکه چند مستاجری و مبتنی بر قصد G5 اشاره دارد. در بخش ۳.۱، توضیح مختصری از اجزای اصلی یک شبکه G5 رایج ارائه شده است تا مسئله مورد بررسی بهتر در متن قرار گیرد. این توضیح، درک واضح‌تری از منابع زیربنایی ویژگی‌های مختلف شبکه مورد بحث در زیربخش‌های بعدی را برای خواننده فراهم می‌کند. بخش ۳.۲ توضیحی در مورد IBN پیشنهادی ارائه می‌دهد. در نهایت، در بخش ۳.۳، منابع جمع‌آوری داده شبکه G-IB5 ارائه شده است. این مربوط به استخراج دادههای ضروری از حسگرهای مختلف شبکه و فرآیند ذخیره‌سازی بعدی است.

۳.۱. معماری زیرساخت مرجع G۵

شکل ۱ یک زیرساخت شبکه چند مستاجری مرجع G۵ را نشان می‌دهد، که در آن سرویس‌های شبکه در همان زیرساخت فیزیکی، نرم‌افزاری و مجازی‌سازی می‌شوند. در چنین زیرساختی، ترافیک بین مستاجران به دلیل قابلیت‌های مجازی‌سازی و تونل‌سازی ترافیک در بخش‌های شبکه، ایزوله باقی می‌ماند. معماری G۵ به پنج بخش شبکه مختلف تقسیم می‌شود: شبکه دسترسی رادیویی (RAN)، بخش شبکه لبه، بخش انتقال، بخش شبکه هسته و بخش بین دامنه‌ای. فقط اجزای صفحه داده در شکل نشان داده شده‌اند.

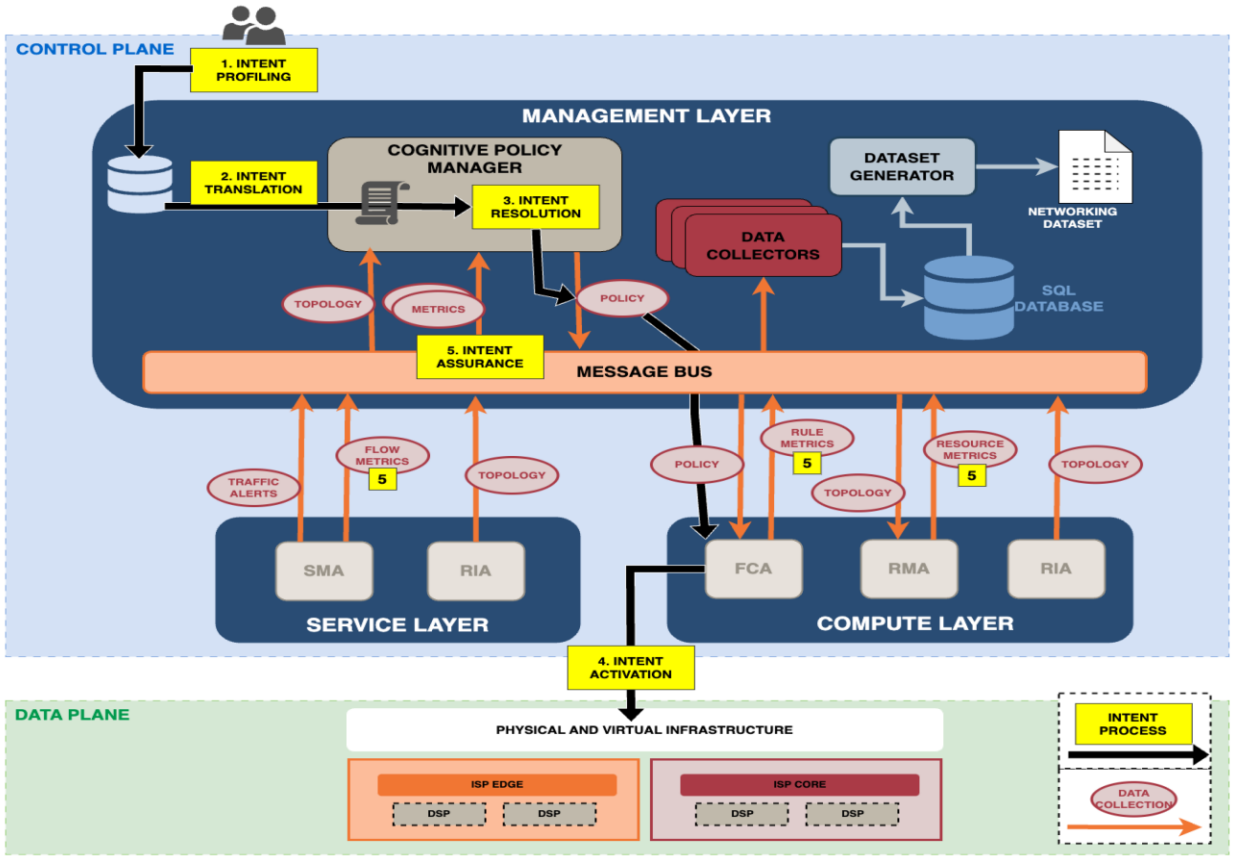


جریان داده از هر بخش عبور می‌کند و هر بخش هدف مشخصی را دنبال می‌کند. با شروع از بخش RAN، این بخش رابط بین دستگاه‌های تجهیزات کاربر (UEها) و شبکه G5 را نشان می‌دهد. این بخش شامل آنتن‌ها (واحدهای رادیویی، RUها)، واحدهای توزیع‌شده (DUها) و سایر تجهیزاتی است که مسئول ارسال و دریافت سیگنال‌های بی‌سیم هستند. RAN از طریق واحدهای متمرکز (CUها) که مجازی‌سازی شده و در شبکه محاسبات لبه موبایل/چنددسترسی (MEC) [18] مستقر شده‌اند، به بخش لبه متصل می‌شود. بخش لبه منابع محاسباتی و ذخیره‌سازی را به کاربران نهایی نزدیک‌تر می‌کند، تأخیر را کاهش می‌دهد و کیفیت خدمات را بهبود می‌بخشد. بخش‌های لبه و هسته از طریق بخش انتقال به هم متصل می‌شوند. شبکه هسته بخش مرکزی شبکه G5 است. این بخش عملکردهای پیشرفته‌ای مانند مدیریت جلسه، مدیریت تحرک و احراز هویت را از طریق اجزای مختلف صفحه کنترل ارائه می‌دهد. این بخش به بخش بین دامنه‌ای متصل است که شامل اتصال بین دامنه‌های اداری یا ارائه دهندگان خدمات مختلف است. در نهایت، همانطور که در شکل 1 نشان داده شده است، سیستم G5 از ذینفعان مختلفی که در [5] توسط G PPP5 (مشارکت عمومی-خصوصی G5) شرح داده شده است، تشکیل شده است. این ذینفعان شامل ارائه دهندگان خدمات زیرساخت (ISP) و ارائه دهندگان خدمات دیجیتال (DSP) هستند که هر دو در تأمین منابع شبکه نقش دارند. معماری ارائه شده از تلاش برای اتوماسیون شبکه که از طریق حلقه‌های شناختی حاصل می‌شود، پشتیبانی می‌کند. قابلیت‌های شبکه خودمختار، شبکه را قادر می‌سازد تا عملیات خود را خود مدیریت و خود بهینه کند. چنین اتوماسیونی با استقرار لایه‌های مختلف شبکه، که لایه محاسبه، سرویس و مدیریت هستند، انجام می‌شود. این لایه‌های شبکه در شکل 1 به صورت جعبه‌های آبی نشان داده شده‌اند. لایه محاسبه نقش مهمی در فعال کردن قابلیت‌های شبکه خودمختار ایفا می‌کند زیرا قدرت پردازش و ظرفیت ذخیره‌سازی لازم برای پشتیبانی از خدمات و برنامه‌ها را فراهم می‌کند. لایه سرویس شامل ایجاد، استقرار و مدیریت خدمات ارائه شده از طریق زیرساخت شبکه G5 است. این لایه بر ارائه مجموعه بزرگی از خدمات که نیازهای متنوع کاربران نهایی و برنامه‌ها را برآورده می‌کند، تمرکز دارد. این لایه از تکنیک‌های اتوماسیون، مجازی‌سازی و هماهنگ‌سازی برای اطمینان از تأمین کارآمد

خدمات، مقیاس‌بندی و سفارشی‌سازی استفاده می‌کند. در نهایت، لایه مدیریت مسئول نظارت و کنترل عملیات و منابع شبکه است. این لایه از تجزیه و تحلیل پیشرفته و الگوریتم‌های هوش مصنوعی برای نظارت، تجزیه و تحلیل و بهینه‌سازی ویژگی‌های مختلف شبکه مانند امنیت، عملکرد و تخصیص منابع استفاده می‌کند. این لایه از داده‌ها و اطلاعات دریافتی از لایه‌های محاسباتی و خدماتی برای تصمیم‌گیری آگاهانه و خودکارسازی فرآیندهای مدیریت شبکه استفاده می‌کند. ادغام این سه لایه در شبکه G5 امکان توسعه قابلیت‌های حلقه بسته مانند سازگاری با شرایط متغیر، محافظت از خود در برابر حملات و بهینه‌سازی استفاده از منابع را فراهم می‌کند [1].

۳.۲. شرح معماری مبتنی بر هدف G۵

شکل ۲ رویکرد IBN ما را برای دستیابی به قابلیت‌های حلقه بسته در شبکه مرجع G۵ نشان می‌دهد. چنین رویکرد IBN با کاهش دخالت متخصص انسانی، فرآیند پیکربندی، تأمین و تضمین شبکه را خودکار می‌کند. معماری پیشنهادی شامل سه لایه توضیح داده شده در بالا است که عبارتند از مدیریت، خدمات و محاسبات، که به تفصیل شرح داده شده‌اند. اجزای متعدد شبکه (حسگرها و محرک‌ها) که در این لایه‌ها اختصاص داده شده‌اند، برای دستیابی به چنین ویژگی‌های خودکاری با هم کار می‌کنند.



سیستم IBN پیشنهادی ما یک پلتفرم حلقه بسته ایجاد می‌کند که در آن الزامات سرویس سطح بالا به صورت خودکار در شبکه هماهنگ و اجرا می‌شوند. کل فرآیند از زمانی که intent وارد شبکه می‌شود تا زمانی که از آن حذف می‌شود، شامل پنج مرحله مختلف است [3]. هر یک از آنها در شکل 2 با کادرهای زرد و فلش‌های سیاه نشان داده شده‌اند. این فرآیند به شرح زیر توضیح داده شده است:

1. پروفایل‌بندی intent. این اولین مرحله IBN است که در آن کاربر با سیستم تعامل می‌کند تا intent مورد نظر را مشخص کند.

2. ترجمه intent. عبارت intent به یک سیاست شبکه ترجمه می‌شود که شامل مجموعه‌ای از قوانین و پیکربندی‌های شبکه است. یک policy مجموعه‌ای از قوانین است که تعریف می‌کند تحت چه شرایطی چه کاری باید انجام شود [19].

3. حل intent. باید در نظر گرفته شود که چندین intent می‌توانند همزمان در شبکه اتفاق بیفتند. به همین دلیل، در طول ترجمه intent، جلوگیری از منجر شدن شبکه به پیکربندی‌های متناقض و متضاد شبکه ضروری است.

4. فعال‌سازی intent. مرحله بعدی پس از تأیید عدم وجود هرگونه تضاد با سایر عبارات intent در شبکه، هماهنگ‌سازی و فعال‌سازی intent است. این مرحله شامل پیکربندی شبکه و ارائه سیاست شبکه درخواستی است. همانطور که در شکل 2 نشان داده شده است، این فرآیند منجر به اجرای یک قانون (یا بیش از یک قانون) می‌شود که در صفحه داده منعکس خواهد شد.

5. تضمین هدف. این مرحله نهایی مستلزم اطمینان از این است که شبکه پس از دستیابی به هدف مورد نظر، واقعاً با آن مطابقت دارد. برای انجام این کار، IBN ما شامل چندین حسگر است که قادر به نظارت بر وضعیت شبکه در زمان تقریباً واقعی هستند. چنین مؤلفه‌هایی معیارها را به لایه مدیریت گزارش می‌دهند (به کادرهای زرد "5" در شکل 2 مراجعه کنید)، که وظیفه اطمینان از تحقق هدف را بر عهده دارد. بسته به نوع هدف، پس از مرحله 5، فرآیند هدف به پایان می‌رسد یا خیر.

با تکمیل این پنج مرحله، حلقه بسته کامل می‌شود. چنین حلقه‌ای به عنوان یک سیستم مستقل عمل می‌کند، به طور مستقل به وظایف شناسایی شده پاسخ می‌دهد و اطمینان حاصل می‌کند که شبکه بدون نیاز به مداخله انسانی در کل فرآیند، با هدف مشخص شده توسط کاربر همسو است.

۳.۳. چارچوب پیشنهادی برای جمع‌آوری داده‌ها

این زیربخش، اجزای نرم‌افزاری مختلف لازم نه تنها برای دستیابی به هدف، بلکه برای فرآیند ایجاد مجموعه داده‌ها را نیز شرح می‌دهد. فرآیند استخراج داده‌ها در شکل ۲ با فلش‌های نارنجی و دایره‌های قرمز به تفصیل شرح داده شده است.

رویکرد شبکه زیرساخت با معماری حفاظت خودمدیریتی پیشنهادی در [20] مطابقت دارد. این رویکرد شامل مجموعه‌ای از اجزای نرم‌افزاری شبکه است که در سه لایه قبلاً تعریف شده توزیع شده‌اند و با هم کار می‌کنند تا یک سیستم حلقه بسته شناختی را مطابقت دهند. هر جزء یک وظیفه خاص را اجرا می‌کند و ترکیب همه آنها منجر به اجرای دقیق قوانین کنترل شبکه می‌شود. چنین قوانین کنترلی قبلاً از یک هدف ترجمه شده‌اند. همانطور که در شکل ۲ با رنگ نارنجی نشان داده شده است، ارتباط و تبادل داده بین اجزا توسط یک نرم‌افزار گذرگاه پیام از طریق معماری انتشار و اشتراک تسهیل می‌شود. دایره‌های قرمز نشان دهنده نوع داده‌ای است که در هر مورد مبادله می‌شود، در حالی که فلش‌های نارنجی نشان دهنده اشتراک یا انتشار آن تبادل داده است. اجزای نرم‌افزاری در زیر شرح داده شده‌اند.

- عامل موجودی منابع (RIA). این یک جزء شبکه است که مسئول انتشار اطلاعات توپولوژیکی شبکه در زمان واقعی است. چنین اطلاعاتی مربوط به دستگاه‌های فیزیکی و مجازی، پورت‌ها و اتصالات بین پورت‌ها و دستگاه‌های موجود در هر دستگاه است. RIA توپولوژی شبکه G5 را کشف می‌کند، در آنجا نمونه‌سازی شده و برای بقیه اجزای شبکه منتشر می‌شود. عملکرد و قابلیت‌های این مؤلفه در [21] ارائه شده است.

- عامل نظارت بر امنیت (SMA). این عامل دو عملکرد متمایز دارد. اول، مسئول افزایش و گسترش قابلیت‌های ارائه شده توسط یک IDS سنتی است. دلیل این محدودیت را می‌توان به ناکافی بودن قابلیت‌های سیستم‌های تشخیص نفوذ شبکه سنتی (NIDS) نسبت داد که قادر به بهره‌برداری کامل از پتانسیل ارائه شده توسط زیرساخت G5 و اطلاعات شبکه همراه آن نیستند. از این رو، SMA با Snort [22] همکاری می‌کند و اطلاعات ارائه شده توسط این NIDS سنتی را گسترش می‌دهد. دوم، اطلاعاتی در مورد جریان‌های شبکه ارائه می‌دهد و فهرستی از تمام جریان‌هایی که از هر یک از رابط‌های شبکه عبور می‌کنند و هشدارهای ترافیکی را ارائه می‌دهد. همچنین معیارهای مرتبط با جریان‌های شبکه گزارش شده را ارائه می‌دهد. این حسگر و قابلیت‌های آن در [23] ارائه شده است. عامل نظارت بر منابع (RMA). این عامل امکان نظارت بر منابع مختلف شبکه را فراهم می‌کند. این عامل، معیارها را از دستگاه‌های شبکه، پورت‌های شبکه (رابط‌های شبکه فیزیکی و مجازی) و فناوری‌های صفحه داده که قبلاً توسط مؤلفه RIA کشف و منتشر شده‌اند، استخراج می‌کند. معیارهای نظارت شده با استفاده از یک فایل پیکربندی پیکربندی می‌شوند و می‌توانند به راحتی با تغییر چنین فایل پیکربندی گسترش یابند.

- مدیر سیاست شناختی (CPM). این مؤلفه چهار وظیفه متمایز را برای تولید سیاست‌های شبکه انجام می‌دهد. ابتدا، عبارات قصد کاربر را در سیاست‌های شبکه ترجمه می‌کند. همزمان، جریان‌ها و معیارهای منابع ارائه شده توسط SMA و RMA را همراه با اطلاعات مکانی RIA تجزیه و تحلیل می‌کند. بنابراین، قادر به ایجاد یک تحلیل فشرده از وضعیت فعلی شبکه است. با در نظر گرفتن سیاست شبکه، با استفاده از چنین تحلیلی تصمیمی می‌گیرد که شامل چه اقدامی، کجا و با چه فناوری صفحه داده‌ای باید انجام شود. چنین اطلاعاتی برای تکمیل سیاست شبکه استفاده خواهد شد. بسته به نوع قصد، سیاست‌های بعدی می‌توانند متنوع باشند. نمونه‌هایی از سیاست‌های شبکه عبارتند

از: انجام یک drop، mirroring ترافیک، تغییر مسیر ترافیک مشخص، اولویت‌بندی یک جریان مشخص و غیره. پس از تصمیم‌گیری در مورد عملی که باید انجام شود، محاسبات لازم برای تکمیل اطلاعات سیاست را انجام می‌دهد که به شرح زیر است: چه عملی باید اجرا شود، در کدام رابط شبکه، چگونه باید اجرا شود و سیاست تا چه مدت فعال خواهد بود. پس از تکمیل تمام این اطلاعات، سیاست را هماهنگ کرده و آن را در گذرگاه پیام منتشر می‌کند [20].

- عامل کنترل جریان (FCA). این عامل، قابلیت‌های کنترل ترافیک شبکه را در اختیار صفحه مدیریت قرار می‌دهد. FCA در تبادل سیاست مشترک است و هنگامی که آن را دریافت می‌کند، سیاست را به پیکربندی‌ها و قوانین خاص شبکه ترجمه می‌کند که می‌توانند توسط زیرساخت شبکه در صفحه داده اجرا شوند. FCA در کل زیرساخت توزیع شده است و یک لایه انتزاعی بر روی فناوری‌های مختلف کنترل صفحه داده مانند OpenFlow، SNMP، کنترل ترافیک لینوکس (TC)، Open Virtual Switch (OVS) و iptables است. هنگامی که یک قانون در شبکه اعمال می‌شود، FCA معیارهای مرتبط با چنین قانونی را نیز به صورت دوره‌ای ارائه می‌دهد. توضیح مفصلی در مورد این عامل و عملکرد آن را می‌توان در [24] یافت.

- گردآورندگان داده. تمام داده‌هایی که از طریق گذرگاه پیام در شبکه رد و بدل می‌شوند، مانند توپولوژی، که توسط RIA استخراج می‌شود، معیارهایی که توسط RMA و SMA و معیارهای قانون گزارش شده توسط FCA توسط گردآورندگان داده جمع‌آوری می‌شوند. گردآورندگان داده مسئول تبدیل تمام داده‌ها در پرس‌وجوهای SQL و درج آنها در پایگاه داده به صورت بلادرنگ هستند. همانطور که در شکل 2 نشان داده شده است، برای هر نوع داده یک گردآورنده تخصصی وجود دارد. گردآورندگان اطلاعات منتشر شده توسط عوامل شبکه را استخراج می‌کنند، داده‌ها را تطبیق می‌دهند و آن را در یک پایگاه داده SQL ذخیره می‌کنند. در نتیجه، لایه مدیریت یک پایگاه داده به‌روز با تمام اطلاعات شبکه منتشر شده به صورت بلادرنگ را نگه می‌دارد.

- مولد مجموعه داده. مولد مجموعه داده آخرین جزء مورد نیاز برای داشتن مجموعه داده حاصل است و مسئول ایجاد مجموعه داده است. این یک جزء نرم‌افزاری است که داده‌های ذخیره شده در پایگاه داده SQL را استخراج، شکل‌دهی، مرتب‌سازی و با CSVها (مقادیر جدا شده با کاما) تطبیق می‌دهد. داده‌های استخراج شده ویژگی‌های مختلف مجموعه داده را تشکیل می‌دهند که در بخش 5 توضیح داده خواهند شد.

پس از ارائه مروری بر چارچوب استخراج و جمع‌آوری داده‌ها، اکنون به تشریح الزامات خاص مربوط به مجموعه داده خواهیم پرداخت. جزئیات بیشتر در مورد چارچوب مورد استفاده را می‌توانید در نشریه اخیر ما [25] بیابید.

۴. شبیه‌سازی سناریو و تولید مجموعه داده‌ها

این بخش زیرساخت بستر آزمایشی برای شبیه‌سازی سناریو را شرح می‌دهد و جزئیات نحوه جمع‌آوری داده‌ها برای تولید مجموعه داده‌های IW-IB-5GNET را شرح می‌دهد.

۴.۱. جزئیات پیاده‌سازی

تمام اجزای نرم‌افزاری شرح داده شده در بخش ۳.۳ در یک زیرساخت محاسبات لبه موبایل واقعی G۵ طراحی، مستقر و اعتبارسنجی شده‌اند. اکثریت قریب به اتفاق آنها در جاوا ۱۷ (RIA، SMA، RMA، CPM و Collectors) پیاده‌سازی شده‌اند، به استثنای FCA و تولیدکننده مجموعه داده‌ها که در پایتون ۳.۸ پیاده‌سازی شده‌اند. مؤلفه SMA از Snort 3.0 در زیر برای انجام تشخیص حمله استفاده می‌کند. RIA از مجموعه‌ای از ابزارها شامل OpenStack، OpenAirInterface 5G W44 2022 (یا بالاتر)، LLDP، CDP و iproute2 (نسخه ۱.۹ یا بالاتر) در پشته لینوکس برای تشخیص توپولوژی شبکه استفاده می‌کند. گذرگاه پیام با RabbitMQ 3.6 پیاده‌سازی شده است. پایگاه داده SQL، MySQL 8.15 است. مدیر قوانین شناختی (Cognitive Rule Manager) یک پیاده‌سازی جاوا مبتنی بر موتور MySQL 8 است تا امکان استفاده از SQL را برای آشکارکردن سیاست‌های تحلیلی، تصمیم‌گیری و برنامه‌ریزی فراهم کند. FCA برای اجرای اقدامات به TC qdisk لینوکس، OVS 2.17.3 و iproute2 (نسخه 1.9 یا بالاتر) متکی است. مولد مجموعه داده (Dataset Generator) در پایتون 3.8.10 پیاده‌سازی شده است.

ابزار شبیه‌سازی مورد استفاده برای ایجاد توپولوژی شبکه، شبیه‌ساز تحقیقات باز مشترک (CORE) [26] است. این ابزار از فضاهای نام شبکه لینوکس (netns) برای شبیه‌سازی (به جای شبیه‌سازی) دستگاه‌ها و شبکه‌های مختلفی که زیرساخت را تشکیل می‌دهند، استفاده می‌کند. هر دستگاه یا شبکه در محیط‌های شبکه و پردازش خصوصی خود عمل می‌کند، در حالی که همچنان از همان سیستم فایل و هسته استفاده می‌کند. علاوه بر این، ابزارهای پل زدن اترنت لینوکس موجود در محیط لینوکس، شبیه‌سازی هر نوع شبکه، از جمله شبکه‌های بی‌سیم موبایل را امکان‌پذیر می‌کنند، بنابراین زیرساخت دقیق شرح داده شده در این تحقیق را به طور واقع‌بینانه‌ای نشان می‌دهند. CORE برای پیاده‌سازی سیستمی استفاده شد که امکان ایجاد، پیکربندی، تأمین، شبیه‌سازی و اجرای سناریوهای آزمایشی مختلف در شبکه‌های چند مستاجری G5 را فراهم می‌کند. مشخصات عمیق‌تر سیستم مورد استفاده را می‌توان در [20] یافت.

در زمینه شبکه G5 شبیه‌سازی شده ما، برجسته کردن واقع‌گرایی ترافیک شبکه تولید شده ضروری است. اهمیت این جمله در این واقعیت نهفته است که کل شبکه شبیه‌سازی شده است، به جز لینکی که UE و مؤلفه RAN را به هم متصل می‌کند و شبیه‌سازی شده است. با وجود این عنصر شبیه‌سازی محدود، صحت ترافیک شبکه دست نخورده باقی می‌ماند. این صحت با نمونه‌سازی دقیق از مؤلفه‌ها، پروتکل‌ها و رفتارهای شبکه حفظ می‌شود، که تضمین می‌کند الگوهای ترافیک شبیه‌سازی شده به طور دقیق سناریوهای دنیای واقعی را منعکس می‌کنند. به عنوان مثال، ما از عناصر شبکه اصلی واقعی ارائه شده توسط Osmocom (SGSNEmu [27] و GGSN [28]) و همچنین OpenAirInterface [29] استفاده می‌کنیم. ما همچنین زیرساخت چند-مستاجری را با استفاده از یک کنترلر SDN سفارشی Openstack Neutron-like که شبکه‌های مستاجر ایزوله را با استفاده از OpenVSwitch پر می‌کند، شبیه‌سازی می‌کنیم. علاوه بر این، ترافیک به گونه‌ای مدل‌سازی می‌شود که از طریق پروتکل‌های تونل‌سازی مورد استفاده در معماری‌های G5 مانند VXLAN و GTP، از هر دو

قابلیت تحرک و چند-مستاجری پشتیبانی کند. شبیه‌سازی کامل توپولوژی End-to-End با استفاده از کانتینرهای لینوکس برای انجام استقرار هر یک از عملکردهای شبکه در دستگاه‌های شبیه‌سازی شده مربوطه انجام شده است تا یک استقرار چند-مستاجری G5 واقع‌بینانه ایجاد شود. بنابراین، تمام داده‌هایی که از شبکه G5 شبیه‌سازی شده ما عبور می‌کنند، ترافیک شبکه واقعی (غیرمصنوعی) را منعکس می‌کنند و محیط شبیه‌سازی شده به طور دقیق پویایی شبکه دنیای واقعی را نشان می‌دهد. این رویکرد نه تنها آزمایش و تجزیه و تحلیل قوی را تسهیل می‌کند، بلکه اعتبار شبیه‌سازی ما را به عنوان ابزاری ارزشمند برای ارزیابی و آزمایش شبکه تقویت می‌کند.

آزمایش‌ها در یک ماشین فیزیکی با توزیع اوبونتو نسخه LTS 20.04 با نسخه هسته 5.15.0 اجرا شدند. از نظر منابع فیزیکی، دارای یک پردازنده 56 هسته‌ای Intel(R) Xeon(R) E5-2660 v4 با فرکانس 2.00 گیگاهرتز و 128 گیگابایت DDR4 با فرکانس 2400 مگاهرتز است.

4.2. طراحی آزمایش

امنیت شبکه یکی از مهمترین نگرانی‌های اپراتورهای G5 است [30]. به همین دلیل، ما تصمیم گرفتیم بر امنیت تمرکز کنیم و مجموعه داده‌ها را بر اساس هدف زیر جمع‌آوری کنیم: "از بین بردن هرگونه جریان ترافیکی که رفتار مخرب یا تلاش‌های دسترسی غیرمجاز را نشان می‌دهد، بدون ایجاد اختلال در ترافیک مشروع شبکه". طبق گزارش Cloudflare [31]، حملات انکار سرویس توزیع‌شده ٪38.18 (DDoS) از ترافیک حمله شبکه جهانی و لایه کاربرد را تشکیل می‌دهد. علاوه بر این، محبوب‌ترین نوع حمله DDoS، UDP است که ٪54.4 از کل را تشکیل می‌دهد. مطابق با این واقعیت، ما تصمیم گرفتیم شبکه خود را در معرض حملات UDP DDoS قرار دهیم، به طوری که مجموعه داده‌های حاصل، وضعیت شبکه را در حین انجام این هدف ثبت کنند. بنابراین، مجموعه داده‌های استخراج‌شده می‌توانند برای تولید ماژول‌های هوش مصنوعی که قادر به تصمیم‌گیری بهینه در طول فرآیند هدف هستند، مورد استفاده قرار گیرند. این تصمیمات بهینه خواهند بود زیرا سیاست خاص شبکه می‌تواند مطابق با وضعیت شبکه در هر زمان معین تولید شود ، نه به صورت پیش‌فرض.

آزمایش‌های متعددی برای دستیابی به طیف گسترده‌ای از داده‌ها طراحی و اجرا شده‌اند. پارامترهای مورد مطالعه در تحقیق ما به شرح زیر است: نوع سناریوی اجرا شده، نوع سیاستی که باید در شبکه اجرا شود، فناوری صفحه داده مورد استفاده برای اجرای سیاست، نرخ بسته‌ای که داده‌ها با آن منتقل می‌شوند و اندازه بسته مورد استفاده در این انتقال‌ها.

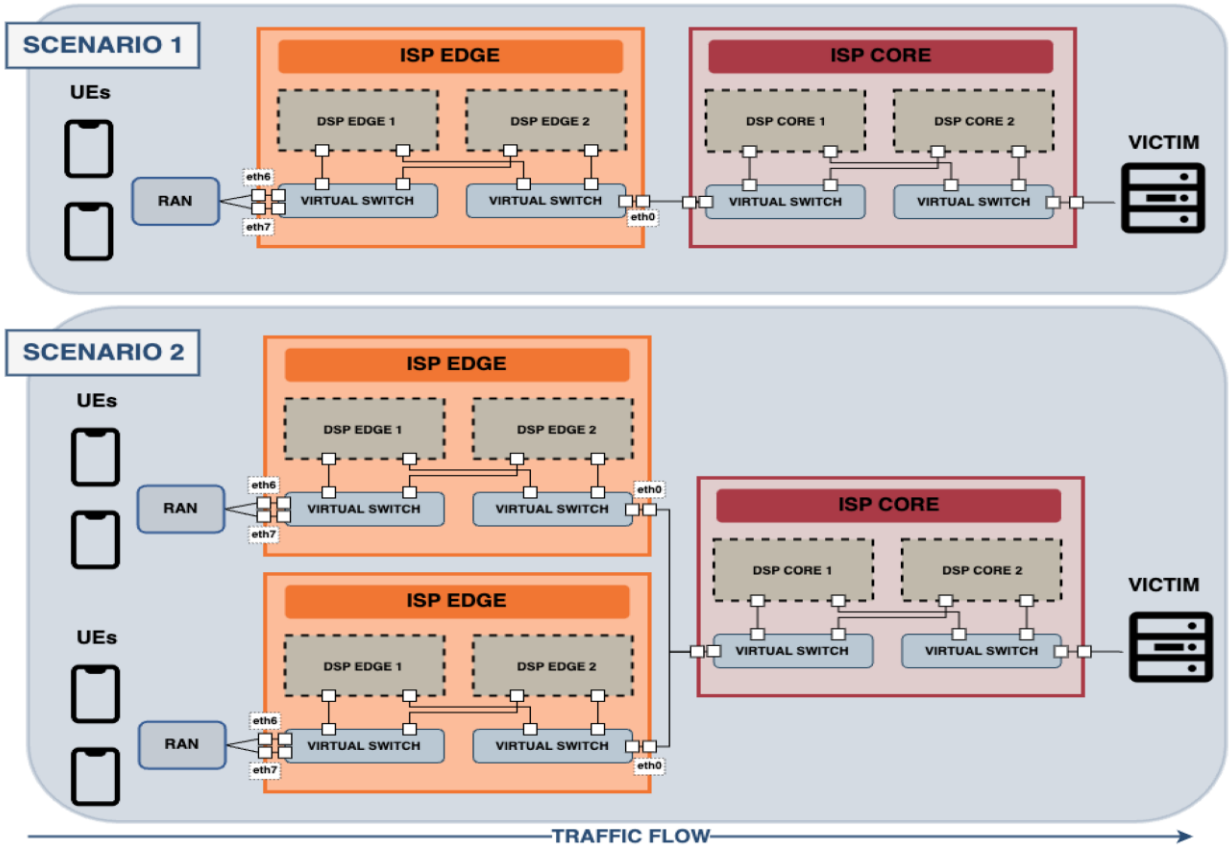
در مجموع چهار سناریوی اساساً متفاوت برای دستیابی به یک مجموعه داده کامل‌تر توسعه داده شده است. این چهار سناریو از نظر تعداد لبه‌ها و همچنین تعداد UEهای متصل به هر لبه تفاوت قابل توجهی دارند. این موارد در زیر شرح داده شده‌اند:

سناریو ۱. از دو UE و یک لبه تشکیل شده است. بنابراین، هر دو UE به یک لبه متصل هستند.

سناریو ۲. از چهار UE و دو لبه تشکیل شده است. دو UE به هر لبه متصل هستند.

سناریو ۳. دارای هشت UE و دو لبه است. بنابراین، چهار UE به هر لبه متصل هستند. سناریو ۴. این سناریو از شانزده UE و دو لبه تشکیل شده است. هشت UE به هر لبه متصل هستند.

طراحی سناریوهای ۱ و ۲ در شکل ۳ نمایش داده شده است. سناریوهای ۳ و ۴ از همان طراحی سناریو ۲ پیروی می‌کنند، از جمله تعداد بیشتری UE متصل به هر لبه. آنها برای سادگی در شکل اضافه نشده‌اند. به همین دلیل، لایه‌های کنترل شبکه نیز در شکل گنجانده نشده‌اند. توجه داشته باشید که تعداد UE های متصل به RAN برای موارد استفاده خاص انتخاب شده است تا تنوع داشته باشند. با این حال، این تعداد و همچنین تعداد لبه‌های متصل به شبکه اصلی قابل گسترش است. این گسترش معماری با استفاده از شبیه‌ساز CORE که یک شبکه با مقیاس‌پذیری گسترده است، به راحتی قابل دستیابی است.



نوع اقدام، جنبه‌های عملکرد، رفتار و عملیات شبکه را که ما بر آنها تمرکز خواهیم کرد و از نزدیک بررسی خواهیم کرد، تعیین می‌کند. در این کار حاضر، ما بر رفتار شبکه تمرکز کرده‌ایم و در عین حال هرگونه جریان ترافیکی که رفتار مخرب یا تلاش‌های دسترسی غیرمجاز را نشان می‌دهد، حذف کرده‌ایم. این شامل انتخاب سیاستی است که باید اجرا شود و فناوری نرم‌افزاری سطح داده که با آن آن را بر روی شبکه اعمال می‌کنیم. این تصمیمات عمیقاً بر رفتار شبکه تأثیر خواهند گذاشت. سه فناوری سطح داده نرم‌افزار شبکه در کار تحقیقاتی ما مورد مطالعه قرار گرفته‌اند که عبارتند از iptables،

OVS و TC. در نهایت، تغییرات در نرخ و اندازه بسته‌ها مستقیماً بر معیارهای عملکرد از نظر توان عملیاتی، از دست دادن بسته‌ها، استفاده از پهنای باند و تراکم تأثیر می‌گذارد. به همین دلیل، تغییرات بسته‌هایی که از شبکه عبور می‌کنند شامل تغییر (۱) اندازه بسته: ۳۲، ۱۲۸، ۲۵۶، ۵۱۲ و ۱۰۲۴ بایت و (۲) نرخ بسته: ۵۰ و ۱۰۰ بسته در ثانیه در هر واحد داده است. این نرخ بسته‌ها در هر ثانیه که به قربانی می‌رسند، بین ۱۰۰ در کمترین حالت و ۱۶۰۰ در بیشترین حالت (۱۶ UE و ۱۰۰ بسته در ثانیه) است که طبق [32]، حملات DDoS با نرخ پایین در نظر گرفته می‌شوند. به طور خاص، حملات DDoS با نرخ پایین ثابت ایجاد شدند. این نوع حمله شامل ارسال بسته‌ها با نرخ پایین ثابت است [33]. تغییر این دو پارامتر، همراه با تعداد UEها، ۴۰ اجرا برای هر فناوری صفحه داده ایجاد می‌کند که در مجموع ۱۲۰ اجرا برای جمع‌آوری داده‌ها برای مجموعه داده‌ها حاصل می‌شود. لازم به ذکر است که بسته به تعداد UEها، هر اجرا شامل تعداد متفاوتی از فایل‌های داده خواهد بود که بسته به تعداد قوانین شبکه فعال در شبکه است. این بدان معناست که هرچه UEهای بیشتری به شبکه حمله کنند، سیاست‌های شبکه بیشتری از هدف برای حذف ترافیک مخرب پردازش می‌شوند.

۴.۳. اجرای آزمایش

این بخش نحوه تولید و جمع‌آوری داده‌ها را برای ایجاد مجموعه داده‌ها توضیح می‌دهد. همانطور که در بخش قبل ذکر شد، هدف نهایی هر آزمایش، جمع‌آوری داده‌ها از نقاط مختلف شبکه در حین حذف ترافیک مخرب از شبکه است. پیکربندی هر آزمایش یکسان است. مدت زمان همه آزمایش‌ها ۳ دقیقه است. پس از تنظیم تمام پارامترهای مشخص شده در بخش قبل و اجرای آزمایش، این مراحل توسط سیستم خودگردان برای تشکیل مجموعه داده‌ها دنبال می‌شود:

هر جزء نرم‌افزاری درگیر در حلقه بسته شناختی (که در بخش ۳.۳ توضیح داده شده است) شروع به کار می‌کند و در حالت آماده به کار عمل می‌کند.

هدف با استفاده از یک رابط خط فرمان (CLI) در شبکه وارد می‌شود (به مرحله ۱ "پروفایلینگ هدف" در شکل ۲ مراجعه کنید). هدف در هر آزمایش یکسان است: "حذف هرگونه جریان ترافیکی که رفتار مخرب یا تلاش‌های دسترسی غیرمجاز را نشان می‌دهد، بدون ایجاد اختلال در ترافیک قانونی شبکه".

CPM هدف را به یک الگوی سیاست ترجمه می‌کند (به مرحله ۲ "ترجمه هدف" در شکل ۲ مراجعه کنید) و منتظر هرگونه هشدار ترافیکی می‌ماند.

UEها شروع به ارسال دو نوع ترافیک می‌کنند. Bonesi [34] ابزاری است که برای ترافیک حمله DDoS استفاده می‌شود، در حالی که hping3 برای ترافیک بی‌خطر استفاده می‌شود (شکل 3 را ببینید).

اجزای نرم‌افزاری در سرویس و لایه‌های محاسباتی لبه ISP شروع به استخراج داده‌ها می‌کنند، زیرا ترافیک از UEها از شبکه عبور می‌کند. داده‌های مبادله شده در شکل 2 با دایره‌های قرمز نشان داده شده‌اند.

SMA ترافیک مخرب را تشخیص می‌دهد و سیستم را از حمله از طریق گذرگاه پیام، با استفاده از تبادل هشدار ترافیک، مطلع می‌کند.

CPM تمام مراحل شرح داده شده در بخش 3.3 را انجام می‌دهد و یک سیاست شبکه ایجاد می‌کند. سپس در گذرگاه پیام منتشر می‌شود. این سیاست مشخص می‌کند که چه کاری، کجا و چگونه انجام شود. در آزمایش ما، یک عمل حذف در یک رابط شبکه با استفاده از فناوری صفحه داده مشخص انجام می‌شود. تصمیم فناوری صفحه داده قبلاً در پیکربندی آزمایش تعریف شده است و در الگوی سیاست گنجانده شده است. از سوی دیگر، تصمیم مکان (رابط شبکه) برای انجام حذف توسط حلقه شناختی محاسبه می‌شود. بنابراین، بسته به فناوری سطح داده‌ای که قرار است استفاده شود، عمل حذف روی یک رابط شبکه خاص انجام خواهد شد. در سناریوی خاص ما، از آنجایی که ما با یک حمله DDoS مقابله می‌کنیم، مطلوب است که آن را در اسرع وقت متوقف کنیم. این بدان معناست که عمل حذف باید تا حد امکان نزدیک به UEها انجام شود تا ترافیک مخرب از شبکه عبور نکند. شکل 3 سه رابط شبکه مختلف را نشان می‌دهد که در آنها امکان انجام عمل حذف وجود دارد. فناوری‌های سطح داده OVS و TC در رابط‌های eth0، eth6 و eth7 در دسترس هستند. با این حال، iptables فقط در eth0 در دسترس خواهد بود. به همین دلیل، هنگامی که فناوری OVS یا TC انتخاب می‌شود، حذف به ترتیب در eth6 و eth7 انجام می‌شود. از سوی دیگر، در صورت انتخاب iptables، عمل حذف در eth0 انجام می‌شود زیرا اجرای این عمل در شبکه زودتر امکان‌پذیر نیست. پس از همه این تصمیمات و سیاست پس از اتمام ، مرحله 3 "تشخیص هدف" تکمیل می‌شود.

خط‌مشی در گذرگاه پیام منتشر می‌شود. نمونه‌ای از پیام خط‌مشی در فهرست 1 نشان داده شده است.

FCA خط‌مشی را دریافت کرده و آن را به یک قانون شبکه تبدیل می‌کند که هدف آن حذف ترافیک مخرب است. همانطور که در فهرست 1 نشان داده شده است، خط‌مشی همچنین مشخص کرده است که کدام رابط شبکه (eth6) و با چه فناوری سطح داده‌ای (TRAFFIC_CONTROL) این قانون را اجرا کند. همانطور که گفته شد، FCA قادر به انجام عمل حذف با فناوری‌های سطح داده زیر است: iptables، OVS یا TC. این مرحله 4 "فعال‌سازی هدف" در شکل 2 را تکمیل می‌کند.

پس از اجرای قانون در شبکه، ترافیک مخرب توسط فناوری سطح داده‌ای که در خط‌مشی تعریف شده است، در رابط شبکه مشخص شده حذف می‌شود. در همین حال، همه اجزا به گزارش معیارهای رفتار شبکه ("تضمین هدف") ادامه می‌دهند. با این مرحله آخر، می‌توانیم حلقه را با موفقیت بسته شده در نظر بگیریم زیرا اکنون شبکه به حالت آماده به کار خود باز می‌گردد.

اجرای آزمایش در 3 دقیقه ادامه می‌یابد. در طول این مدت، UEها به ارسال ترافیک مخرب ادامه می‌دهند. چنین ترافیکی در لبه شبکه متوقف می‌شود. در این میان، تمام اطلاعات و معیارهای تولید شده در طول اجرای آزمایش در پایگاه داده ذخیره شده و به مجموعه داده حاصل تبدیل شده‌اند.

Listing 1. Example of network policy.

```
{
        “Policy”:    {
              “actionType”: “INSERT”,
              “actionName”: “DROP”,
              “priority”: 1,
              “flowId”: “4BA92944”,
              “reportedTime”: 1659106981511
        },
        “Params”:    [
              {
                    “paramName”: “interfaceName”,
                    “paramValue”: “eth6”
              },
              {
                    “paramName”: “technology”,
                    “paramValue”: “TRAFFIC_CONTROL”
              },
              {
                    “paramName”: “device”,
                    “paramValue”: “edge1”
              }
        ]
}
```

۵. شرح مجموعه داده

این بخش شرح مفصلی از مجموعه داده پیشنهادی و ایجاد شده در این تحقیق، با تأکید بر ساختار آن، نوع داده‌های جمع‌آوری‌شده و ویژگی‌های نمونه‌ها و ویژگی‌های آن ارائه می‌دهد.

۵.۱. مجموعه داده IW-IB-5GNET

مجموعه داده IW-IB-5GNET از چندین فایل تشکیل شده است. از آنجایی که این یک مجموعه داده مبتنی بر هدف است، هر فایل با یک سیاست شبکه مرتبط است که از یک هدف ترجمه شده است. در نتیجه، تعداد فایل‌های تولید شده با تعداد کل سیاست‌های شبکه فعال در هر آزمایش مطابقت خواهد داشت. برای هر فایل، ویژگی‌ها از نظر توپولوژیکی مرتب شده‌اند. این بدان معناست که ویژگی‌ها با رویکردی از بالا به پایین سازماندهی شده‌اند، که با ویژگی‌های دستگاه شروع می‌شود و به دنبال آن ویژگی‌های رابط شبکه، ویژگی‌های فناوری صفحه داده، صف‌ها، جریان‌های شبکه و قوانین شبکه قرار می‌گیرند. تجمیع ویژگی‌های مربوط به تمام سطوح توپولوژی شبکه G5، آن را به یک مجموعه داده در سطح زیرساخت تبدیل می‌کند.

پس از اجرای هر آزمایش شرح داده شده در بخش ۴، تمام نمونه‌های هر فایل حاصل در یک فایل csv واحد ادغام شده‌اند. این فایل، مجموعه داده IW-IB-5GNET را تشکیل می‌دهد. در نتیجه ۱۲۰ اجرا، در مجموع ۷۰۰ فایل csv جمع‌آوری شده است. این مجموعه داده دارای ابعاد نهایی ۶۴۲۹۰ × ۱۰۷ است، یعنی در مجموع ۶۴۲۹۰ نمونه و ۱۰۷ ویژگی. گزیده‌ای از مجموعه داده شامل ۱۰۰۰ نمونه به صورت آنلاین در دسترس است (به مطالب تکمیلی در انتهای سند مراجعه کنید).

۵.۲. شرح ویژگی‌ها

مجموعه داده IW-IB-5GNET شامل ۱۰۷ ویژگی مربوط به لبه و هسته یک شبکه G۵ است. جدول ۲ تمام ویژگی‌ها و موقعیت‌های آنها را در مجموعه داده فهرست می‌کند. ویژگی‌ها از نظر توپولوژیکی مرتب شده‌اند، با شروع از فرادادهها و معیارهای مرتبط با دستگاه (به موارد ۱ تا ۳ با رنگ خاکستری در جدول ۲ مراجعه کنید)، و پس از آن پورت دستگاه (به موارد ۴ تا ۵، زرد مراجعه کنید)، فناوری‌های صفحه داده و صف‌های آنها (به موارد ۶ تا ۳۳، آبی مراجعه کنید)، جریان‌های ترافیک (به موارد ۳۴ تا ۴۹، قرمز مراجعه کنید) و در نهایت، قوانین کنترل شبکه (به موارد ۵۰ تا ۱۰۶، سبز و نارنجی مراجعه کنید) قرار دارند. ویژگی‌های ۵۰ تا ۶۱ شامل معیارها و فرادادههای یک قانون شبکه خاص هستند که بر روی یک رابط شبکه خاص (مشخص شده در ویژگی ۴، iface) اجرا و نظارت می‌شوند. ویژگی‌های ۶۲ تا ۱۰۶ شامل اطلاعات تحلیلی مربوط به تمام قوانین شبکه‌ای هستند که در حال حاضر در شبکه اجرا می‌شوند. در مورد ویژگی‌های RAN، لازم به ذکر است که ما تصمیم گرفته‌ایم هیچ مقداری را که از یک منبع قابل اعتماد نمی‌آید، لحاظ نکنیم. بنابراین، هیچ ویژگی لایه فیزیکی گنجانده نشده است، زیرا تنها لینک در شبکه است که شبیه‌سازی می‌شود.

جدول ۲. لیست ویژگی‌ها در مجموعه داده IW-IB-5GNET. رنگ، سطح توپولوژی را نشان می‌دهد: (خاکستری-دستگاه)، (زرد-رابط)، (آبی-فناوری)، (قرمز-جریان) و (سبز، نارنجی-سیاست).

Table 2. List of features in IW-IB-5GNET dataset. Color indicates topology level: (grey-device), (yellow-interface), (blue-technology), (red-flow) and (green, orange-policy).

No.	Feature	No.	Feature	No.	Feature	No.	Feature
1	Hostname	28	TC_rx_bytes	55	currentMatchedPackets	82	OVS_mean_crr
2	ContextSwitchesPerSecond	29	TC_RX_dropped	56	lastMatchedTime	83	OVS_median_crr
3	AbstractionLayer	30	TC_RX_packets	57	ruleComplexity	84	OVS_st_crr
4	Iface	31	TC_TX_bytes	58	totalMatchedBytes	85	OVS_q1_crr
5	Iface_speed	32	TC_TX_dropped	59	totalMatchedPkts	86	OVS_q3_crr
6	IPTAB_activated	33	TC_TX_packets	60	actionType	87	OVS_mean_total
7	IPTAB_complexity	34	encapsulationLayer	61	actionName	88	OVS_median_total
8	IPTAB_maxRules	35	encapsulationType1	62	IPTAB_min_crr_currentMatchedPkts	89	OVS_st_total
9	IPTAB_rx_bytes	36	encapsulationType2	63	IPTAB_max_crr_currentMatchedPkts	90	OVS_q1_total
10	IPTAB_rx_packets	37	sense	64	IPTAB_min_total_totalMatchedPkts	91	OVS_q3_total
11	IPTAB_tx_bytes	38	l3Protocol	65	IPTAB_max_total_totalMatchedPkts	92	TC_min_crr_currentMatchedPkts
12	IPTAB_tx_packets	39	dstIP	66	IPTAB_numberRulesActivatedTotal	93	TC_max_crr_currentMatchedPkts
13	OVS_activated	40	macSrc	67	IPTAB_mean_crr	94	TC_min_total_totalMatchedPkts
14	OVS_complexity	41	macDst	68	IPTAB_median_crr	95	TC_max_total_totalMatchedPkts
15	OVS_maxRules	42	l4Protocol	69	IPTAB_st_crr	96	TC_numberRulesActivatedTotal
16	OVS_rx_bytes	43	tos	70	IPTAB_q1_crr	97	TC_mean_crr
17	OVS_rx_dropped	44	outTos	71	IPTAB_q3_crr	98	TC_median_crr
18	OVS_rx_packets	45	dstPort	72	IPTAB_mean_total	99	TC_st_crr
19	OVS_tx_bytes	46	state	73	IPTAB_median_total	100	TC_q1_crr
20	OVS_tx_dropped	47	totalpktCount	74	IPTAB_st_total	101	TC_q3_crr
21	OVS_tx_packets	48	totalBits	75	IPTAB_q1_total	102	TC_mean_total
22	TC_activated	49	packetSize	76	IPTAB_q3_total	103	TC_median_total
23	TC_queueDiscipline	50	programmableTechnology	77	OVS_min_crr_currentMatchedPkts	104	TC_st_total
24	TC_queueLenght	51	activatedRuleTimeSecs	78	OVS_max_crr_currentMatchedPkts	105	TC_q1_total
25	TC_complexity	52	averageMatchedBytes	79	OVS_min_total_totalMatchedPkts	106	TC_q3_total
26	TC_crr_bwd_guaranteed	53	averageMatchedPackets	80	OVS_max_total_totalMatchedPkts	107	timestamp
27	TC_maxRules	54	currentMatchedBytes	81	OVS_numberRulesActivatedTotal		

Table 3. Description of boolean features in IW-IB-5GNET dataset.

Feature Name	Representation	Description
IPTAB_activated	[True, False]	Presence of iptables in the interface.
OVS_activated	[True, False]	Presence of ovs in the interface.
TC_activated	[True, False]	Presence of linux tc in the interface.
Sense	[ingress, egress]	Direction of network traffic flow.

بر اساس ماهیت و نوع اطلاعاتی که ارائه می‌دهند، ویژگی‌ها را می‌توان به چهار نوع مختلف طبقه‌بندی کرد که در زیر فهرست شده‌اند.

ویژگی‌های بولی. آن‌ها به مقادیر دودویی اشاره دارند که نشان‌دهنده وجود یا عدم وجود یک ویژگی خاص هستند. مجموعه داده IW-IB-5GNET در مجموع چهار ویژگی بولی دارد که در جدول 3 فهرست شده‌اند.

ویژگی‌های فراداده. آن‌ها از ویژگی‌های دسته‌بندی‌شده، عددی و متنی تشکیل شده‌اند. آن‌ها در یک دسته با هم ترکیب می‌شوند زیرا ویژگی‌های هر آزمایش خاص را نشان می‌دهند. اکثر آن‌ها ویژگی‌های اساسی شبکه را توصیف می‌کنند و برای درک هر مورد استفاده خاص ضروری هستند. علاوه بر این، اکثر مقادیر آن در طول آزمایش‌ها تغییر نمی‌کنند. با این حال، مهم است که آن‌ها را در مجموعه داده نگه دارید تا توصیف کاملی از شبکه در هر اجرای آزمایش خاص داشته باشید.
ویژگی‌های فراداده در جدول 4 فهرست شده‌اند. جدول نوع داده هر ویژگی را شرح می‌دهد. در مورد ویژگی‌های دسته‌بندی‌شده، مقادیری که می‌توانند به دست آورند مشخص شده است. علاوه بر این، شرح مختصری از آنچه هر یک از این ویژگی‌ها نشان می‌دهند، گنجانده شده است.

ویژگی‌های عددی. آن‌ها مقادیر عددی پیوسته یا گسسته را نشان می‌دهند. در مجموع ۷۴ ویژگی عددی در مجموعه داده IW-IB-5GNET وجود دارد. چنین ویژگی‌هایی مربوط به معیارهای اندازه‌گیری شده در زمان واقعی، در سطوح مختلف توپولوژی شبکه هستند: میزبان دستگاه، رابط‌ها، جریان‌ها، فناوری‌های صفحه داده، صف‌ها و قوانین. نام‌های داده شده به این معیارها به اندازه کافی توصیفی هستند تا خواننده بداند که آنها چه چیزی را نشان می‌دهند.

ویژگی‌های تاریخ/زمان. آنها نقاط خاصی را در زمان نشان می‌دهند. تنها یک ویژگی تاریخ در مجموعه داده IW-IB-5GNET وجود دارد، مهر زمانی، که نشان دهنده لحظه‌ای است که استخراج مشخصی انجام شده است. این ویژگی با استفاده از مهر زمانی یونیکس نمایش داده می‌شود.

برای نتیجه‌گیری از توصیف داده‌ها، همانطور که خواننده می‌تواند مشاهده کند، هیچ هدف خاصی مرتبط با مجموعه داده‌ها وجود ندارد. این تصمیم آگاهانه برای اطمینان از تطبیق‌پذیری و سازگاری مجموعه داده‌ها با موارد استفاده مختلف کنترل و مدیریت شبکه گرفته شده است. با برچسب‌گذاری نکردن مجموعه داده‌ها با یک هدف خاص، آزادی استفاده از آن برای اهداف مختلف مدیریت و بهینه‌سازی شبکه، همانطور که در بخش 1 توضیح داده شده است، فراهم می‌شود. این رویکرد به ما امکان می‌دهد مجموعه داده‌ها را برای رسیدگی به طیف وسیعی از نیازهای خاص بررسی و به کار ببریم و نوآوری و انعطاف‌پذیری را در شیوه‌های مدیریت شبکه تقویت کنیم. جدول 5 شامل مجموعه‌ای از موارد استفاده است که مجموعه داده‌های ما می‌تواند برای آنها استفاده شود. این جدول هر مورد استفاده و همچنین هدف کلی آن: مدیریت و بهینه‌سازی را شرح می‌دهد. علاوه بر این، مشخص شده است که بسته به مورد استفاده، کدام برچسب را به ستون هدف اختصاص دهیم. در نهایت، ویژگی‌های مختلفی که، از قبل، می‌توانند هنگام تجزیه و تحلیل هر مورد استفاده خاص مرتبط‌تر باشند، مورد تأکید قرار گرفته‌اند. توجه داشته باشید که این بدان معنا نیست که این ویژگی‌ها تنها ویژگی‌های مهم هستند، بلکه ما سعی می‌کنیم تطبیق‌پذیری مجموعه داده‌ها را از نظر ویژگی‌های آن هنگام برخورد با موارد استفاده مختلف منعکس کنیم. اگرچه ساختار مجموعه داده‌ها برای همه این موارد استفاده اعمال می‌شود، داده‌های فعلی به‌دست‌آمده با آزمایش‌های شرح داده شده در بخش ۴ را می‌توان برای مدیریت امنیت و بهینه‌سازی QoS استفاده کرد. برای بقیه مثال‌های ارائه شده در جدول ۵، آزمایش‌های خاص‌تری لازم است.

Table 4. Description of metadata features in W-IB-5GNET dataset.

Feature Name	Data Type	Categorical Values	Additional Info
Hostname	text	-	Data extraction host name.
AbstractionLayer	number	[0, 1, 2]	Level of virtualization.
Iface	text	-	Interface name where data extraction was performed.
Iface_speed	number	-	Network interface speed.
IPTAB_complexity	number	[8]	Level of complexity iptables rules.
IPTAB_maxRules	number	[4096]	iptables rule limit.
OVS_complexity	number	[4]	Level of complexity ovs rules.
OVS_maxRules	number	[16,384]	ovs rule limit.
TC_queueDiscipline	text	[noqueue, fq_codel, atm, htb, prio]	Primary iface qdisc queue discipline.
TC_queueLenght	number	-	Max number of packets allowed in the queue.
TC_complexity	number	[8]	Level of complexity tc rules.
TC_maxRules	number	[4096]	tc rule limit.
encapsulationLayer	number	[0, 1, 2]	Number of encapsulations of a traffic flow.
encapsulationType1	number	[gtp, vxlan]	Type of first encapsulation.
encapsulationType2	number	[gtp, vxlan]	Type of second encapsulation.
L3Protocol	number	[ipv4, ipv6, icmp, arp]	Layer 3 protocol of flow.
dstIP	number	-	Flow destination IP.
macSrc	number	-	Flow source mac address.
macDst	number	-	Flow destination mac address.
L4Protocol	number	[tcp, udp]	Layer 4 protocol of flow.
Tos	number	[0]	Type of service.
OutTos	number	[0]	Out type of service.
dstPort	number	-	Flow destination port.
State	text	[active, dropped, inactive]	Flow state description.
programmableTechnology	text	[TC, OVS, IPTABLES]	Data-plane technology used to do the action.
ruleComplexity	number	[1, 2, 3. . .]	Rule performance complexity.
actionType	text	[INSERT, SET, DELETE]	Definition of action type.
actionName	text	[DROP, PRIORITY, QUEUE, SLICE]	Definition of action name.

Table 5. Use-case examples in which to use the IW-IB-5GNET dataset.

	Purpose	Use-Case Description	Target	Relevant Features
Management	Security	Anomaly detection. Monitoring of unusual activities to respond to security incidents.	0: Benign traffic 1: Malicious traffic	34 to 49
	Traffic	Balancing traffic load across different network interfaces to prevent congestion.	Interface	4 to 49
	QoS	Defining QoS policies according to the status of the network and active network policies based on user intents.	QoS policy definition	6 to 33 62 to 106
Optimisation	Resource	Detecting redundant, unused rules to reduce processing overhead and policy congestion.	0: keep active rule 1: delete/change active rule	50 to 106
	QoS	Predictive modeling to determine the best technology/method for implementing and enforcing active network policies.	Optimal technology for active policy	34 to 61

۶. نتایج

در این بخش، تحلیل و اعتبارسنجی مجموعه داده‌های حاصل را ارائه می‌دهیم. این بخش داده‌های به‌دست‌آمده از طریق جداول و نمودارها را ارائه می‌دهد که می‌تواند به درک بهتر مجموعه داده‌ها کمک کند. هدف نهایی این بخش، نشان دادن کیفیت و قابلیت اطمینان مجموعه داده‌های IW-IB-5GNET است. برای توسعه چنین تحلیل‌هایی، ما از پایتون ۳.۹.۶ و کتابخانه‌های آن برای تحلیل داده‌ها استفاده کردیم: Pandas، Numpy، SciPy، Seaborn و Matplotlib.

۶.۱. پیش‌پردازش مجموعه داده‌ها

قبل از ارزیابی عملکرد مجموعه داده‌های IW-IB-5GNET، ما برخی پیش‌پردازش‌های داده‌ها را انجام دادیم تا اطمینان حاصل شود که مجموعه داده‌ها تمیز، سازگار و مناسب برای تحلیل هستند. این مرحله مقدماتی برای کاهش تأثیر بالقوه نویز، خطاها در طول اجرای آزمایش و بی‌نظمی‌ها در داده‌ها، که می‌تواند به طور قابل توجهی بر دقت و قابلیت اطمینان تحلیل‌های بعدی تأثیر بگذارد، ضروری بود.

ابتدا، همانطور که در بخش ۵ ذکر شد، لازم بود تمام فایل‌های CSV از آزمایش‌های مختلف در یک فایل واحد تجمیع شوند و فرآیند جمع‌آوری آمار و انجام تحلیل ساده شود. وقتی همه داده‌ها در یک فایل CSV واحد قرار گرفتند، چهار مرحله زیر در خط لوله پیش‌پردازش داده‌های ما انجام شد:

تمام ردیف‌هایی که ستون‌هایشان تکراری بودند را حذف کنید. این نشان می‌دهد که تمام ستون‌های آنها مقدار یکسانی دارند.

تمام ردیف‌هایی که شرایط زیر را دارند را حذف کنید: activatedRuleTime = lastMatchedTime. ستون‌هایی که این شرط را برآورده می‌کنند، نتیجه یک عملیات نامعتبر در اجرای آزمایش هستند، زیرا نشان می‌دهند که یک قانون درج شده در شبکه با هیچ بسته‌ای مطابقت نداشته است.

ستون‌هایی را که مقادیر آنها در تمام تکرارها خالی بود، تجزیه و تحلیل و حذف کنید.

تمام ردیف‌هایی را که ستون‌های آنها حاوی مقادیر منفی بودند، حذف کنید. هیچ یک از ویژگی‌ها برای منفی بودن طراحی نشده بودند، بنابراین وجود این مقادیر، در صورت وجود، به دلیل یک عملیات نامعتبر در طول اجرای آزمایش بود.

پیش‌پردازش داده‌ها منجر به مجموعه داده IW-IB-5GNET شد که ابعاد آن قبلاً در بخش قبلی ذکر شده است: 64290 × 107. میزان استفاده از حافظه آن 51.2 مگابایت است. تجزیه و تحلیل انواع مختلف ویژگی‌های ارائه شده در بخش 5 در زیربخش‌های جداگانه در زیر مورد بحث قرار گرفته است. ۶.۲. ارزیابی ویژگی‌های بولی

جدول ۶ خلاصه آماری ویژگی‌های بولی در مجموعه داده‌ها را نشان می‌دهد. این جدول تعداد مقادیر غیر تهی در هر ستون، تعداد مقادیر منحصر به فرد در ستون، بیشترین مقدار تکرار شونده در هر ستون و فراوانی بالاترین مقدار را مشخص

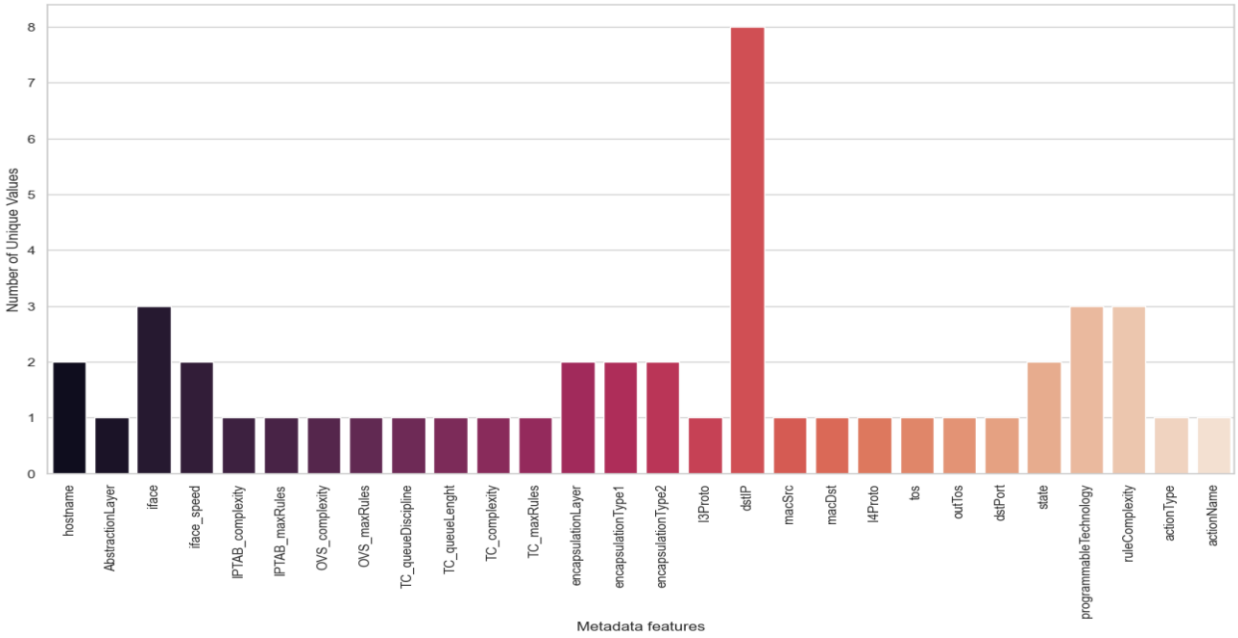
می‌کند. این خلاصه برای درک سریع توزیع و ویژگی‌های داده‌های بولی در مجموعه داده‌های ما مفید است. به عنوان مثال، وجود فناوری‌های صفحه داده OVS و TC را در تمام آزمایش‌های اجرا شده که در آنها عمل حذف انجام شده است، شناسایی می‌کند (به OVS_activated و TC_activated در جدول ۶ مراجعه کنید). از سوی دیگر، نشان می‌دهد که موارد بیشتری وجود دارد که در آنها به دلیل عدم وجود یک قانون در رابط شبکه تحت نظارت، امکان اجرای آن در iptables وجود ندارد (به IPTAB_activated مراجعه کنید). آمار همچنین نشان می‌دهد که بیشتر جریان‌های شبکه حذف شده در ورودی هستند (به بخش مربوطه مراجعه کنید).

Table 6. Statistical values of boolean features in IW-IB-5GNET dataset.

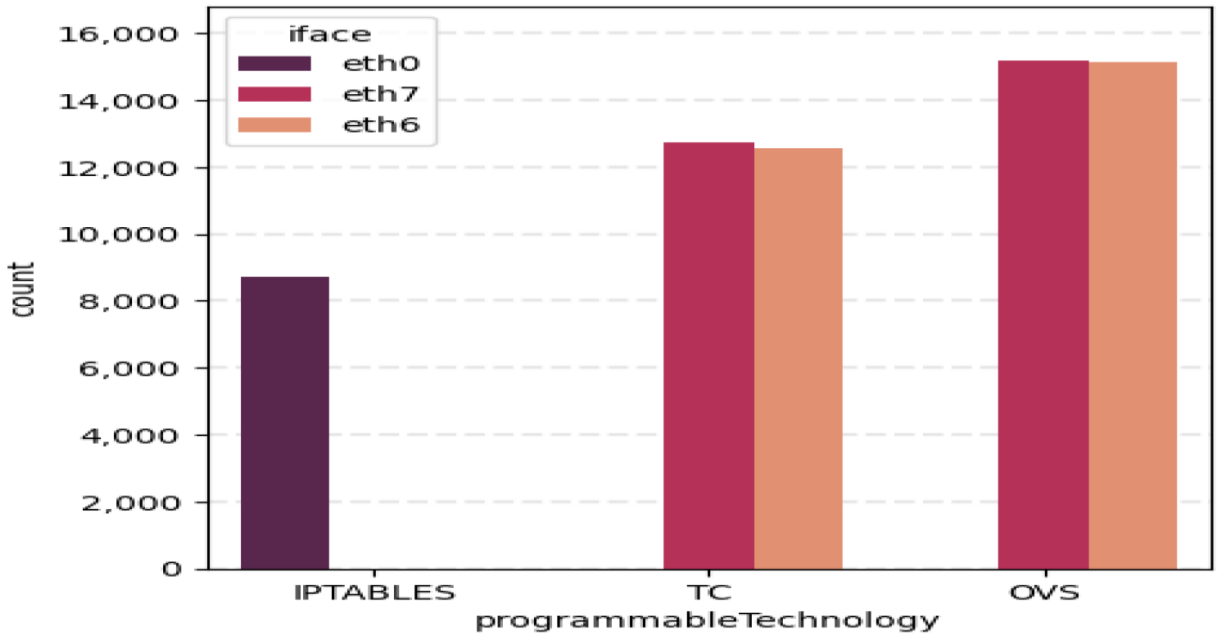
Feature Name	Count	Unique	Top	Freq
IPTAB_activated	64,290	2	False	55,569
OVS_activated	64,290	1	True	64,290
TC_activated	64,290	1	True	64,290
sense	64,290	2	ingress	55,569

۶.۳. ارزیابی ویژگی‌های فراداده

همانطور که در جدول ۴ توضیح داده شده است، ۲۸ ویژگی شامل متغیرهای دسته‌بندی شده به عنوان فراداده در مجموعه داده‌های ما هستند. اکثر آنها درک بهتری از زیرساخت شبکه، ویژگی‌های جریان ترافیک و قانون کنترل شبکه فعال در هر لحظه خاص استخراج به ما ارائه می‌دهند. شکل ۴ تعداد مقادیر منحصر به فرد هر ویژگی ارائه شده در جدول ۴ را نشان می‌دهد. به عنوان مثال، در تمام آزمایش‌ها از سه رابط شبکه مختلف برای حذف جریان‌های ترافیک استفاده شد. از دیگر ویژگی‌های مرتبط، می‌توان به کپسوله‌سازی جریان‌های ترافیک به صورت ارائه شده اشاره کرد. دو نوع کپسوله‌سازی مختلف وجود داشت. فناوری قابل برنامه‌ریزی، فناوری صفحه داده مورد استفاده برای اجرای اقدام در شبکه را به ما می‌دهد. مقدار آن سه است، زیرا ما از سه فناوری صفحه داده مختلف استفاده کرده‌ایم: iptables، OVS و TC. به طور مشابه، سه پیچیدگی قانون مختلف وجود داشت که با اجرای هر فناوری صفحه داده مرتبط است. اکثر این ویژگی‌ها به طور مداوم یک مقدار یکنواخت را در کل مجموعه داده‌ها حفظ می‌کنند. این ثبات ناشی از این واقعیت است که، همانطور که در بخش قبلی توضیح داده شد، فراداده‌ها اطلاعات مربوط به شبکه را در اختیار ما قرار می‌دهند که در طول زمان نسبتاً ثابت باقی می‌مانند. اگرچه این اطلاعات ممکن است فوراً جذاب نباشند، اما حفظ آنها در مجموعه داده‌ها مناسب است. دلیل این امر این است که اگر آنها انواع دیگری از آزمایش‌ها را انجام دهند، این فراداده تغییر می‌کند و ارتباط آن با مجموعه داده‌ها افزایش می‌یابد.



اکنون با تمرکز بر ویژگی‌های خاص، شکل ۵ اطلاعات جالبی را در مورد ارتباط متغیرهای iface و programmableTechnology آشکار می‌کند. نمودار شمارش به سه فناوری صفحه داده (iptables، OVS و TC) که برای اجرای سیاست شبکه استفاده می‌شوند، تقسیم شده است. علاوه بر این، چنین تقسیم‌بندی با در نظر گرفتن محل اجرای این سیاست از نظر رابط شبکه، طبقه‌بندی می‌شود. بنابراین، می‌توانیم مشاهده کنیم که تمام اقدامات انجام شده با iptables در رابط شبکه eth0 اجرا می‌شوند، در حالی که TC و OVS بین eth6 و eth7 متفاوت هستند. این نمودار عملکرد موفقیت‌آمیز حلقه کنترل مستقل را تأیید می‌کند، زیرا سیاست‌های شبکه در رابط‌هایی که در بخش ۴.۳ مورد تجزیه و تحلیل قرار گرفته‌اند، اجرا می‌شوند.



۶.۴. ارزیابی ویژگی‌های عددی

همانطور که در بخش ۵.۲ توضیح داده شد، در مجموع ۷۴ ویژگی، متغیرهای عددی مجموعه داده IW-IB-5GNET را تشکیل می‌دهند. برخی از آنها در طول فرآیند پیش‌پردازش داده‌ها حذف شدند. علاوه بر این، ویژگی‌های ۶۲ تا ۱۰۶ (به جدول ۲ مراجعه کنید) از قبل اطلاعات تحلیلی را نشان می‌دهند. بنابراین، در این تحلیل داده‌ها در نظر گرفته نشدند.

ابتدا، تحلیل داده‌ها را با برخی آمار توصیفی ارائه شده در جدول ۷ شروع می‌کنیم. این جدول آمار ۲۲ ویژگی عددی غیر تهی در مجموعه داده ما را نشان می‌دهد. دو ستون اول نام ویژگی و واحد آن را نشان می‌دهند. آمار شامل میانگین، انحراف معیار، حداقل و حداکثر معیارها است. همانطور که در جدول مشاهده می‌شود، بیشتر این ویژگی‌ها بر نظارت بر بسته‌ها یا بایت‌ها در ثانیه که از نقاط مختلف شبکه عبور می‌کنند، متمرکز هستند. همچنین می‌توانیم اطلاعات مربوط به تعداد بسته‌هایی که در طول آزمایش از شبکه عبور می‌کنند و اندازه آنها بر حسب بایت را مشاهده کنیم. با این اندازه‌گیری‌ها می‌توان نقاط تراکم در امتداد شبکه و همچنین فناوری‌های صفحه داده و رابط شبکه‌ای که در آن قرار دارند را شناسایی کرد. در نهایت، هفت متغیر آخر اطلاعاتی در مورد اثربخشی قانون فعلی اعمال شده بر شبکه مطابق با هدف مورد نظر در اختیار ما قرار می‌دهند. ما این اثربخشی را بر اساس هر یک از فناوری‌های صفحه داده‌ای که عمل حذف با آنها انجام شده است، برجسته کرده‌ایم. این بازتاب را می‌توان در شکل 6 مشاهده کرد. در این شکل، یک نمودار جعبه‌ای برای هر یک از فناوری‌های صفحه داده نشان داده شده است: iptables، TC و OVS. از سوی دیگر، در محور y، تعداد کل بسته‌هایی که با قانون حذف مطابقت دارند، نشان داده شده است. از این نمودار می‌توان نتیجه گرفت که برای مدت زمان مشابه آزمایش (3 دقیقه)، تعداد بسته‌هایی که OVS قادر به پردازش آنهاست، کمی بیشتر از TC و متعاقباً iptables است. این

نتایج نشان می‌دهد که OVS از نظر زمان پردازش قانون شبکه سریع‌تر است و می‌توان آن را هنگام انجام وظایف بهینه‌سازی شبکه در نظر گرفت.

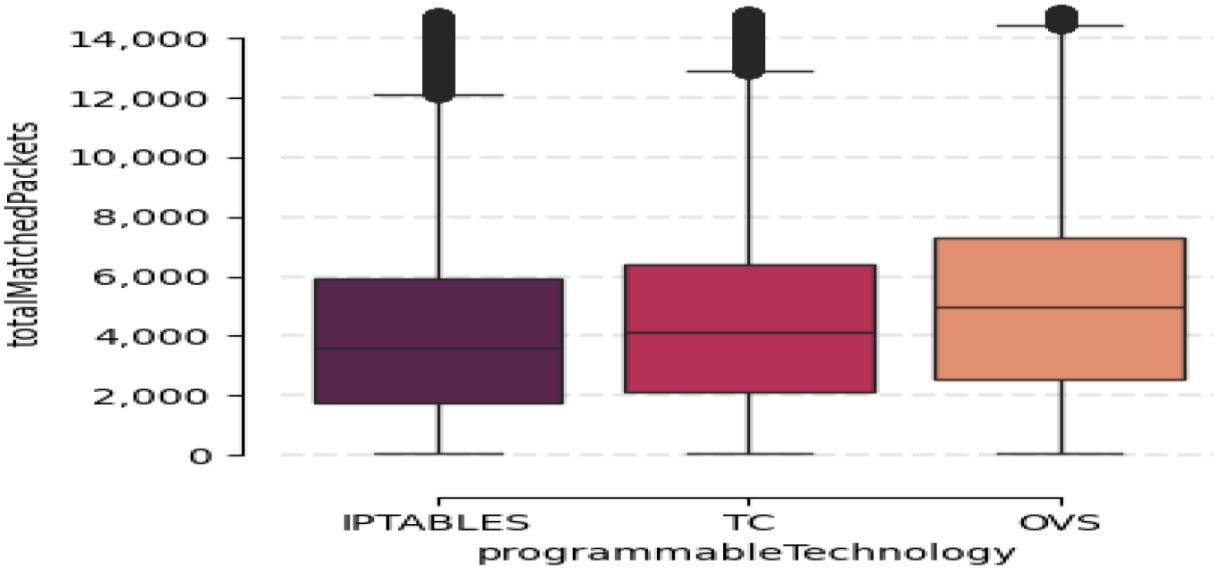
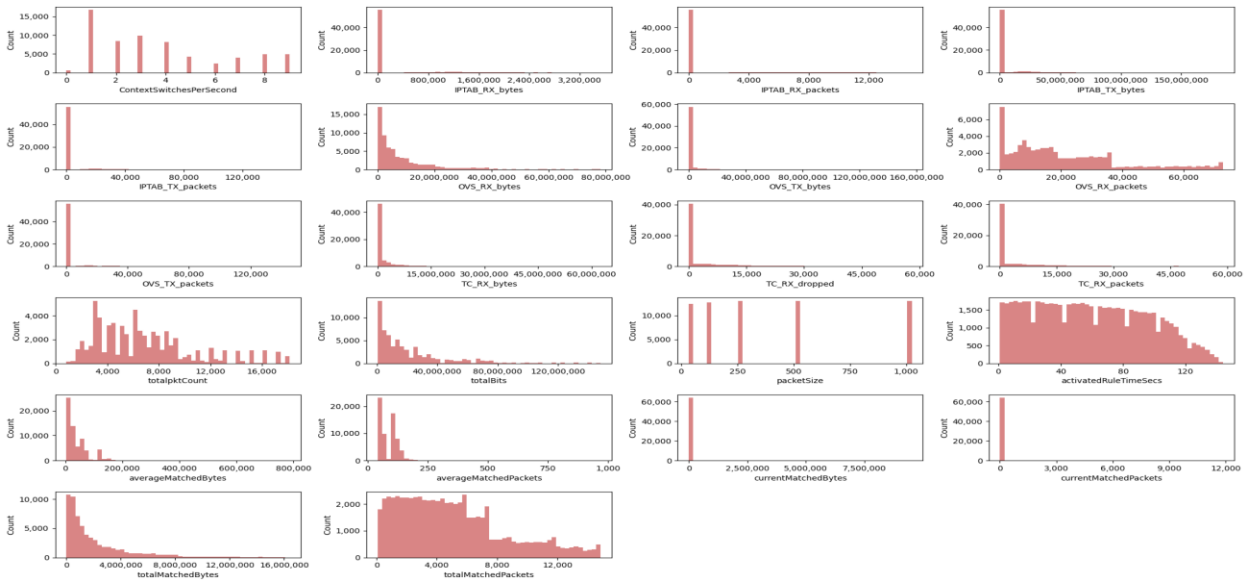


Table 7. Descriptive statistics values of numerical features in IW-IB-5GNET dataset.

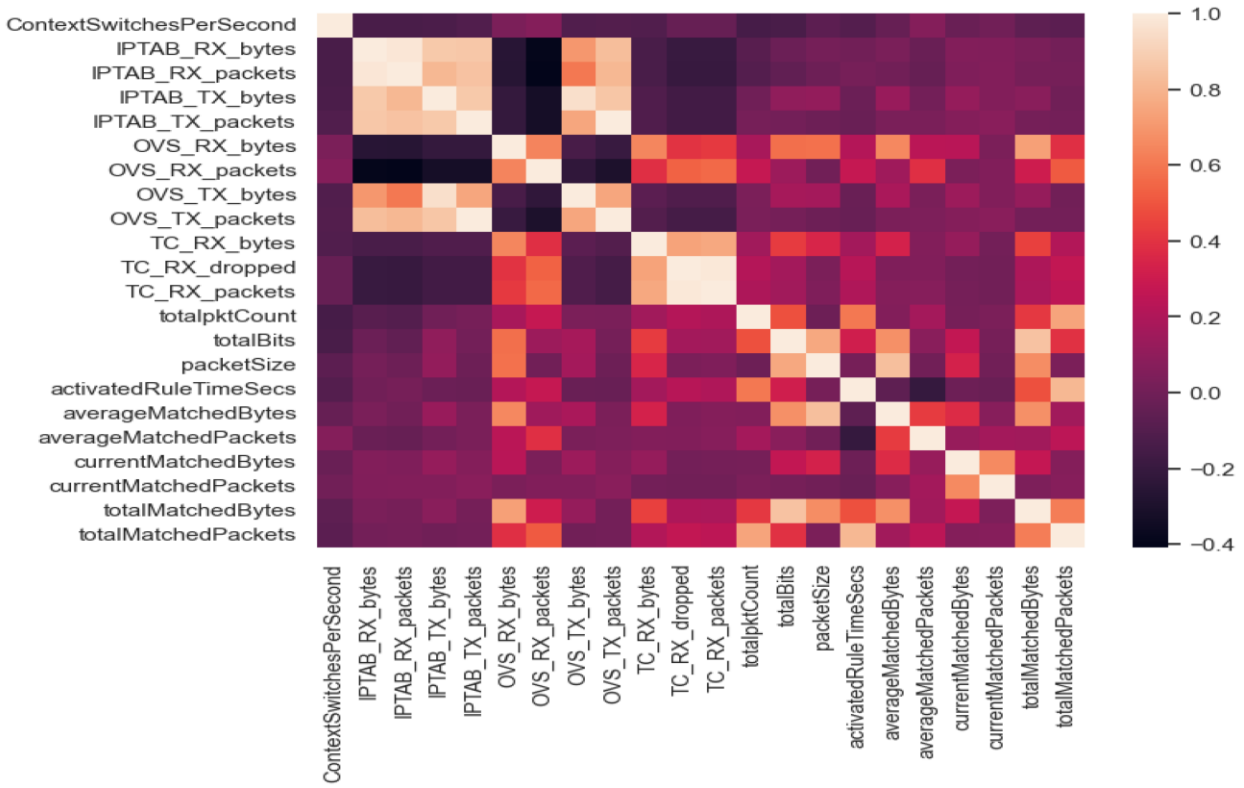
Feature	Unit	Mean	Std	Min	Max
ContextSwitchesPerSecond	Switch/s	3.78	2.65	0	9
IPTAB_RX_bytes	Bytes/s	178,080	495,268	0	3,524,780
IPTAB_RX_packets	Packets/s	987	2661	0	14,866
IPTAB_TX_bytes	Bytes/s	4,031,111	13,392,294	0	184,598,347
IPTAB_TX_packets	Packets/s	5345	17,555	0	151,650
OVS_RX_bytes	Bytes/s	9,290,003	12,603,834	9926	78,227,821
OVS_RX_packets	Packets/s	20,772	18,092	93	72,736
OVS_TX_bytes	Bytes/s	2,135,533	9,334,686	6354	156,582,130
OVS_TX_packets	Packets/s	5014	16,145	611	145,377
TC_RX_bytes	Bytes/s	2,642,547	7,151,074	0	62,570,352
TC_RX_dropped	Packets/s	5145	9556	0	57,989
TC_RX_packets	Packets/s	5682	10,507	0	58,842
totalpktCount	Packets	6866	3809	840	18,120
totalBits	Bits	21,578,431	25,187,493	416,976	148,316,160
packetSize	Bytes	395	355	32	1024
activatedRuleTimeSecs	Seconds	59	36	1	144
averageMatchedBytes	Bytes/s	40,171	37,985	3665	785,862
averageMatchedPackets	Packets/s	89	39	42	969
currentMatchedBytes	Bytes/s	37,469	91,559	0	9,513,560
currentMatchedPackets	Packets/s	83	159	0	11,859
totalMatchedBytes	Bytes	2,299,883	2,792,487	5772	16,381,370
totalMatchedPackets	Packets	5031	3518	70	14,855

پس از اعتبارسنجی فنی ویژگی‌های عددی، شکل 7 را ارائه می‌دهیم. این شکل، نمودارهای توزیع داده ستون‌های عددی در مجموعه داده IW-IB-5GNET را نشان می‌دهد. طیف گسترده‌ای از اشکال در توزیع داده‌ها را می‌توان در شکل مشاهده کرد. ابتدا، می‌توانیم مقادیر گسسته مانند ContextSwitchesPerSecond یا packetSize را شناسایی کنیم. با تمرکز بر نمودار packetSize در شکل 7، پنج اندازه بسته مختلف انتخاب شده برای آزمایش‌ها را می‌توان مشاهده کرد. این اندازه‌ها 32، 128، 256، 512 و 1024 بایت هستند. بسیاری از هیستوگرام‌ها با اشکال نمایی نیز در شکل وجود دارند، مانند ویژگی‌های totalBits و totalMatchedBytes. ویژگی totalBits تعداد بیت‌هایی را که در واقع از یک رابط شبکه خاص عبور می‌کنند نشان می‌دهد، در حالی که totalMatchedBytes تعداد کل بایت‌هایی را که با قانون شبکه‌ای که تحت نظارت است مطابقت دارند، تعیین می‌کند. در نهایت، انواع اشکال غیر یکنواخت را مشاهده می‌کنیم. اکثر این شکل‌ها دارای حداکثر مقدار بالا و مقادیر درهم‌ریخته بسیار کوچکتری در اطراف آن هستند. نمونه‌هایی از این موارد عبارتند از totalpktCount یا totalMatchedPackets. با نگاهی به نمودار کل بسته‌های منطبق در شکل ۷، می‌توانیم مشاهده کنیم که اکثر مقادیر در حدود میانگین، یعنی ۵۰۳۱، هستند. علاوه بر این، حداکثر مقدار آن ۱۴۸۵۵ است که نشان می‌دهد در برخی از آزمایش‌ها، این قانون با تعداد زیادی بسته مطابقت دارد.



برای نتیجه‌گیری از ارزیابی فنی مجموعه داده IW-IB-5GNET، ضرایب همبستگی ویژگی‌های عددی محاسبه شده‌اند. تکنیک مورد استفاده، محاسبه ضرایب همبستگی پیرسون (PCC) بود که همبستگی خطی بین دو مجموعه داده را اندازه‌گیری می‌کند. مقادیر ضرایب بین 1- و 1 متغیر است که نشان دهنده قدرت و جهت رابطه خطی بین دو متغیر است [35]. شکل 8 ماتریس همبستگی به‌دست‌آمده از این ضرایب را نشان می‌دهد. مشاهده می‌شود که ویژگی‌هایی با همبستگی بالا با رنگ نارنجی بسیار روشن نمایش داده شده‌اند. در مقابل، ویژگی‌هایی با همبستگی منفی بالا با بنفش تیره نمایش داده شده‌اند. ویژگی‌هایی با همبستگی کم بین آنها با رنگ‌های متوسط نمایش داده شده‌اند. به عنوان مثال، همبستگی مثبت

واضحی بین بایت‌ها و بسته‌های هر جفت ویژگی مرتبط با همان نمونه تحت نظارت وجود دارد (یعنی هرچه بسته‌های بیشتری دریافت شود، بایت‌های بیشتری دریافت می‌شود). نمونه‌ای از این را می‌توان در ویژگی‌های دوم و سوم در شکل 8، به ترتیب IPTAB_RX_bytes و IPTAB_RX_bytes، یافت، که در آن رنگ نارنجی بسیار روشن نشان‌دهنده همبستگی مثبت آنها قابل مشاهده است. با تمرکز بر همبستگی‌های بسیار منفی، ویژگی averageMatchedPackets را با ویژگی activatedRuleTimeSecs بررسی می‌کنیم. هرچه یک قانون زمان بیشتری در شبکه فعال باشد، تعداد کمتری از بسته‌های متوسط با آن قانون مطابقت دارند. به طور کلی، ما یک رابطه خطی واضح در قسمت پایین سمت راست ماتریس همبستگی مشاهده می‌کنیم. این مربوط به معیارهای مرتبط با قوانین شبکه است که در طول اجرای آزمایش‌ها بسیار متغیر هستند. از سوی دیگر، وابستگی خطی کمی بین معیارهای مرتبط با iptables با توجه به بقیه ویژگی‌ها وجود دارد. این به این دلیل است که iptables در رابط خروجی شبکه لبه قرار دارد. بنابراین، ترافیک در بسیاری از آزمایش‌ها عبور نمی‌کند زیرا ترافیک قبل از رسیدن به رابط خروجی قطع می‌شود.



به طور خلاصه، ماتریس همبستگی اطلاعات زیادی در مورد رابطه داده‌های ما به ما می‌دهد. بسته به نوع مشکلی که انتظار داریم با استفاده از مجموعه داده‌های خود به آن بپردازیم، باید به برخی الگوها توجه کنیم. به عنوان مثال، چندخطی بودن (دو یا چند متغیر با یکدیگر همبستگی بالایی دارند) می‌تواند در تحلیل رگرسیون مشکل‌ساز باشد، زیرا می‌تواند منجر به

تخمین‌های ضریب ناپایدار شود. بنابراین، بسته به هدف نهایی مجموعه داده‌ها، نتایج به عنوان کافی یا ناکافی در نظر گرفته می‌شوند و ممکن است اقدامات متفاوتی لازم باشد.

7. بحث

این بخش با هدف تثبیت و مستندسازی بینش‌ها و بررسی‌های حیاتی حاصل از تحقیقات ما انجام می‌شود. این مجموعه با هدف ارائه منبعی جامع برای محققان آینده تهیه شده است و آنها را قادر می‌سازد تا از ملاحظات ما به عنوان پایه‌ای ارزشمند برای تحقیقات علمی خود بهره‌مند شوند.

در طول طراحی و پیاده‌سازی چارچوب پیشنهادی، به اهمیت بالای اجرای مکانیسم‌هایی برای همسوسازی IDS از تمام اجزای توپولوژی شبکه پی بردیم. اینها جریان‌های شبکه، پورت‌های شبکه، فناوری‌های صفحه داده موجود در هر پورت شبکه و میزبان‌های دستگاه هستند. این امر برای امکان دستکاری بعدی ویژگی‌های حاصل از هر یک از این موجودیت‌های مختلف بسیار مهم است. ما همچنین دشواری واقعی دستیابی به این مجموعه داده مبتنی بر هدف و در سطح زیرساخت را درک و دریافته‌ایم، زیرا نه تنها به یک زیرساخت کامل با سطح کافی از ادغام بین اجزا، بلکه به یک حلقه کنترل بسته کامل و کاملاً کاربردی که بر روی آن اجرا می‌شود نیز نیاز دارد. همچنین با در نظر گرفتن این نکته که علاوه بر موارد فوق، یک سیستم استخراج ویژگی خودکار نیز در حال اجرا است.

در زمینه تحقیقات ما، استفاده از شبیه‌سازهای شبکه نقش محوری در تولید سناریوهای متنوع برای مجموعه‌های داده ایفا می‌کند. این شبیه‌سازها ما را قادر می‌سازند تا محیط‌های کنترل‌شده را با دقت و صحت بازآفرینی کنیم و بررسی دقیق شرایط مختلف شبکه و تأثیر آنها بر مطالعه ما را تسهیل کنیم. در نهایت، در نمونه اولیه خود، اهمیت ایجاد ارتباط بین اجزای شبکه در بین مهرهای زمانی را تشخیص دادیم، زیرا آنها نقش مهمی در تسهیل آموزش هوش مصنوعی در مراحل بعدی ایفا می‌کنند. در نتیجه، یک اصل طراحی اساسی که پدیدار شد، گنجاندن مهرهای زمانی در هر رابط موجود در سیستم پیشنهادی بود. این انتخاب طراحی، فرآیند ردیابی را امکان‌پذیر می‌کند. این رویکرد به درک زمینه و زمان تولید رویداد در سیستم کمک می‌کند.

8. نتیجه‌گیری

در این تحقیق، نیاز به یک مجموعه داده جدید که بتواند پیچیدگی‌های شبکه‌های B5G، از جمله توپولوژی‌های آنها در سطوح مختلف و ماهیت پویای قوانین کنترل شبکه را ثبت کند، تشخیص داده شده است. این مقاله یک مجموعه داده شبکه جدید و جامع، IW-IB-5GNET، را ارائه داده است که در سطح زیرساخت و مبتنی بر هدف است و به نیاز مبرم به راه‌حل‌های داده‌محور قوی‌تر و سازگارتر در مدیریت و بهینه‌سازی شبکه در شبکه‌های B5G می‌پردازد. این راه‌حل‌ها می‌توانند توسط ISPها و DSPها برای بهبود مدیریت و بهینه‌سازی سیاست‌های شبکه خود در هر دو بخش لبه و هسته استفاده شوند. مجموعه داده ما چندین مزیت کلیدی از نظر مدیریت و بهینه‌سازی شبکه ارائه می‌دهد. اولاً، دسترسی گسترده به

زیرساخت آن تضمین می‌کند که کل اکوسیستم شبکه را در بر می‌گیرد و دیدگاه گسترده‌ای از پویایی و وضعیت شبکه ارائه می‌دهد. این شمول با پیچیده‌تر و به‌هم‌پیوسته‌تر شدن شبکه‌ها حیاتی است. دوم اینکه، این مجموعه داده مبتنی بر قصد و نیت است و نه تنها جنبه‌های فنی شبکه را مستند می‌کند، بلکه اهداف و اقدامات کنترلی اساسی که پیکربندی‌ها و سیاست‌های شبکه را هدایت می‌کنند را نیز در نظر می‌گیرد. مهم است که ماهیت این مجموعه داده که از یک حلقه بسته در یک شبکه B5G استخراج شده است، برجسته شود.

نتایج تجربی و تحلیلی، تنوع گسترده‌ای را در توزیع داده‌ها و همچنین رایج‌ترین مقادیر و همبستگی‌های خطی آنها نشان می‌دهد. این نتایج، وضعیت شبکه را در لایه‌های مختلف آن نشان می‌دهد که از آنها معیارهای عملکرد ارزشمندی استخراج می‌شود. به طور خاص، جدول 7 و شکل 6 و شکل 7 به آن اجازه می‌دهند تا یک تحلیل اکتشافی از داده‌ها انجام دهد، که گامی ضروری قبل از اجرای هر مدلی است. علاوه بر این، شکل 8 به ما کمک می‌کند تا اقداماتی مانند کاهش ابعاد را برای بهینه‌سازی مدل‌های داده‌محور انجام دهیم. نتایج می‌تواند برای تولید مدل‌های مبتنی بر هوش مصنوعی برای بهینه‌سازی سیاست‌های شبکه و همچنین مدل‌های هوش مصنوعی برای بهبود QoS بسیار مفید باشد، به عنوان مثال، ایجاد یک مدل طبقه‌بندی برای بهینه‌سازی قوانینی که در حال حاضر در شبکه اجرا می‌شوند، به طوری که مدل بتواند پیش‌بینی کند کدام فناوری برای اجرای یک سیاست شبکه بهینه است. مثال دیگر، مدلی است که قادر به تشخیص قوانین شبکه‌ای است که استفاده نمی‌شوند و بنابراین می‌توان آنها را برای بهبود تراکم سیاست شبکه حذف کرد. با این وجود، مجموعه داده‌های ما نیز محدودیت‌هایی دارد. ما می‌دانیم که این مجموعه داده مقیاس محدودی دارد و به اندازه کافی گسترده نیست که بتواند کل پیچیدگی سناریوهای دنیای واقعی را پوشش دهد. علاوه بر این، از نظر پوشش هدف، فاقد تنوع است و فعلاً بر روی نوع خاصی از هدف تمرکز دارد.

در کارهای آینده، پتانسیل مجموعه داده IW-IB-5GNET را بیشتر بررسی خواهیم کرد. ما کاربرد آن را در حوزه‌های مختلف، با تمرکز ویژه بر چالش‌های مدیریت شبکه، بهینه‌سازی و QoS، بررسی خواهیم کرد. ما نه تنها اثربخشی آن را ارزیابی خواهیم کرد، بلکه مدل‌های شناخته شده هوش مصنوعی را نیز به کار خواهیم گرفت. علاوه بر این، هدف ما کار بر روی محدودیت‌های ذکر شده در بالا و گسترش مقیاس و پوشش مجموعه داده است. نه تنها با افزایش نوع حملات DDoS، بلکه با اعمال انواع دیگر عبارات هدف.

مشارکت‌های نویسندگان

مفهوم‌سازی، J.A.-H. و J.M.A.-C.؛ گردآوری داده‌ها، J.A.-H.؛ تحلیل رسمی، J.A.-H.؛ تأمین بودجه، Q.W. و J.M.A.-C.؛ تحقیق، J.A.-H.؛ روش‌شناسی، J.A.-H.؛ مدیریت پروژه، Q.W.؛ منابع، J.A.-H. و J.M.A.-C.؛ نرم‌افزار، J.A.-H.؛ نظارت، J.M.A.-C.؛ اعتبارسنجی، J.A.-H.؛ مصورسازی، J.A.-H.؛ نگارش - پیش‌نویس اصلی، J.A.-H.؛ نگارش - بررسی و ویرایش، Q.W. و J.M.A.-C.. همه نویسندگان نسخه منتشر شده مقاله را خوانده و با آن موافقت کرده‌اند. بودجه

این تحقیق توسط کمیسیون اروپا تحت دو پروژه تأمین مالی شده است: RIGOUROUS (طراحی و استقرار امن محاسبات ابری G6) و ARCADIAN-IoT (چارچوب مدیریت اعتماد، امنیت و حریم خصوصی مستقل برای اینترنت اشیا) با شماره‌های کمک هزینه (HORIZON-JU-SNS-2022-STREAM-B-01-04/101095933) و (H2020-SU-DS-2020/101020259).