

Deteksi DDoS Berbasis Peningkatan di Sistem Internet untuk Segala Hal

Ivan Cvitić^{ID}, Dragan Peraković^{ID}, Anggota Senior, IEEE, Brij B. Gupta, Anggota Senior, IEEE Bahasa Indonesia: dan Kim-Kwang Raymond Choo^{ID}, Anggota Senior, IEEE

Abstrak—Serangan Distributed Denial-of-Service (DDoS) tetap menjadi tantangan untuk diatasi dalam sistem yang ada, termasuk jaringan internal yang terdiri dari berbagai perangkat Internet of Things (IoT). Dalam artikel ini, kami menyajikan model deteksi lalu lintas DDoS yang menggunakan metode boosting dari pohon model logistik untuk berbagai kelas perangkat IoT. Secara khusus, versi model yang berbeda akan dibuat dan diterapkan untuk setiap kelas perangkat karena karakteristik lalu lintas jaringan dari setiap kelas perangkat mungkin memiliki variasi yang halus. Sebagai studi kasus, kami menjelaskan bagaimana perangkat dalam lingkungan rumah pintar yang umum dapat dikategorikan ke dalam empat kelas yang berbeda (dan dalam konteks kami, Kelas 1—tingkat predikabilitas lalu lintas yang sangat tinggi, Kelas 2—tingkat predikabilitas lalu lintas yang tinggi, Kelas 3—tingkat predikabilitas lalu lintas sedang, dan Kelas 4—tingkat predikabilitas lalu lintas yang rendah). Temuan dari evaluasi kami menunjukkan bahwa akurasi pendekatan yang kami usulkan adalah antara 99,92% dan 99,99% untuk keempat kelas perangkat ini. Dengan kata lain, kami menunjukkan bahwa kami dapat menggunakan kelas perangkat untuk membantu kami mendeteksi lalu lintas DDoS secara lebih efektif.

Istilah Indeks—Kecerdasan buatan, keamanan siber, Distributed Denial of Service (DDoS), pembelajaran mesin ensemble, IDS, Internet of Things (IoT), pembelajaran yang diawasi.

Aku. AKUPENGANTAR

Pengguna dan sistem INTERNET OF THINGS (IoT) menjadi hal yang lumrah dan kakehnya, semakin banyak menjadi sasaran penyerang, misalnya, dengan mengidentifikasi dan mengeksplorasi kerentanan dalam perangkat lunak dan perangkat keras IoT, atau implementasinya, untuk memfasilitasi aktivitas yang tidak sah dan berbahaya. Perangkat tersebut juga telah dieksplorasi untuk membuat jaringan botnet guna menghasilkan lalu lintas Distributed Denial-of-Service (DDoS). DDoS merupakan ancaman siber berorientasi jaringan yang kritis, yang trennya terus meningkat selama bertahun-tahun.

Naskah diterima pada tanggal 25 Januari 2021; direvisi pada tanggal 31 Maret 2021 dan Mei 14, 2021; diterima 15 Juni 2021. Tanggal publikasi 21 Juni 2021; tanggal versi terkini 24 Januari 2022. Karya ini didukung oleh Universitas Zagreb, Kroasia, melalui Proyek "Tantangan Jaringan Informasi dan Komunikasi, Teknologi, Layanan, dan Peralatan Pengguna dalam Membangun Lingkungan Masyarakat 5.0—Fase 2" di bawah Hibah 210219; ZUID2020/2021. Karya Kim-Kwang Raymond Choo didukung oleh Cloud Technology Endowed Professorship.(Penulis koresponden: Ivan Cvitić, Brij B. Gupta.)

Ivan Cvitić dan Dragan Peraković berasal dari Fakultas Ilmu Transportasi dan Lalu Lintas, Departemen Lalu Lintas Informasi dan Komunikasi, Universitas Zagreb, 10000 Zagreb, Kroasia (e-mail: ivan.cvitic@fpz.unizg.hr ; dragan.perakovic@fpz.unizg.hr).

Brij B. Gupta bekerja di Institut Teknologi Nasional Kurukshetra, Kurukshetra 136119, India, dan juga di Universitas Asia, Taichung 413, Taiwan (e-mail: bbgupta@nitkkr.ac.in).

Kim-Kwang Raymond Choo bekerja di Departemen Sistem Informasi dan Keamanan Siber, Universitas Texas di San Antonio, San Antonio, TX 78249 AS (e-mail: raymond.choo@fulbrightmail.org).

Pengidentifikasi Objek Digital 10.1109/JIOT.2021.3090909

dekade terakhir [1], [2]. Misalnya, serangan DDoS yang menargetkan Amazon AWS pada Q1 tahun 2020 dilaporkan memiliki volume puncak 2,3 Tbps [3].

Perangkat dan sistem IoT tidak hanya ditemukan di lingkungan organisasi atau pemerintahan, tetapi juga di rumah kita (misalnya, rumah pintar). Rumah pintar adalah salah satu aplikasi IoT yang paling cepat berkembang, dan perangkat yang digunakan sangat heterogen. Perangkat tersebut sering kali dikirimkan dengan mekanisme keamanan yang minimal atau tidak ada sama sekali, dan dalam upaya untuk membuat perangkat ini ramah pengguna, persyaratan keamanan sering kali dikurangi [4]. Selain itu, sebagian besar perangkat di rumah pintar tidak mahal dan tidak memiliki kemampuan komputasi yang signifikan dan, akibatnya, perangkat tersebut dapat dengan mudah dikompromikan untuk memfasilitasi berbagai aktivitas jahat, termasuk menghasilkan lalu lintas DDoS [5]. Dalam ekosistem rumah pintar yang umum, terdapat beberapa kelompok pemangku kepentingan, seperti pengguna akhir (pemilik rumah atau penyewa di dalam rumah), penyedia layanan internet/telekomunikasi, produsen perangkat, dan penyedia layanan (misalnya, penyedia layanan pihak ketiga seperti layanan keamanan yang dipantau). Para pemangku kepentingan ini umumnya memiliki kepentingan pribadi untuk tidak terlibat dalam aktivitas siber yang jahat, atau agar perangkat, sistem, platform, dan/atau infrastruktur mereka dieksplorasi untuk memfasilitasi aktivitas jahat. Misalnya, kepentingan penyedia layanan internet/telekomunikasi adalah untuk segera mendeteksi setiap perilaku/aktivitas tak berizin dalam lingkungan rumah pintar, melindungi infrastruktur jaringan mereka sendiri, dan mencegah perangkat/sistem yang disusupi digunakan sebagai landasan peluncuran terhadap perangkat dan sistem lain (dengan implikasi hukum dan finansial terkait).

Tantangannya adalah bagaimana merancang sistem deteksi DDoS yang efektif yang dapat digunakan dalam lingkungan rumah pintar yang semakin beragam dan dinamis. Misalnya, berdasarkan karakteristik lalu lintas jaringan yang dihasilkan, seseorang dapat mengidentifikasi jenis perangkat yang umum ditemukan di lingkungan rumah pintar [5]. Sejalan dengan itu, model untuk mengklasifikasikan perangkat IoT ke dalam kelas yang telah ditentukan sebelumnya disajikan dalam penelitian kami sebelumnya [6], di mana kami mendefinisikan kelas perangkat IoT berdasarkan perilaku lalu lintas dan predikabilitas perilakunya (yaitu, koefisien variasi rasio data yang diterima dan dikirim). Berdasarkan penelitian sebelumnya ini, kami menyajikan dua hipotesis berikut. Pertama, adalah mungkin untuk mendefinisikan profil lalu lintas yang sah (normal) untuk kelas perangkat IoT, berdasarkan karakteristik arus lalu lintas. Hipotesis kedua adalah, berdasarkan profil lalu lintas sah masing-masing kelas perangkat IoT, kami dapat mengembangkan model pembelajaran mesin yang diawasi yang dapat secara efektif mendeteksi DDoS

lalu lintas sebagai anomali jaringan yang dihasilkan dari perangkat IoT individual. Oleh karena itu, kami mengembangkan model deteksi DDoS untuk sistem IoT yang dinamis dan heterogen, yang dapat diimplementasikan dalam lingkungan rumah pintar. Kami juga mencatat bahwa model deteksi DDoS yang disajikan dalam artikel ini menggunakan metode penguatan pohon model logistik (LMT), di mana versi model yang berbeda diterapkan untuk setiap kelas perangkat.

Kontribusi penelitian kami dapat diringkas sebagai berikut.

- 1) Kumpulan data IoT yang sah dan lalu lintas DDoS anomali yang dihasilkan dalam penelitian ini akan tersedia untuk umum bagi komunitas ilmiah yang luas (dan tidak ada kumpulan data seperti itu di [7]—lihat juga bagian kedua).
- 2) Proses yang kami tetapkan untuk membentuk profil lalu lintas normal untuk kelas perangkat IoT.
- 3) Model deteksi DDoS yang kami usulkan, yang menggunakan kelas perangkat untuk mendeteksi lalu lintas DDoS. Kami berpendapat bahwa pendekatan semacam itu lebih efektif, seperti yang akan kami tunjukkan nanti dalam artikel ini.

Sisa artikel ini disusun sebagai berikut. Bagian II mengulas secara singkat literatur DDoS terkait. Bagian III menjelaskan metodologi pengumpulan data, praproses set data, dan pengembangan model deteksi DDoS berdasarkan metode pohon model logistik dari kumpulan pembelajaran mesin terbimbing. Bagian IV menunjukkan analisis temuan, yang menunjukkan bahwa akurasi model tinggi untuk semua kelas perangkat (yaitu, tingkat akurasi antara 99,92% dan 99,99%). Kami juga akan membahas implikasi dari pekerjaan kami. Di Bagian V, kami akan menyimpulkan artikel ini dan membahas kemungkinan penelitian di masa mendatang.

II. RGEMBIRARPENELITIAN

Telah banyak aplikasi teknik pembelajaran mesin untuk mendeteksi lalu lintas DDoS, yang dapat dikategorikan menjadi teknik yang berdasarkan pada pengawasan (menggunakan pengetahuan yang ada untuk mengklasifikasikan kejadian yang tidak diketahui di masa mendatang) dan teknik yang berdasarkan pada tanpa pengawasan (mencoba menentukan kelas kejadian yang sesuai tanpa pengetahuan sebelumnya). Misalnya, Doshi dan *lain-lain*.[8] mengembangkan model klasifikasi biner lalu lintas pada lalu lintas yang sah dan lalu lintas DDoS menggunakan lima metode pembelajaran mesin yang berbeda. Fitur khusus lalu lintas Smart Home IoT (SHIoT) diamati melalui perubahan karakteristik lalu lintas, seperti ukuran paket, waktu interim paket, protokol yang digunakan, dan perubahan jumlah alamat protokol Internet (IP) tujuan yang dikomunikasikan perangkat ini pada interval waktu yang berbeda. Penelitian yang disajikan dalam [9] juga mengusulkan pendekatan lalu lintas DDoS yang dihasilkan oleh perangkat IoT di lingkungan perusahaan, menggunakan metode Deep Autoencoders berdasarkan jaringan saraf tiruan. dan *lain-lain*.[10] menyarankan bahwa efisiensi deteksi lalu lintas DDoS lebih tinggi jika dilakukan di tepi lingkungan IoT yang diamati. Cvitić dan *lain-lain*.[11] mengusulkan model deteksi DDoS konseptual yang mempertimbangkan kelas perangkat IoT.

Meskipun akurasi deteksinya tinggi dan banyaknya keuntungan dari pendekatan-pendekatan yang ada, namun terdapat beberapa kekurangan dan

tantangan tetap ada. Tantangan utama adalah kurangnya kumpulan data relevan yang dapat digunakan untuk melatih model deteksi berbasis pembelajaran mesin [12], [13]. Meskipun ada sejumlah kumpulan data yang berisi DDoS dan lalu lintas normal, ini seringkali usang dan akibatnya mengurangi akurasi deteksi, karena tidak mencerminkan karakteristik lalu lintas saat ini karena perangkat yang lebih baru, konsep jaringan (misalnya, jaringan yang ditentukan perangkat lunak), dan layanan telah digunakan [14], [15]. Misalnya, Doshi dan *lain-lain*.[8] menggunakan tiga perangkat dengan lalu lintas yang dikumpulkan selama periode 10-m, sedangkan penelitian di [9] menggunakan sembilan perangkat, yang lima di antaranya adalah webcam atau kamera keamanan. Namun, kumpulan data dari [9] tidak tersedia untuk umum dalam bentuk aslinya. Ini hanya tersedia sebagai *.file csv* yang berisi fitur lalu lintas yang telah diekstraksi. Hal ini membatasi bagi peneliti lain karena tidak memiliki lalu lintas yang dihasilkan dalam bentuk aslinya yang disimpan dalam format yang memungkinkan peneliti untuk mengekstrak dan menghitung fitur yang berbeda dari yang diekstraksi oleh Meidan dan *lain-lain*.[9] Saharkizan dan *lain-lain*.[16] menggunakan kumpulan data yang diperoleh melalui simulasi dalam pendekatan yang diusulkan berdasarkan metode long short-term memory (LSTM) untuk mendeteksi serangan di jaringan IoT. Dalam [17], model dua tingkat digunakan untuk menganalisis arus lalu lintas jaringan. Fitur arus lalu lintas dipilih secara empiris, dan kumpulan data publik yang ada digunakan untuk mengevaluasi model deteksi. Salman dan *lain-lain*.[18] menyajikan sebuah model untuk mengidentifikasi perangkat IoT dan mendeteksi serangan pada perangkat IoT menggunakan beberapa metode pembelajaran mesin (yaitu, pohon keputusan, hutan acak, dan metode pembelajaran mendalam). Penelitian ini menggunakan kumpulan data lalu lintas yang dikumpulkan menggunakan tujuh perangkat IoT. Akurasi deteksi maksimum dari model yang dikembangkan adalah 94,47%. Pendekatan deteksi DDoS lainnya mencakup pendekatan yang disajikan dalam [19]–[21]. Pengamatan lain dari karya-karya ini adalah bahwa kumpulan data umumnya sangat kecil dan tidak mewakili sistem dunia nyata.

Membuat testbed yang kuat untuk menghasilkan set data yang realistik merupakan hal yang menantang, memakan waktu, dan mahal sebagian karena kemungkinan kombinasi konfigurasi yang berbeda. Set data yang ada juga berbeda dalam cara pembuatannya, yang dapat berupa sintetis, simulasi, atau nyata [22]. Set data sintetis dibuat untuk memenuhi persyaratan dan kondisi spesifik yang juga dipenuhi oleh set data nyata. Set data yang ada yang digunakan dalam literatur juga umumnya bertanggal (misalnya, dibuat antara tahun 1998 dan 2012) dan, oleh karena itu, mungkin tidak mewakili jaringan komunikasi saat ini. Bahkan set data yang lebih baru jarang menyertakan lalu lintas IoT—lihat juga Tabel I. Contoh set data yang ada termasuk yang dari University of New South Wales di Australia [23], yang terdiri dari sejumlah perangkat SHIoT. Untuk pengembangan sistem deteksi anomali, sangat penting untuk memiliki set data yang berisi lalu lintas normal/sah yang dihasilkan oleh perangkat IoT. Dari set data tersebut, dimungkinkan untuk menentukan profil perilaku lalu lintas normal untuk perangkat individual atau seluruh kelas perangkat IoT.

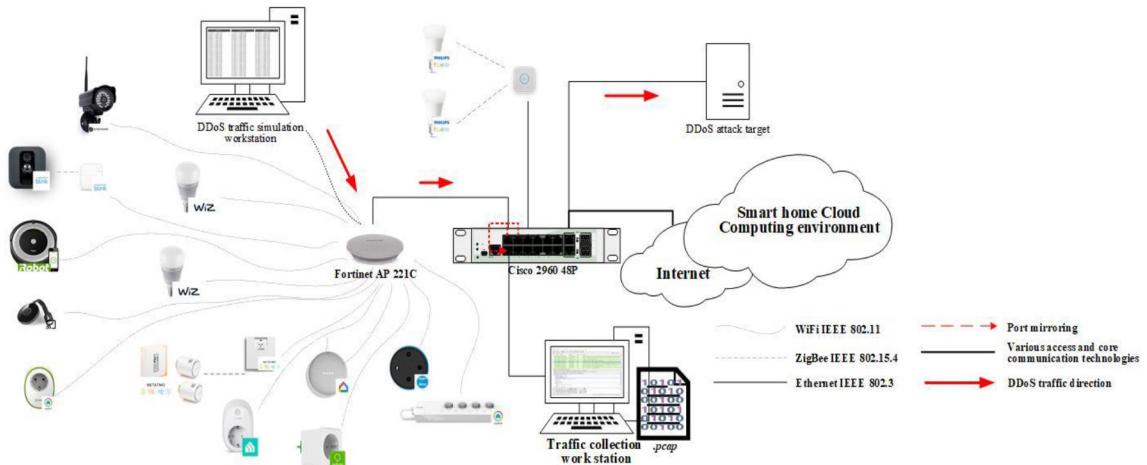
III. Hal.DITUTUPKANAPENDEKATAN

A. Pengaturan Tempat Uji

Pengaturan lingkungan laboratorium rumah pintar kami ditunjukkan pada Gambar 1, dan juga disajikan dalam [24]. Lingkungan kami

TABEL I
Foto Napshot dari Bahasa Inggris MENGENAI PADA SAYAHAT DAN SAYAHAT DATA SETS

Dataset	Devices (number, types, etc.)	Setup (synthetic, simulated/emulated, or real)	Collected traffic	Year created	Types of traffic acquired
DARPA'98 [41]	Non-IoT, conventional devices (PC, servers) No. of devices: N/A	Simulated (small network) – simulated Air Force Base network	Collecting time: 2 weeks	1998	Attack traffic (38 types of attacks) DoS (11 types), R2L (unauthorized access from remote machine, 14 types) U2R (unauthorized access to local root, 7 types), probe (6 types)
KDDcup99 [42]	Non-IoT, convencional devices(PC, servers) No. of devices: N/A	Simulated traffic in a military environment (small network)	Attack instances - 3,925,650 Benign instances - 972,781	1999	DoS (SYN flood), R2L, U2R, probe
CAIDA [43]	Non-IoT, conventional devices (PC, servers) No. of devices: N/A	Real	Collecting time: 1 hour	2007	DDoS traffic
NSL-KDD [44]	Non-IoT, conventional devices (PC, servers) No. of devices: N/A	Emulated (small network)	Number of instances (train set): 4,898,431 Number of instances (test set): 311,027	2009	Attack traffic, normal/legitimate traffic
TUIDS [45]	Non-IoT, conventional devices (PC, servers) No. of devices: N/A	Emulated	Collecting time: 4 week (5.3 GB) No. of packets – 432875 No. of flows - 400131	2011/2012	Attack traffic (16 attack types) Normal traffic
CICIDS2017 [46]	Non-IoT, conventional devices (PC, servers) 25 users behaviour profiles	Real	Collecting time: 24 hours (4.6 GB)	2017	Attack traffic (High-volume and low-volume application-level DDoS)
CSE-CIC-IDS2018 [47]	Non-IoT, conventional devices (PC, servers) No. of devices: 50 machines	Real	N/A	2018	Attack traffic (seven scenarios: Brute-force, Heartbleed, Botnet, DoS, DDoS, Web attacks, and infiltration of the network from inside)
N-BaIoT	IoT devices infected with Mirai and Bashlite No. of devices: 9	Real	Number of instances: 7062606	2018	Attack traffic (spam, UDP flood, TCP flood, Scan, ACK flood) Normal/legitimate traffic
Bot-IoT [48]	IoT devices No. of devices: 5 simulated IoT devices	Simulated (Ostinato and Node-red tool)	Collecting time: 4 week No. of packets – 432875 No. of flows - 400131	2019	Normal IoT, Attack IoT (Information gathering, DoS, Keylogger)
CICDDoS2019 [49]	Non-IoT, conventional devices (PC, servers)	Simulated	No. of instances for attack – 73360900 No. of instances for benign - 9543	2019	Attack traffic (PortMap, NetBIOS, LDAP, MSSQL, UDP, UDP-Lag, SYN, NTP, DNS, SNMP, SSDP, WebDDoS, TFTP)
University of New South Wales [31]	IoT devices No. of devices: 28 real IoT devices	Real	Collecting time: 26 weeks Daily average traffic – 365 MB	2019	Normal/legitimate traffic



Gbr. 1. Uji coba rumah pintar.

terdiri dari 41 perangkat SHIoT yang beragam, dan infrastruktur komunikasi dasar serta platform perangkat lunak-perangkat keras juga disiapkan untuk memungkinkan pengumpulan lalu lintas yang dapat digunakan untuk melatih model deteksi DDoS.

Selain data primer yang dikumpulkan dalam penelitian ini, kami juga menggunakan data sekunder dari [23], termasuk sejumlah besar perangkat SHIoT yang berbeda (yaitu, heterogenitas perangkat yang lebih besar). Titik akses nirkabel Fortinet AP 221C, sakelar Cisco 2960 Catalyst 48 Power over Ethernet (PoE), HP Pavilion dm1, dan Microsoft HP 10 10.0.17134 build

17.134 stasiun kerja telah disiapkan untuk menangkap lalu lintas menggunakan pencerminan port, arsitektur prosesor x64, AMD E-350, dua inti 1600-MHz, RAM 4-GB) dengan perangkat lunak Wireshark versi 2.6.3 terpasang. Port komunikasi fisik sakelar (FA0/1 dan FA0/3) tempat titik akses nirkabel dan hub IoT untuk perangkat Phillips Hue terhubung dikonfigurasi untuk pencerminan port. Port-port ini disiapkan sebagai sumber, yang memastikan bahwa semua lalu lintas ke dan dari mereka dicerminkan (dipetakan) ke port kontak tujuan (FA0/2). Stasiun kerja pengumpulan lalu lintas terhubung ke port ini. Dengan

TABEL II
HAIASLISAYASAH DAN DDoS TRAFIK DATASETS'
Karakteristik

	Number of files	Number of collected packets	The amount of data collected (GB)	Collection period (hours)
Primary (sum)	103	456,174,601	344.59	2,472.01
Secondary (sum)	41	99,334,088	38.16	986.45
DDoS-UDP	245	269,806,374	19.95	10.75
DDoS-TCP	73	85,373,401	5.88	17.12
DDoS-ICMP	195	217,593,439	16.1	8.75

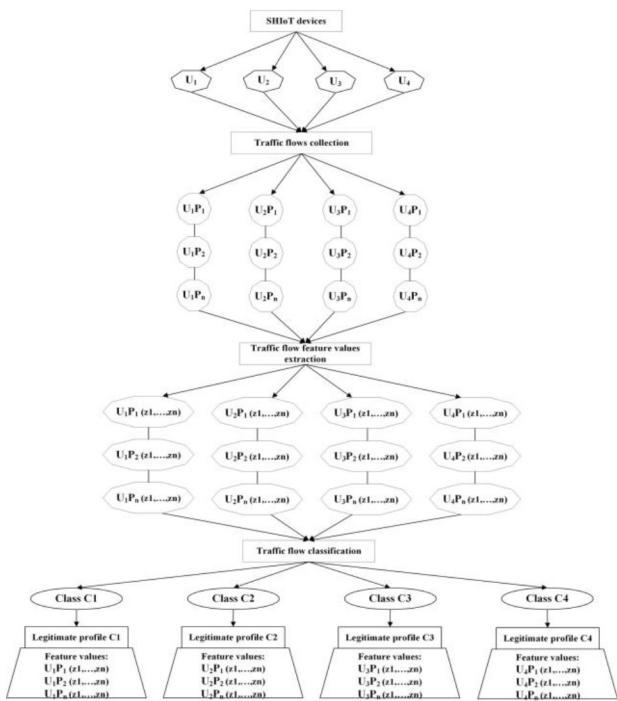
profil lalu lintas yang sah dari perangkat SHIoT, sangat penting untuk memiliki kumpulan data yang mencakup lalu lintas DDoS. Kedua kumpulan ini menjadi dasar untuk mengembangkan model yang efektif guna mendeteksi anomali lalu lintas jaringan seperti lalu lintas DDoS yang dihasilkan oleh perangkat SHIoT.

Mengingat bahwa lalu lintas yang sah berasal dari sumber primer dan sekunder, di mana penulis tidak memiliki akses ke perangkat sumber sekunder, tantangan utama adalah manipulasi perangkat SHIoT untuk menghasilkan lalu lintas DDoS. Oleh karena itu, dalam penelitian ini, untuk menghasilkan lalu lintas DDoS BoNeSi (alat perangkat lunak sumber terbuka) digunakan [25]. Stasiun kerja virtual digunakan untuk menghasilkan lalu lintas DDoS dan membuat kumpulan data lalu lintas tidak sah. Konfigurasi stasiun kerja virtual adalah sebagai berikut: Sistem operasi Linux Ubuntu 19.04 dengan RAM khusus 4 GB, prosesor Intel Core i7-5500U (4x2,40 GHz). Pada Gambar 1, mesin virtual dan alat BoNeSi menunjukkan perangkat SHIoT di jaringan rumah pintar lokal yang menghasilkan lalu lintas DDoS. Untuk alasan praktis, alat BoNeSi digunakan untuk mensimulasikan lalu lintas tidak sah yang dihasilkan oleh perangkat SHIoT untuk meminimalkan risiko membahayakan perangkat sebenarnya. BoNeSi bukan sekadar generator lalu lintas jaringan (seperti yang ditunjukkan oleh dokumentasi alat tersebut), tetapi juga merupakan generator dan alat simulator DoS dan DDoS yang kuat dan efisien. Oleh karena itu, pilihan kami adalah menggunakan alat untuk mensimulasikan lalu lintas yang mirip dengan yang dihasilkan oleh perangkat SHIoT individual sebagai bagian dari botnet. Selain itu, lalu lintas yang tidak sah dihasilkan dalam lingkungan yang terisolasi untuk menghindari pelanggaran hukum Republik Kroasia, Uni Eropa, dan Amerika Serikat. Untuk penelitian ini, tujuan serangan tidak sepenting sumber serangan. Tiga jenis lalu lintas DDoS pada tingkat infrastruktur dihasilkan dan dikumpulkan (UDP, TCP, dan ICMP) karena lebih sering terjadi daripada serangan pada lapisan aplikasi.

Dalam hal jumlah file terkumpul yang berisi siklus 24 jam lalu lintas yang dihasilkan, jumlah paket terkumpul, jumlah data terkumpul, dan keseluruhan waktu pengumpulan data, karakteristik data lalu lintas sah dan DDoS yang awalnya terkumpul ditunjukkan dalam Tabel II.

B. Menentukan Profil Lalu Lintas yang Sah untuk Kelas Perangkat SHIoT

Seperti dibahas sebelumnya, SHIoT adalah lingkungan yang dinamis dan ada di mana-mana, tempat perangkat IoT konsumen baru dengan fungsionalitas berbeda terus-menerus diperkenalkan ke pasar.



Gbr. 2. Proses penentuan profil lalu lintas yang sah untuk perangkat SHIoT kelas.

Oleh karena itu, perangkat SHIoT yang baru dan tidak dikenal mungkin memiliki fungsionalitas yang berbeda dari perangkat SHIoT yang tersedia saat ini.

Hal ini menghadirkan tantangan dalam mengidentifikasi perangkat tersebut dan mengetahui perilaku sahnya, yang menjadi dasar untuk mendeteksi anomali perilaku seperti menghasilkan lalu lintas DDoS.

Untuk mengembangkan model deteksi lalu lintas DDoS berdasarkan kelas perangkat SHIoT yang telah ditetapkan sebelumnya, profil lalu lintas yang sah dari setiap kelas perangkat harus ditetapkan. Dalam pengembangan model deteksi anomali apa pun berdasarkan metode pembelajaran mesin yang diawasi, diperlukan sekumpulan data yang akan mewakili lalu lintas yang sah dan sekumpulan data yang akan mewakili lalu lintas yang tidak sah.

Kelas-kelas perangkat SHIoT yang telah ditetapkan [5] memungkinkan pembentukan profil lalu lintas yang sah dari kelas perangkat tertentu, yang penting dalam pengembangan model deteksi anomali selanjutnya. Dengan demikian, nilai-nilai karakteristik lalu lintas perangkat SHIoT menjadi bagian dari profil yang sah dari kelas perangkat yang diamati. Profil lalu lintas yang sah dari kelas perangkat SHIoT tertentu ditetapkan oleh nilai-nilai karakteristik arus lalu lintas yang ditetapkan oleh model klasifikasi ke kelas perangkat SHIoT tertentu, seperti yang ditunjukkan pada Gambar 2.

Biarkan perangkat SHIoT diwakili oleh $kamu_x$, dan arus lalu lintas yang dihasilkan oleh perangkat tersebut oleh $kamu_xPT_{kamu}$. Setiap perangkat $kamu_x$ direpresentasikan sebagai sekumpulan aliran lalu lintas $kamu_xPT_{kamu}$, yaitu, setiap perangkat berisi serangkaian aliran lalu lintas, $kamu_x = \{kamu_xPT_1, \dots, kamu_xPT_{kamu}\}$. Kemudian, profil lalu lintas yang sah dari setiap kelas C definisikan sebagai sekumpulan arus lalu lintas yang diidentifikasi oleh model klasifikasi sebagai bagian dari kelas C , yaitu $C_m = \{kamu_xPT_1, \dots, kamu_xPT_{kamu}\}$, $m \in \{1, 2, 3, 4\}$. Ketika setiap aliran lalu lintas diwakili oleh

TABEL III
NJARINGAN TRAFIK FRENDAH FSIFAT D DESKRIPSI

Feature name	ID	Feature description	Feature name	ID	Feature description
flowID	z1	Traffic flow ID	max_flowptl	z42	Maximum length of a flow
srcIP	z2	Source IP address	mean_flowptl	z43	Mean length of a flow
src_port	z3	Source communication port	std_flowptl	z44	Standard deviation length of a flow
dstIP	z4	Destination IP address	min_flowiat	z45	Minimum inter-arrival time of packet
dst_port	z5	Destination communication port	max_flowiat	z46	Maximum inter-arrival time of packet
proto	z6	Used communication protocols in traffic flow	mean_flowiat	z47	Mean inter-arrival time of packet
timestamp	z7	Date ad time of traffic flow start	std_flowiat	z48	Standard deviation inter-arrival time of packet
Federation	z8	Duration of the flow in Microsecond	flow_fin	z49	Number of packets with FIN
total_fpackets	z9	Total packets in the forward direction	flow_syn	z50	Number of packets with SYN
total_bpackets	z10	Total packets in the backward direction	flow_RST	z51	Number of packets with RST
total_fpktl	z11	Total size of packet in forward direction	flow_psh	z52	Number of packets with PUSH
total_bpktl	z12	Total size of packet in backward direction	flow_ack	z53	Number of packets with ACK
min_fpktl	z13	Minimum size of packet in forward direction	flow_urg	z54	Number of packets with URG
min_bpktl	z14	Minimum size of packet in backward direction	flow_cwr	z55	Number of packets with CWE
max_fpktl	z15	Maximum size of packet in forward direction	flow_ece	z56	Number of packets with ECE
max_bpktl	z16	Maximum size of packet in backward direction	downUpRatio	z57	Download and upload ratio
mean_fpktl	z17	Mean size of packet in forward direction	avgPacketSize	z58	Average size of packet
mean_bpktl	z18	Mean size of packet in backward direction	fAvgSegmentSize	z59	Average size observed in the forward direction
std_fpktl	z19	Standard deviation size of packet in forward direction	fAvgBytesPerBulk	z60	Average number of bytes bulk rate in the forward direction
std_bpktl	z20	Standard deviation size of packet in backward direction	fAvgPacketsPerBulk	z61	Average number of packets bulk rate in the forward direction
total_fiat	z21	Total time between two packets sent in the forward direction	fAvgBulkRate	z62	Average number of bulk rate in the forward direction
total_biat	z22	Total time between two packets sent in the backward direction	bAvgSegmentSize	z63	Average size observed in the backward direction
min_fiat	z23	Minimum time between two packets sent in the forward direction	bAvgBytesPerBulk	z64	Average number of bytes bulk rate in the backward direction
min_biat	z24	Minimum time between two packets sent in the backward direction	bAvgPacketsPerBulk	z65	Average number of packets bulk rate in the backward direction
max_fiat	z25	Maximum time between two packets sent in the forward direction	bAvgBulkRate	z66	Average number of bulk rate in the backward direction
max_biat	z26	Maximum time between two packets sent in the backward direction	sflow_fpacket	z67	The average number of packets in a sub flow in the forward direction
mean_fiat	z27	Mean time between two packets sent in the forward direction	sflow_fbytes	z68	The average number of bytes in a sub flow in the forward direction
mean_biat	z28	Mean time between two packets sent in the backward direction	sflow_bpacket	z69	The average number of packets in a sub flow in the backward direction
std_fiat	z29	Standard deviation time between two packets sent in the forward direction	sflow_bbytes	z70	The average number of bytes in a sub flow in the backward direction
std_biat	z30	Standard deviation time between two packets sent in the backward direction	min_active	z71	Minimum time a flow was active before becoming idle
fphs_cnt	z31	Number of times the PSH flag was set in packets travelling in the forward direction (0 for UDP)	mean_active	z72	Mean time a flow was active before becoming idle
bpsh_cnt	z32	Number of times the PSH flag was set in packets travelling in the backward direction (0 for UDP)	max_active	z73	Maximum time a flow was active before becoming idle
furg_cnt	z33	Number of times the URG flag was set in packets travelling in the forward direction (0 for UDP)	std_active	z74	Standard deviation time a flow was active before becoming idle
burg_cnt	z34	Number of times the URG flag was set in packets travelling in the backward direction (0 for UDP)	min_idle	z75	Minimum time a flow was idle before becoming active
total_fhlen	z35	Total bytes used for headers in the forward direction	mean_idle	z76	Mean time a flow was idle before becoming active
total_bhflen	z36	Total bytes used for headers in the forward direction	max_idle	z77	Maximum time a flow was idle before becoming active
fPktsPerSecond	z37	Number of forward packets per second	std_idle	z78	Standard deviation time a flow was idle before becoming active
bPktsPerSecond	z38	Number of backward packets per second	Init_Win_bytes_forward	z79	The total number of bytes sent in initial window in the forward direction
flowPktsPerSec ond	z39	Number of flow packets per second	Init_Win_bytes_backwa rd	z80	The total number of bytes sent in initial window in the backward direction
flowBytesPerSecond	z40	Number of flow bytes per second	Act_data_pkt_forward	z81	Count of packets with at least 1 byte of TCP data payload in the forward direction
min_flowptl	z41	Minimum length of a flow	min_seg_size_forward	z82	Minimum segment size observed in the forward direction

karakteristiknya dari, dapat diamati sebagai sekumpulan nilai fitur yang mewakili arus lalu lintas yang diamati, $kamu \times PT_{kamu} = \{z(kamu) \times PT_{kamu}\}$. Bahasa Indonesia: . . . , $\{z(kamu) \times PT_{kamu}\}_N\}$.

Selain fakta bahwa dimungkinkan untuk menentukan kelas perangkat SHIoT (lihat [5]), juga memungkinkan untuk mengklasifikasikan perangkat, yaitu, arus lalu lintas yang dihasilkan oleh perangkat tersebut menggunakan model klasifikasi yang dikembangkan dan fitur arus lalu lintas dengan akurasi klasifikasi yang tinggi (99,7956%). Hasil tersebut memungkinkan pembuatan profil lalu lintas yang sah dari kelas perangkat SHIoT tertentu [6].

C. Pembentukan Data Set untuk Pengembangan Model Deteksi Trafik DDoS

Kelas perangkat SHIoT yang ditentukan oleh penelitian ini memungkinkan identifikasi afiliasi kelas perangkat berdasarkan aliran lalu lintas yang dihasilkan oleh perangkat tersebut. Hal ini juga memungkinkan pembuatan profil lalu lintas yang sah karena setiap aliran lalu lintas

yang ditetapkan ke salah satu dari empat kelas yang ditentukan oleh model klasifikasi menjadi bagian dari suatu set yang mewakili profil lalu lintas yang sah dari kelas yang sama. Untuk mengembangkan suatu model guna mendeteksi lalu lintas jaringan DDoS (yang tidak sah), digunakan metode LMT. Untuk penerapan metode dan pemrosesan data, kami menggunakan perangkat lunak WEKA, serta set data yang mewakili profil lalu lintas normal yang dihasilkan dari model klasifikasi perangkat SHIoT dan set data lalu lintas DDoS yang tidak sah.

Empat set data (yaitu, C1DDoS, C2DDoS, C3DDoS, dan C4DDoS) yang berisi vektor gabungan karakteristik profil lalu lintas yang sah dari setiap kelas perangkat SHIoT dan lalu lintas DDoS dibangun. Awalnya, keempat set berisi nilai semua fitur aliran lalu lintas independen (total 83) yang tercantum dalam Tabel III. Untuk ekstraksi fitur, kami menggunakan alat CICFlowMeter [26]. Jumlah dan distribusi aliran lalu lintas yang sah dan DDoS dalam set data diseimbangkan dan berdasarkan profil lalu lintas yang sah yang berasal dari model klasifikasi perangkat SHIoT yang ditunjukkan pada [5].

TABEL IV
PBUATANPRESENTASI DARI DATA SETSKAMUSED DI DALAM DPERKEMBANGAN SEBUAH DDHAI SDEKSEKUSI MODEL

Feature vector	z8	z9	z10	z11	z12	...	z23	...	z36	...	z83	Class
C1DDoS												
1	110,176,901	5	4	372	648	...	13,800,000	...	113,851	...	54,900,000	C1
2	114,974,487	13	1	0	0	...	8,844,190,000,000	...	0	...	5,402,915	DDoS
C2DDoS												
1	32,421,069	12	12	2,973	6,626	...	1,409,612	...	76	...	0	C2
2	119,994,320	8	5	724	913	...	9,999,527	...	214,652	...	29,800,000	C2
C3DDoS												
1	91,127,887	3	3	33	95	...	18,225,577	...	120,725	...	90,875,582	C3
2	47,780	5	5	84	474	...	5,309	...	2	...	0	C3
C4DDoS												
1	119,436,915	16	18	5,158	527	...	3,619,300	...	136	...	6,167,447	C4
2	119,436,449	16	18	5,158	527	...	3,619,286	...	136	...	6,138,410	C4

Seperti halnya pengembangan model pembelajaran mesin, tujuannya adalah untuk menggunakan fitur-fitur independen tersebut, yang perubahannya memiliki dampak terbesar pada perubahan fitur dependen. Penting juga untuk mengurangi fitur-fitur yang dapat menyebabkan bias model. Oleh karena itu, seperti halnya pengembangan model klasifikasi perangkat SHIoT, fitur-fitur independen z1–z7 mewakili pengenal aliran lalu lintas dan berisi informasi tentang alamat IP sumber dan tujuan, protokol yang digunakan, dan waktu pembuatan aliran lalu lintas yang dihapus dari set data awal. Hasilnya, 76 fitur independen diperoleh, yang akan diamati untuk pengembangan model lebih lanjut, dan yang penyajian sebagiannya ditunjukkan pada Tabel IV. Tabel tersebut sebagian menunjukkan set data yang digunakan untuk mengembangkan model deteksi DDoS. Setiap set terdiri dari fitur-fitur independen dari setiap nilai aliran lalu lintas dan fitur dependen terkait yang mewakili kelas. Dalam hal ini, kelas tersebut adalah biner, yaitu, dapat mengambil dua nilai (0, 1), yang menunjukkan aliran lalu lintas sebagai sah untuk kelas yang diamati atau tidak sah, yaitu, aliran lalu lintas yang dibuat sebagai hasil dari pembuatan lalu lintas DDoS.

Pendekatan ini diperlukan untuk pengembangan model lebih lanjut dengan penerapan metode pembelajaran mesin terbimbing. Kami kemudian memanfaatkan metode LMT dalam pengembangan model deteksi DDoS kami. Metode LMT, yang dikembangkan pada tahun 2003 [27], adalah metode peningkatan pembelajaran mesin terbimbing yang merupakan gabungan dari dua metode klasifikasi yang umum digunakan: 1) regresi logistik dan 2) pohon keputusan, untuk meningkatkannya. Prinsip kerja dasar metode ini terdiri dari pembuatan pohon keputusan dan pembentukan model regresi logistik di simpul pohon. Model regresi logistik saling membangun menjadi satu model tunggal. Dengan cara ini, metode regresi logistik memperkirakan probabilitas milik vektor fitur individual ke kelas yang ditentukan. Untuk fitur numerik (seperti yang ditemukan dalam set data yang disajikan), fitur yang mewakili simpul di mana pembagian adalah yang "paling murni" dipilih. Ini menyiratkan bahwa jumlah maksimum vektor fitur milik satu kelas ketika nilai fitur yang dipilih berada di bawah ambang batas nilai yang ditentukan dan ke kelas lain jika fitur yang dipilih diamati di atas ambang batas nilai yang ditentukan. Model LMT terdiri dari struktur pohon keputusan

berisi node internal N dan satu set simpul terminal T . S mewakili keseluruhan kumpulan data dengan semua fitur [28]. Pohon keputusan kemudian membagi kumpulan tersebut menjadi subset (wilayah) yang terpisah S_T . Setiap wilayah diwakili oleh simpul terminal pohon seperti yang ditunjukkan berikut ini:

$$S = \bigcup_{T \in T} S_T \quad \text{Basis: } S \cap S_T = \emptyset \text{ untuk } T \neq T' \quad (1)$$

Di mana

- S himpunan semua vektor fitur; subhimpunan terpisah dari S
- S_T vektor fitur; simpul terminal dari sekumpulan simpul
- T terminal T .

Tidak seperti pohon keputusan klasik, metode LMT mengasosiasikan fungsi regresi logistik, f , alih-alih penunjukan kelas, dengan simpul terminal $T \in T$. Regresi logistik mempertimbangkan subset $Bahasa Indonesia: Z \subseteq Bahasa Indonesia: Z$ dari semua fitur independen dalam set data dan memodelkan probabilitas untuk menjadi anggota kelas berdasarkan

$$\Pr(G_j | Bahasa Indonesia: X = x) = f_j \quad \text{Basis: } f_j(x) \quad (2)$$

$$\sum_{i=1}^M f_i(x) = \text{sebuah} \begin{cases} \text{Jika } 0 \\ \text{dari } 1 \end{cases} \quad \text{Basis: } f_i(x) \quad (3)$$

Di mana

- f_j koefisien fitur independen j dari Bahasa Indonesia:
- x fitur independen dari sekumpulan fitur independen $Bahasa Indonesia: Z = \{dari \dots\}$
- i Bahasa Indonesia: $\dots, dari i\}$.

Model LMT akhir mengambil bentuk yang diberikan oleh

$$\sum_{T \in T} \left\{ \begin{array}{ll} 1 & \text{jika } x \in S_T \\ 0 & \text{angka lain tidak.} \end{array} \right. \quad (4)$$

Menurut Landwehr dan lain-lain. [28], tujuan metode ini adalah untuk mengadaptasi data sehingga pohon keputusan logistik digeneralisasikan (dipangkas) ke tingkat satu model regresi logistik, yaitu, ke simpul akar pohon keputusan jika memungkinkan, mengingat kumpulan data tempat metode diterapkan.

Pemilihan fitur independen yang relevan dari kumpulan semua fitur saat menggunakan metode LMT tidak perlu dilakukan

dilakukan secara terpisah karena metode ini menyesuaikan (menyesuaikan) fungsi regresi ke setiap fitur independen menggunakan kesalahan kuadrat terkecil. Menurut kriteria ini, model akhir mencakup fitur-fitur yang menghasilkan kesalahan kuadrat terkecil, seperti yang ditunjukkan pada Tabel V. Dengan menggunakan lingkungan perangkat lunak WEKA, metode LMT yang dijelaskan diimplementasikan pada keempat set data kami (yaitu, C1DDoS, C2DDoS, C3DDoS, dan C4DDoS) untuk mengembangkan model LMT untuk setiap kelas perangkat SHIoT.

1) Model LMT untuk Perangkat SHIoT Kelas C1:Dengan menerapkan metode LMT menggunakan lingkungan pemrograman WEKA, fitur-fitur independen dengan pengaruh terbesar terhadap fitur dependen dipilih, dan model regresi logistik dikembangkan karena pohon keputusan digeneralisasi ke simpul akar. Oleh karena itu, pada simpul akar pohon keputusan, model LMT yang sesuai didefinisikan.

$$\Pr(G=C1 | X=X) = \frac{\text{Bahasa Inggris}:F_{C1}(X)}{\text{Bahasa Inggris}:F_{C1}(X)+\text{Bahasa Inggris}:F_{Serangan DDoS}(X)}$$

$$\Pr(G=\text{Serangan DDoS} | X=X) = \frac{\text{Bahasa Inggris}:F_{Serangan DDoS}(X)}{\text{Bahasa Inggris}:F_{C1}(X)+\text{Bahasa Inggris}:F_{Serangan DDoS}(X)}.$$

Keduanya F_{C1} Dan $F_{Serangan DDoS}$ Fungsi regresi logistik digunakan untuk menentukan probabilitas untuk masuk ke dalam suatu kelas dengan memodelkan pengaruh fitur independen terhadap fitur dependen. Untuk kelas C1, model regresi logistik mengambil bentuk yang ditunjukkan oleh

$$F_{C1}(X) = -1.37 + 0.\text{tanggal } 02\text{Bahasa Indonesia:z14} + 0.01\text{Bahasa Indonesia:z18} + 3.29\text{Bahasa Indonesia:z38}$$

$$+ \text{ angka } 0.01\text{Bahasa Indonesia:z46} + (3\text{Bahasa Indonesia:z72})\text{Bahasa Indonesia:z50} + (-1.\text{tanggal } 08)\text{Bahasa Indonesia:z10} + (-0.05)$$

$$+ (-0.2)\text{Bahasa Indonesia:z54} + 0.88\text{Bahasa Indonesia:z58} + 0.57\text{Bahasa Indonesia:z59} + 0.57\text{Bahasa Inggris:z274}$$

$$F_{Serangan DDoS}(X) = -F_{C1}(X) = 1.37 + (-0.\text{tanggal } 02)\text{Bahasa Indonesia:z14} + (-0.01)$$

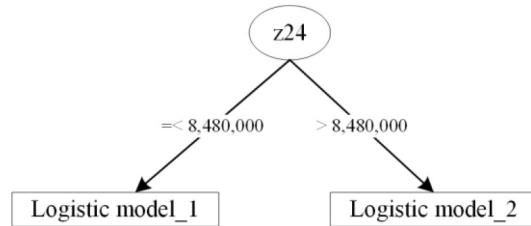
$$+ \text{ Bahasa Inggris:z18} + (-3.29)\text{Bahasa Indonesia:z38} + (-0.01)\text{Bahasa Indonesia:z46} + 3.72$$

$$+ \text{ Bahasa Inggris:z46} + 0.01\text{Bahasa Indonesia:z50} + (-0.39)\text{Bahasa Indonesia:z51} + (-0.28)\text{Bahasa Indonesia:z54} + (-0.97)$$

$$+ (-0.57)\text{Bahasa Indonesia:z58} + 0.57\text{Bahasa Indonesia:z59} + 0.57\text{Bahasa Inggris:z274}$$

Model ini mencakup fitur-fitur independen yang metode deviasi kuadrat terkecilnya menentukan pengaruh terbesar pada perubahan fitur dependen. Pengaruh fitur independen terhadap fitur dependen didefinisikan oleh koefisien yang cukup untuk setiap fitur. Koefisien yang ditetapkan menunjukkan bahwa satu unit peningkatan fitur independen akan mengubah fitur dependen dengan logaritma tata letak koefisien regresi logistik, sementara variabel independen lainnya akan tetap tidak berubah. Misalnya, koefisien yang ditetapkan untuk fitur independen z14 adalah -0,02 dan merupakan estimasi perubahan (peningkatan atau penurunan; sebagaimana ditentukan oleh tanda), dalam hal ini, penurunan jumlah logaritma fitur dependen jika fitur independen z14 meningkat sebesar satu unit dan fitur independen lainnya dalam model tetap tidak berubah.

2) Model LMT untuk Perangkat SHIoT Kelas C2:Model LMT deteksi DDoS untuk perangkat SHIoT kelas C2 dikembangkan dengan cara yang sama seperti model yang dijelaskan sebelumnya untuk kelas C1. Karena perangkat SHIoT yang berbeda termasuk dalam kelas yang berbeda, secara intuitif jelas bahwa arus lalu lintas yang dihasilkan oleh perangkat SHIoT kelas C2 berbeda dalam hal nilai fitur dari arus lalu lintas perangkat SHIoT kelas C1. Oleh karena itu, model yang dikembangkan untuk kelas perangkat ini, meskipun didasarkan pada metode yang sama, memiliki perbedaan tertentu. Hal ini terutama



Gbr. 3. Contoh penerapan metode LMT dalam klasifikasi vektor fitur.

mengacu pada tampilan pohon keputusan dan fitur-fitur independen yang disertakan dalam model, serta koefisien yang ditambahkan ke fitur-fitur ini. Ini berarti bahwa fitur-fitur independen yang mempengaruhi perubahan fitur dependen berbeda dari kelas ke kelas perangkat SHIoT.

Di sisi lain, kelas yang berbeda mungkin memiliki fitur independen yang relevan, tetapi mereka juga memiliki perbedaan koefisien dengan tingkat pengaruh yang berbeda. Untuk kelas C2, pohon keputusan berbeda dari kelas C1 karena tidak mungkin untuk mendefinisikan model regresi logistik pada simpul akar yang akan memberikan kinerja model LMT yang memuaskan. Dalam kasus ini, pohon keputusan digeneralisasikan ke tiga simpul (satu simpul akar dan dua simpul terminal), seperti yang ditunjukkan sebelumnya pada Gambar 3. Oleh karena itu, dua model logistik didefinisikan pada simpul terminal. LM1, menurut ekspresi (9) dan (10) dan LM2, menurut ekspresi (11) dan (12), yang diterapkan

tergantung pada kondisi yang terpenuhi saat melakukan percabangan pohon keputusan

$$F_{C2}(X) = -16.07 + 3.42\text{Bahasa Indonesia:z10} + 4.35\text{Bahasa Indonesia:z38} + 0.01\text{Bahasa Indonesia:z50} + 0.57\text{Bahasa Inggris:z24} + (-0.05)\text{Bahasa Inggris:z46} + (-0.39)\text{Bahasa Inggris:z51} + (-0.28)\text{Bahasa Inggris:z54} + 0.97\text{Bahasa Inggris:z58} + 14.58\text{Bahasa Inggris:z59}$$

$$F_{Serangan DDoS}(X) = -F_{C2}(X) = 16.07 + (-3.42)\text{Bahasa Indonesia:z10} + (-4.35)\text{Bahasa Indonesia:z38} + (-0.01)\text{Bahasa Indonesia:z50} + (-0.57)\text{Bahasa Inggris:z24} + (0.05)\text{Bahasa Inggris:z46} + (0.39)\text{Bahasa Inggris:z51} + (0.28)\text{Bahasa Inggris:z54} + (-0.97)\text{Bahasa Inggris:z58} + (-14.58)\text{Bahasa Inggris:z59}$$

$$F_{C2}(X) = -20.68 + 2.32\text{Bahasa Indonesia:z38} + 0.01\text{Bahasa Indonesia:z46} + (-2.\text{tanggal } 06)\text{Bahasa Indonesia:z50} + 0.39\text{Bahasa Inggris:z24} + 0.28\text{Bahasa Inggris:z51} + 0.84\text{Bahasa Inggris:z54} + 0.84\text{Bahasa Inggris:z58}$$

$$F_{Serangan DDoS}(X) = -F_{C2}(X) = 20.68 + (-2.32)\text{Bahasa Indonesia:z38} + (-0.01)\text{Bahasa Indonesia:z46} + (2.\text{tanggal } 06)\text{Bahasa Indonesia:z50} + (-0.39)\text{Bahasa Inggris:z24} + (-0.28)\text{Bahasa Inggris:z51} + (-0.84)\text{Bahasa Inggris:z54} + (-0.84)\text{Bahasa Inggris:z58}$$

Perlu dicatat bahwa model LMT untuk mendeteksi anomali lalu lintas jaringan untuk perangkat SHIoT yang termasuk dalam kelas C2 terdiri dari pohon keputusan yang pada simpul terminalnya terdapat dua model logistik, dan penggunaannya bergantung pada kondisi mana yang memenuhi vektor fitur yang diamati mengenai nilai fitur independen z24. Hal ini juga bergantung pada kondisi ini di mana fitur independen akan disertakan dalam model logistik dan koefisien yang terkait dengan fitur-fitur ini.

3) Model LMT untuk Perangkat SHIoT Kelas C3:Untuk perangkat SHIoT kelas C3 guna mendeteksi anomali lalu lintas jaringan, model LMT dikembangkan berdasarkan prinsip yang diterapkan pada kelas C1 dan C2. Sedangkan untuk kelas C1, pohon keputusan digeneralisasikan ke simpul akar yang dikaitkan dengan satu model logistik.

TABEL V
DIMPLY DARISAYATIDAK TERGANTUNG PADA ORANG LAINFITURSAYATERMASUK DALAM TTL

LMT-C1		LMT-C2		LMT-C3		LMT model					
						Logistics models					
LM1	LM1	LM2	LM1	LM1	LM2	LM3	LM4	LM5	LM6		
z14	z10	z38	z14	z10	z10	z10	z10	z10	z10		
z18	z38	z46	z38	z16	z16	z16	z16	z16	z38		
z38	z41	z50	z45	z20	z20	z20	z20	z38	z50		
z46	z46	z51	z46	z36	z36	z36	z38	z41	z51		
z50	z50	z54	z50	z38	z38	z38	z41	z42	z54		
z51	z51	z58	z51	z41	z41	z41	z42	z45	z58		
z58	z54		z54	z42	z42	z42	z45	z50	z74		
z74	z58		z58	z45	z43	z44	z45	z51	z54		
	z74		z74	z50	z44	z45	z51	z54			
				z51	z45	z46	z54		z58		
				z54	z46	z50	z58		z74		
				z58	z50	z51	z74				
				z74	z51	z58					
					z54	z73					
					z58	z74					
					z74						

Bentuk akhir model LMT, dengan fitur dan koefisien independen paling signifikan untuk kelas C3, ditunjukkan oleh

$$Fc_3(X) = -1.01 + 0. \text{tangga} 03 \text{Bahasa Indonesia:z14} + 2.91 \text{Bahasa Indonesia:z38} + 0.01 \text{Bahasa Indonesia:z50}$$

$$+ \text{ angka } 0 \text{ tangga } 02 \text{ Bahasa Indonesia:z46} + /2 \text{ Bahasa Indonesia:z50} + /1.82 \text{ Bahasa Indonesia:z54}$$

$$+ 1.12 \text{ Bahasa Indonesia:z54} + 0.87 \text{ Bahasa Indonesia:z58} + 0. \text{tangga } 04 \text{ Bahasa Indonesia:z50}$$

$$\text{Perangkat DDoS}(X) = -Fc_3(X) = 1.01 + /0. \text{tangga } 03 \text{ Bahasa Indonesia:z14} + /2.91 \text{ Bahasa Indonesia:z38}$$

$$+ /0.01 \text{ Bahasa Indonesia:z45} + /0. \text{tangga } 02 \text{ Bahasa Indonesia:z46} + 2 \text{ Bahasa Indonesia:z50}$$

$$+ 1.82 \text{ Bahasa Indonesia:z51} + /1.12 \text{ Bahasa Indonesia:z54} + /0.87 \text{ Bahasa Indonesia:z58}$$

$$+ /0. \text{tangga } 04 \text{ Bahasa Indonesia:z50}$$

(14)

Model tersebut mencakup total sembilan fitur independen ($z14, z38, z45, z46, z50, z51, z54, z58, z74$) yang ditentukan dengan metode kuadrat terkecil yang memiliki dampak terbesar terhadap perubahan fitur dependen.

4) Model LMT untuk Perangkat SHIoT Kelas C4: Perangkat Kelas C4, karena lebih tinggi $C_{kamusindeks}$, menghasilkan lalu lintas dan arus lalu lintas yang karakteristiknya lebih sulit dibedakan dari anomali lalu lintas jaringan seperti lalu lintas DDoS.

Tingkat predikabilitas lalu lintas yang lebih rendah disebabkan oleh mode operasi perangkat, seperti tingkat interaksi pengguna yang tinggi, pemutaran konten audio/video, dan sejenisnya. Hal ini menghasilkan model LMT yang lebih kompleks yang tidak dapat digeneralisasi ke simpul akar, tetapi terdiri dari 11 simpul atau enam simpul terminal. Model regresi logistik didefinisikan pada setiap cabang pohon keputusan yang berakhir di simpul terminal.

Dalam kasus ini, ini berarti bahwa model LMT terdiri dari total lima titik percabangan dan enam model regresi logistik. Model LMT yang berisi pohon keputusan dan model regresi logistik terkait dengan fitur independen relevan terpilih dan koefisien terkait, seperti yang ditunjukkan pada Gambar 4.

D. Prinsip Kerja Model yang Dikembangkan untuk Mendeteksi Trafik Jaringan DDoS Ilegal

Pekerjaan model deteksi lalu lintas DDoS ilegal yang dikembangkan berlangsung dalam dua fase. Fase pertama adalah

prasyarat untuk deteksi lalu lintas DDoS selanjutnya pada fase kedua dan melibatkan klasifikasi perangkat SHIoT berdasarkan arus lalu lintas yang dihasilkan. Hasil model klasifikasi multikelas menunjukkan bahwa perangkat SHIoT dapat diklasifikasikan ke dalam salah satu dari empat kelas yang telah ditetapkan sebelumnya terkait arus lalu lintas yang dihasilkannya dengan akurasi 99,79%.

Setelah perangkat berhasil diklasifikasikan, arus lalu lintas yang baru dihasilkan diperiksa berdasarkan model LMT untuk mendeteksi lalu lintas DDoS yang tidak sah, yang menentukan apakah arus lalu lintas ini termasuk dalam kelas yang dikenali atau mewakili anomali lalu lintas jaringan.

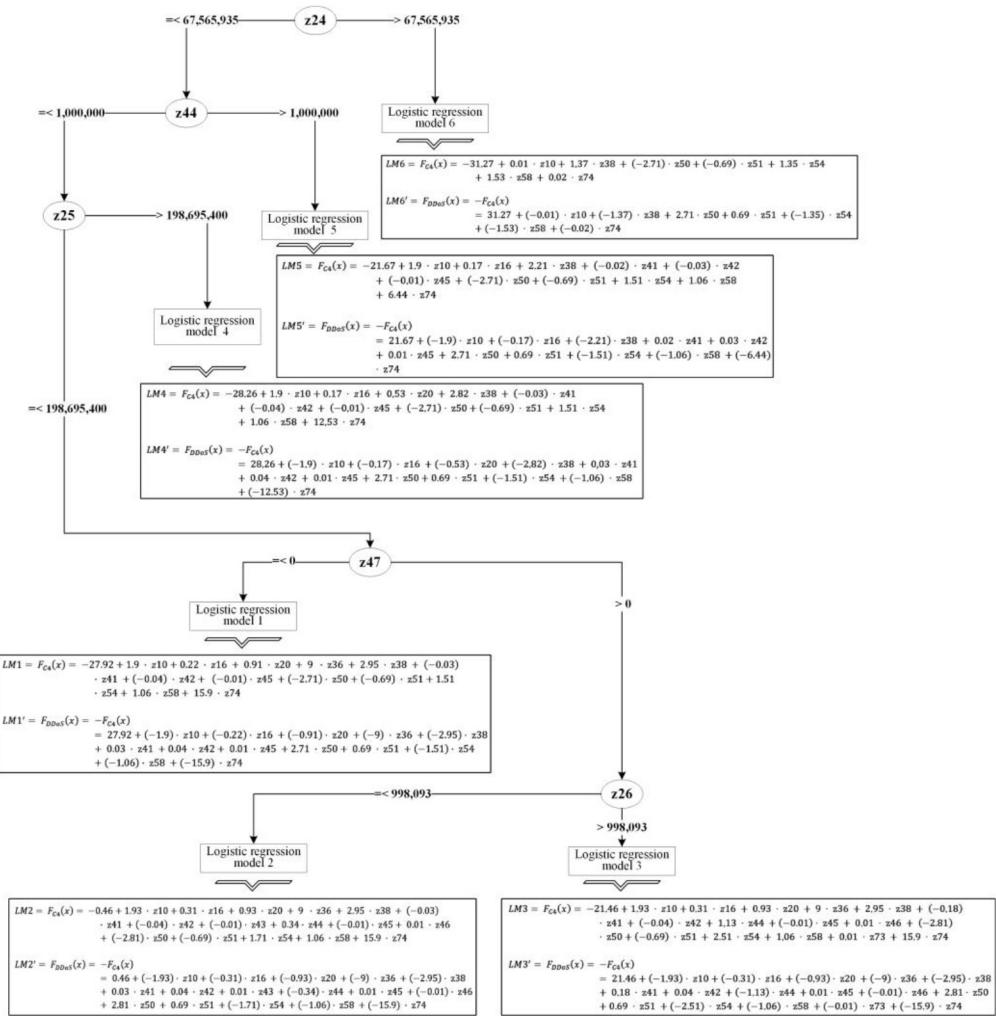
Dasar untuk pengembangan model deteksi lalu lintas DDoS untuk kelas tertentu adalah profil lalu lintas yang sah dari kelas tertentu, yang dihasilkan dari pekerjaan model klasifikasi multikelas pada fase pertama.

Dengan demikian, nilai arus lalu lintas yang diklasifikasikan ke dalam kelas-kelas tertentu yang telah ditetapkan sebelumnya juga menjadi bagian dari profil lalu lintas yang sah dari kelas-kelas ini. Bergantung pada kelas perangkat SHIoT yang sesuai, model LMT individual dapat mendeteksi penyimpangan atau anomali dari profil lalu lintas normal yang ada dengan akurasi tinggi (LMT-C1 = 99,99%, LMT-C2 = 99,92%, LMT-C3 = 99,97%, dan LMT-C4 = 99,95%) dan menggunakan serangkaian karakteristik arus lalu lintas independen yang berbeda.

IV. RASIALANALISIS DAN DISKUSI

Pengembangan model deteksi DDoS berdasarkan karakteristik lalu lintas dan kelas perangkat menunjukkan pentingnya mengenali kelas tempat perangkat SHIoT berada sebagai aktivitas mendasar untuk lebih mengenali anomali dalam lalu lintas jaringan seperti lalu lintas DDoS. Menurut model yang disajikan di bagian sebelumnya, jelas bahwa tidak semua fitur independen sama pentingnya dalam mendeteksi anomali untuk kelas tertentu. Demikian pula, fitur tertentu dalam satu kelas mungkin relevan sementara dilihat dari aspek kelas lain, fitur tersebut tidak harus relevan.

Contohnya adalah melihat setiap kelas berbeda menurut jumlah fitur independen yang relevan, dan terbukti pula bahwa fitur yang sama tidak relevan dalam pendekripsi anomali untuk setiap kelas.



Gambar 4. Model LMT dari model deteksi DDoS untuk kelas C4.

Lebih jauh lagi, nilai ambang batas fitur independen individual yang menentukan percabangan pohon keputusan berbeda untuk setiap kelas. Seperti yang ditunjukkan pada Gambar 3 dan 4, percabangan pada pohon keputusan terjadi berdasarkan nilai ambang batas fitur z24, yang mewakili deviasi standar

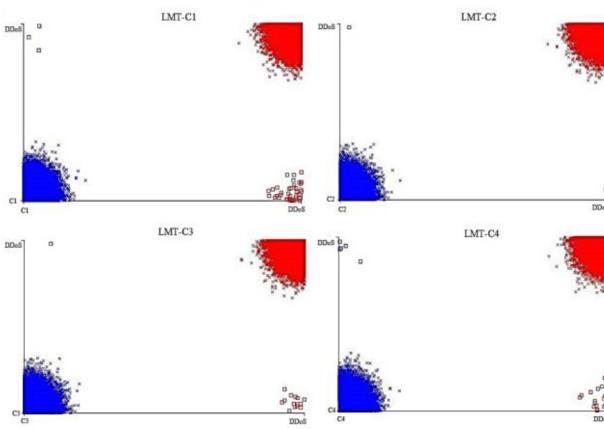
Waktu antar kedatangan paket dalam aliran lalu lintas yang diamati dinyatakan dalam mikrodetik ($\hat{\mu}$). Dalam melakukan hal ini, algoritma C4.5 digunakan, yang memilih nilai ambang batas fitur independen yang memungkinkan pembagian vektor fitur paling murni dalam set [29]. Jadi, misalnya, nilai ambang batas fitur z24 dalam model LMT untuk kelas C2 berbeda dari nilai ambang batas fitur yang sama untuk kelas C4.

Untuk mengevaluasi perilaku model terhadap data yang tidak termasuk dalam proses pembelajaran, setiap versi model LMT divalidasi menggunakan k -fold validasi silang dengan $k=10$. Validasi silang adalah teknik matematika untuk mengevaluasi keberhasilan model pembelajaran mesin pada data baru yang tidak diketahui. Pendekatan ini digunakan untuk menguji keluaran model pada data yang tidak digunakan selama proses pembelajaran. Model diperluas secara berulang k kali lipat dari kumpulan data dengan cara ini. Kumpulan data dibagi menjadi k bagian dalam

TABEL VI
AKURATANDBERKEMBANGMODEL DAN Bahasa Inggris: KAPLIKASI CEFSIEN

Model	LMT-C1		LMT-C2	
	Accurately classified examples	99,921%	59,660	99,996%
Misclassified examples	44	0,0784%	2	0,0034%
Kappa coefficient (κ)	0,9984		0,9999	
Total	56,136		59,662	
Model	LMT-C3		LMT-C4	
Accurately classified examples	58,646	99,974%	59,879	99,958%
Misclassified examples	15	0,0256%	25	0,0417%
Kappa coefficient (κ)	0,9995		0,9992	
Total	58,661		59,904	

setiap iterasi. Sisanya $k-1$ bagian dari set dikelompokkan menjadi subset untuk pembelajaran model, sementara satu bagian dari set digunakan untuk menguji model [30]. Metrik validasi (akurasi, statistik kappa, rasio positif-benar (TPR), rasio positif-salah (FPR), presisi, F -pengukuran, Karakteristik Operasi Penerima ROC, dan Kurva Penarikan Presisi PRC) sering digunakan untuk menguji model klasifikasi pembelajaran mesin.



Gambar 5. Visualisasi kesalahan model klasifikasi LMT untuk kelas yang sesuai.

A. Akurasi Model Klasifikasi LMT yang Dikembangkan

Contoh positif-benar (TP), contoh negatif-benar (TN), contoh positif-salah (FP), dan contoh negatif-salah (FN) mencerminkan bagian dari contoh yang diklasifikasikan dengan benar dalam kumpulan semua contoh.

$$\text{Akut} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (15)$$

Di mana

Akut proporsi contoh yang diklasifikasikan secara akurat dalam kumpulan semua contoh;

T.P. jumlah contoh positif benar;

Bahasa Inggris: jumlah contoh negatif benar;

Bahasa Inggris: jumlah contoh positif salah;

Bahasa Inggris: jumlah contoh negatif salah.

Berdasarkan akurasi klasifikasi, keempat model menunjukkan kinerja tinggi, yang berarti bahwa berdasarkan aliran yang diamati, keempat model dapat menentukan dengan akurasi tinggi apakah aliran lalu lintas merupakan hasil komunikasi perangkat yang sah, atau perangkat tersebut menghasilkan lalu lintas DDoS.

Berdasarkan Tabel VI, akurasi tinggi dari keempat versi model LMT yang dikembangkan untuk setiap kelas perangkat SHIoT dapat diamati. Kesalahan dalam klasifikasi keempat versi model LMT divisualisasikan dan ditunjukkan pada Gambar 5.

Gambar 5 menunjukkan bahwa model deteksi paling akurat untuk kelas C2 dan kinerja terendah diamati pada model LMT-C1. Dari gambar yang diberikan, diamati bahwa kesalahan untuk keempat model lazim terjadi dalam mengklasifikasikan instans lalu lintas DDoS, yang menunjukkan perlunya pemodelan kelas ini yang lebih baik dalam penelitian mendatang.

Untuk menunjukkan keakuratan klasifikasi dengan lebih jelas, matriks kebingungan digunakan untuk semua versi model yang dikembangkan. Matriks kebingungan adalah metrik kinerja untuk model klasifikasi pembelajaran mesin dengan dua atau lebih kelas sebagai keluaran, dan berfungsi sebagai dasar untuk metrik lainnya. Dengan demikian, model LMT untuk kelas perangkat C1 menunjukkan keakuratan 99,9216%, atau 56.092 arus lalu lintas yang diklasifikasikan secara akurat, sebagai DDoS atau arus lalu lintas yang secara sah milik perangkat SHIoT dari kelas C1. Sebanyak 44 arus lalu lintas salah diklasifikasikan, yaitu 0,0784% dalam total set 56.136.

TABEL VII
CKONFLIKMATRIX DARILMT MODEL UNTUKCGadis-gadisC1DANC2

<i>Predicted class affiliation</i>		<i>Actual class affiliation</i>
Class C1	DDoS	
28,065	3	Class C1
41	28,027	DDoS
<i>Predicted class affiliation</i>		
Class C2	DDoS	
29,830	1	Class C2
1	29,830	DDoS

TABEL VIII
CKONFLIKMATRIX DARILMT MODEL UNTUKCGadis-gadisC3DANBahasa Indonesia: C4

<i>Predicted class affiliation</i>		<i>Actual class affiliation</i>
Class C3	DDoS	
29,329	1	Class C3
14	29,317	DDoS
<i>Predicted class affiliation</i>		
Class C4	DDoS	
29,947	5	Class C4
20	29,932	DDoS

44 arus lalu lintas yang salah diklasifikasikan, 41 diprediksi termasuk dalam arus lalu lintas sah kelas C1, sementara tiga arus lalu lintas diklasifikasikan sebagai lalu lintas DDoS, seperti yang ditunjukkan oleh matriks kebingungan pada Tabel VII.

Selain akurasi yang tinggi, model LMT untuk kelas perangkat C1 juga menunjukkan koefisien kappa (*aku*=angka 0,9984), yang menunjukkan kinerja model tinggi.

Versi model LMT yang dikembangkan untuk kelas C2 menunjukkan akurasi tinggi (99,9966%), ditunjukkan pada Tabel VI. Ini menyiratkan 59.660 arus lalu lintas yang diklasifikasikan secara akurat dalam satu set 59.662 arus lalu lintas. Kesalahan klasifikasi adalah 0,0034%, yaitu, dua arus lalu lintas, dengan satu salah ditetapkan ke kelas C2 dan yang lainnya ke lalu lintas DDoS, yang terbukti dari matriks kebingungan yang ditunjukkan pada Tabel VII. Jumlah koefisien kappa adalah 0,9999, yang menunjukkan keberhasilan tinggi dari versi model LMT ini.

Model klasifikasi LMT yang dikembangkan untuk kelas C3 memberikan akurasi 99,9744%, seperti yang ditunjukkan pada Tabel VI. Oleh karena itu, dari 58.661 aliran lalu lintas, 15 salah diklasifikasikan, atau 0,0256%, sementara 58.646 diklasifikasikan secara akurat. Menurut matriks kebingungan yang ditunjukkan pada Tabel VIII, satu aliran lalu lintas salah diklasifikasikan sebagai lalu lintas DDoS, sementara 14 aliran lalu lintas salah diklasifikasikan sebagai bagian dari lalu lintas kelas C3 yang sah.

Besarnya koefisien kappa sebesar 0,9995, seperti pada versi model LMT sebelumnya, menunjukkan kinerjanya yang tinggi. Versi terbaru model LMT, yang dikembangkan untuk kelas C4, menunjukkan akurasi sebesar 99,9583% yang menyiratkan 59.879 arus lalu lintas yang diklasifikasikan dengan benar. Oleh karena itu, 25 arus lalu lintas salah diklasifikasikan, lima sebagai lalu lintas DDoS dan 20 sebagai lalu lintas kelas C4 yang sah, seperti yang ditunjukkan oleh matriks kebingungan pada Tabel VIII. Keberhasilan model yang diukur dengan koefisien kappa adalah 0,9992, terlihat pada Tabel VI.

B. Analisis Kinerja Berdasarkan Hasil Model Positif dan Negatif

Analisis lebih lanjut dan evaluasi kinerja model LMT yang dikembangkan dilakukan dengan menggunakan metrik berbasis

TABEL IX

HAIULASANLMT Model Bahasa Indonesia: VALIDASI MKENYAMANAN(TPR DAN(FPR))

Class	True positive rate (TPR)			
	LMT-C1	LMT-C2	LMT-C3	LMT-C4
C1/C2/C3/C4	1	1	1	1
DDoS	0,999	1	1	0,999
Class	False positive rate (FPR)			
	LMT-C1	LMT-C2	LMT-C3	LMT-C4
C1/C2/C3/C4	0,001	0	0	0,001
DDoS	0	0	0	0

TABEL X

HAIULASANLMT Model Bahasa Indonesia: VALIDASI MKENYAMANAN(PRESI DAN F-MKENYAMANAN)

Class	Precision			
	LMT-C1	LMT-C2	LMT-C3	LMT-C4
C1/C2/C3/C4	0,999	1	1	0,999
DDoS	1	1	1	1
Class	F-measure (F1 score)			
	LMT-C1	LMT-C2	LMT-C3	LMT-C4
C1/C2/C3/C4	0,999	1	1	1
DDoS	0,999	1	1	1

pada hasil positif dan negatif. Mengingat bahwa setiap ukuran memiliki kelebihan dan kekurangan, keberhasilan model klasifikasi berdasarkan pembelajaran mesin harus disampaikan melalui banyak metode berturut-turut.

Pengukuran pertama adalah tingkat TPR. TPR mencerminkan contoh-contoh yang dikategorikan dengan benar dari suatu kelas dalam kumpulan semua contoh yang dikaitkan dengan kelas tersebut.

$$TPR = \frac{T.P.}{TP + FN} \quad (16)$$

Dalam persamaan di atas, TPR adalah tingkat positif sebenarnya.

Tabel IX menunjukkan hasil TPR untuk semua versi model LMT dengan TPR untuk semua kelas lalu lintas yang sah adalah 1. Nilai TPR untuk kelas DDoS dalam model LMT-C2 dan LMT-C3 adalah 1. Model LMT-C1 dan LMT-C4 mencatat penurunan kinerja minimal dengan TPR sebesar 0,999. Ukuran evaluasi kinerja penting berikutnya adalah rasio contoh FP (FPR) yang ditunjukkan dalam tabel yang sama.

Tingkat FP merupakan rasio contoh kelas yang salah diklasifikasikan dalam kumpulan semua contoh yang ditetapkan ke kelas tersebut terhadap (17). Berdasarkan pengukuran ini, semua model menunjukkan hasil yang baik untuk kelas lalu lintas yang sah dan kelas DDoS.

$$FPR = \frac{\text{Baixa Ingris}}{FP + TN} \quad (17)$$

Dalam persamaan di atas, FPR adalah rasio positif palsu.

Menurut (18), perhitungan presisi digunakan untuk menyatakan jumlah contoh yang dikategorikan dengan benar dalam kaitannya dengan jumlah total contoh yang termasuk dalam kelas itu.

Menurut (19), *F*-mengukur atau *Bahasa Indonesia: F1*skor mewakili rata-rata harmonik dari presisi dan TPR [30]. Rata-rata harmonik lebih intuitif daripada rata-rata aritmatika klasik untuk menghitung rata-rata rasio, menurut [31]

$$PPV = \frac{T.P.}{TP + FP} \quad (18)$$

TABEL XI

HAIULASANLMT Model Bahasa Indonesia: VALIDASI MKENYAMANAN(ROC DAN(RRT))

Class	ROC			
	LMT-C1	LMT-C2	LMT-C3	LMT-C4
C1/C2/C3/C4	0,999	1	1	1
DDoS	0,999	1	1	1
Class	PRC			
	LMT-C1	LMT-C2	LMT-C3	LMT-C4
C1/C2/C3/C4	0,998	1	1	1
DDoS	0,999	1	1	1

Pada persamaan di atas, PPV adalah nilai prediksi positif

$$F1 = \frac{2 \cdot PPV_{\text{Bahasa Indonesia: TPR}}}{PPV + TPR}. \quad (19)$$

Berdasarkan nilai yang ditunjukkan pada Tabel X, kedua ukuran menunjukkan kinerja tinggi dari semua versi model LMT. Penurunan kinerja minimal diamati untuk LMT-C1 dan LMT-C4 (0,999) untuk kelas C1 dan C4 untuk ukuran presisi dan untuk LMT-C1 untuk kelas C1 dan DDoS untuk peringkat F1 (0,999).

Keempat versi model LMT memiliki kinerja tinggi yang dapat dilihat dari pengukuran ROC dan PRC yang diterapkan, yang hasilnya dapat dilihat pada Tabel XI. Sebagai salah satu pengukuran terpenting dan paling sering digunakan yang menunjukkan kualitas model klasifikasi, hasil pengukuran ROC menunjukkan kualitas tinggi dari semua versi model LMT yang dikembangkan. Buktiannya adalah nilai rasio rasio TPR dan TNR, yaitu 1 untuk model LMT-C2, LTM-C3, dan LMT-C4, dan 0,999 untuk model LMT-C1.

Karena kumpulan data tersebut terstratifikasi, maka pengukuran PRC sebagai alternatif terhadap pengukuran ROC, yang dapat menilai dengan lebih baik dampak sejumlah besar contoh negatif pada kinerja model, memberikan nilai yang hampir sama untuk semua model LMT yang diamati.

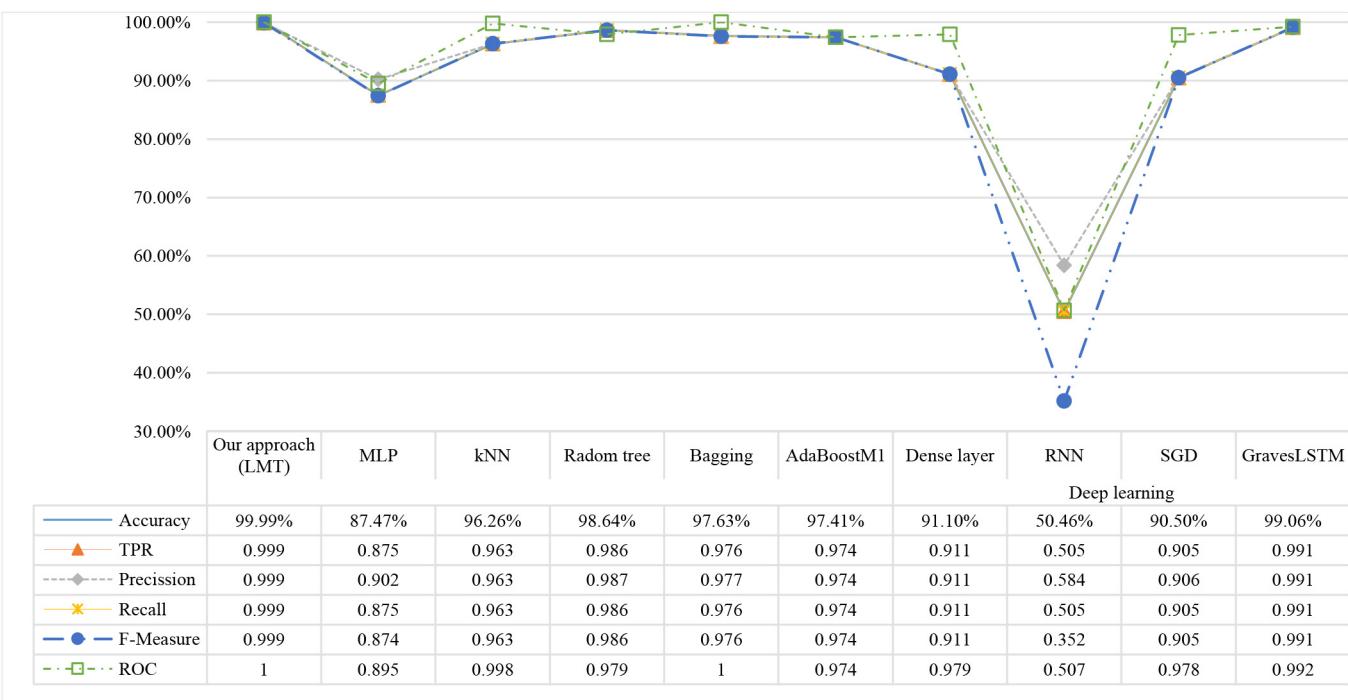
Hasil yang disajikan dari model yang dikembangkan mengonfirmasi hipotesis kedua dari penelitian ini. Berdasarkan profil lalu lintas sah yang ditetapkan dari kelas perangkat IoT tertentu di lingkungan rumah pintar, mendeteksi lalu lintas tidak sah yang dihasilkan oleh perangkat tersebut adalah mungkin.

Ringkasan perbandingan ditunjukkan pada Tabel XII, dan dapat diamati bahwa pendekatan kami mencapai akurasi, presisi, ingatan, dan *F*-ukuran. Kami juga mengamati bahwa kami mempertimbangkan jumlah perangkat SHIoT tertinggi, dan manfaat utama pendekatan kami adalah kemampuannya untuk mendeteksi lalu lintas anomali yang dihasilkan oleh perangkat IoT yang sebelumnya tidak terlihat.

Singkatnya, untuk mengevaluasi efektivitas metode LMT yang diterapkan dalam penelitian ini, kami menerapkan beberapa metode pembelajaran mesin yang sering digunakan pada set data yang sama. Secara khusus, kami membandingkan kinerja proposal kami dengan multilayer perceptron (MLP), *k*-Nearest Neighbors (*k*NN), Random Tree (RT), Bagging, AdaBoostM1, stochastic gradient descent (SGD), lapisan padat, Recurrent Neural Network (RNN), dan GravesLSTM, dalam hal akurasi, TPR, Presisi, Recall, *F*-ukuran, dan ROC. Untuk menerapkan metode yang disebutkan, kami menggunakan paket WekaDeepLearning4j untuk platform WEKA [32]. Dari hasil perbandingan yang disajikan pada Gambar 6, orang dapat melihat bahwa pendekatan kami secara umum mengungguli metode terapan lainnya.

TABEL XII
KOMPARASIT
KamiTHHAIDI SANACPERSAINGANAPENDEKATAN

Research	Used method	Dataset	Acuracy	Precision	Recall	F1 - Measure	Requires profiling each individual device to detect anomaly
[16]	LSTM	Modbus network traffic data	0.9962	0.9935	0.9941	0.993	n/a
[18]	RNN, ResNet, ConvNet	7 real IoT devices	0.852 - 0.9997	0.6248 - 0.9997	0.6740 - 0.9997	0.6208 - 0.9997	no
[19]	J48	8 real IoT devices	n/a	0.9	0.899	0.888	yes
[42]	Convolutional neural networks (CNN)	CICDDoS2019	n/a	0.87	0.86	0.86	n/a
[43]	shingling-based graph sketching	28 real IoT devices	n/a	0.98	0.92	0.92	yes
[8]	KN, LSVM, DT, RF, NN	3 real IoT devices	0.991 - 0.999	0.983 - 0.999	0.870 - 0.999	0.927 - 0.999	yes
[44]	LGBM, DNN, SVM	8 real IoT devices / 5 Non IoT devices	n/a	n/a	0.370 - 1	n/a	yes
Our approach	Bosting based LMT	41 real IoT devices (555,508,689 packets, 393.33 GB of traffic, 3,458.47 hours)	0.9921 - 0.9996	0.999 - 1	0.999 - 1	0.999 - 1	no (works with previously unseen devices)



Gbr. 6. Perbandingan pendekatan kami dengan metode pembelajaran mesin dan pembelajaran mendalam pesaing lainnya.

V.C. Bahasa Indonesia KESIMPULAN DAN FUTUR Kami Bahasa Inggris

Model deteksi DDoS yang disajikan dalam artikel ini menyimpang dari pendekatan deteksi anomali lalu lintas jaringan yang umum. Misalnya, pendekatan sebelumnya sebagian besar didasarkan pada pembuatan profil lalu lintas yang sah yang diasumsikan berlaku untuk semua perangkat terminal. Pendekatan semacam itu logis dalam lingkungan yang terdiri dari perangkat konvensional, yang lalu lintasnya menghasilkan karakteristik yang mencerminkan pengoperasian aplikasi yang terpasang pada perangkat dan cara pengguna menggunakan perangkat tersebut.

Namun, perangkat IoT yang murah agak terbatas dalam hal fungsinya, yang tercermin dalam karakteristik lalu lintas yang dihasilkannya. Ada juga IoT

perangkat yang lebih mampu secara komputasi. Oleh karena itu, pendekatan non-IoT yang ada mungkin tidak cocok, sebagian karena keragaman perangkat IoT (dan akibatnya, perilaku). Dengan kata lain, beberapa perangkat akan selalu menghasilkan lalu lintas yang sama, sementara perangkat lain yang mampu mendukung interaksi yang lebih besar dengan pengguna dapat menghasilkan lalu lintas yang tidak teratur. Yang memperparah tantangan ini adalah pertumbuhan signifikan dalam jumlah perangkat di lingkungan IoT.

Dengan kata lain, pendekatan deteksi DDoS berdasarkan karakteristik perangkat individual memerlukan pembelajaran ulang atau bahkan pengembangan ulang model dasar untuk setiap perangkat baru yang muncul di pasaran. Pendekatan semacam itu sangat rumit dan tidak cukup generik dalam lingkungan yang semakin kompleks dan

lingkungan IoT yang dinamis. Ini adalah batasan yang kami coba atasi dalam artikel ini.

Pendekatan kami mengasumsikan bahwa tidak ada satu profil lalu lintas sah yang menyeluruh untuk perangkat IoT, dan alih-alih berfokus pada perangkat tertentu, kami berfokus pada kelas perangkat (tergantung pada karakteristik lalu lintas yang dihasilkannya). Dengan cara ini, profil lalu lintas sah dibentuk untuk setiap kelas perangkat berdasarkan model deteksi DDoS yang dikembangkan. Pendekatan ini berpotensi untuk mengklasifikasikan perangkat masa depan berdasarkan karakteristik aliran lalu lintas yang dihasilkannya, yang dapat digunakan untuk menentukan apakah perangkat tersebut berperilaku dalam batasan yang sah atau menghasilkan lalu lintas DDoS. Secara khusus, dalam pendekatan kami, model deteksi lalu lintas DDoS didasarkan pada metode pohon keputusan logistik dari serangkaian metode pembelajaran mesin. Masalah mendeteksi lalu lintas DDoS berdasarkan kelas perangkat telah direduksi menjadi klasifikasi biner, di mana versi berbeda dari model yang sama dikembangkan untuk setiap kelas perangkat SHIoT. Inilah sebabnya mengapa setiap kelas lalu lintas perangkat SHIoT memiliki karakteristik yang berbeda, yang terbukti dari versi model yang disajikan, masing-masing berbeda dalam jumlah fitur independen yang digunakan, ukuran pohon keputusan, dan nilai ambang batas percabangannya. Evaluasi kinerja kami menunjukkan bahwa pendekatan tersebut mencapai kinerja tinggi, dalam hal akurasi, TPR, FPR, peringkat F1, presisi, ROC, dan PRC. Misalnya, akurasi model untuk masing-masing kelas adalah C1 = 99,9216%, C2 = 99,9966%, C3 = 99,9744%, dan C4 = 99,9583%.

Pendekatan kami dapat menguntungkan berbagai pemangku kepentingan dalam ekosistem IoT. Misalnya, pengguna biasanya ingin perangkat mereka berfungsi sebagaimana mestinya di lingkungan rumah pintar. Menghasilkan lalu lintas DDoS dapat memengaruhi fungsionalitas perangkat atau membuatnya sama sekali tidak dapat diakses. Oleh karena itu, pengguna berkepentingan untuk segera mendeteksi perilaku abnormal perangkat. Mengingat bahwa operator telekomunikasi sering kali juga merupakan penyedia layanan rumah pintar, mereka juga berkepentingan untuk mendeteksi perilaku perangkat yang tidak sah secara tepat waktu guna melindungi infrastruktur jaringan mereka sendiri. Terakhir, produsen perangkat tersebut harus memastikan pengoperasian perangkat yang benar guna meningkatkan kepuasan pelanggan dan memperkuat kehadiran pasar mereka.

Meskipun penelitian kami telah menunjukkan potensi mendeteksi lalu lintas ilegal dengan akurasi tinggi berdasarkan klasifikasi perangkat ke dalam kelas yang telah ditentukan sebelumnya dan membuat profil lalu lintas yang sah untuk setiap kelas menggunakan metode boosting machine learning, ada sejumlah kemungkinan perluasan di masa mendatang untuk pekerjaan ini. Misalnya, kami bermaksud untuk mengevaluasi pendekatan yang kami usulkan dalam pengaturan lain, seperti perawatan kesehatan, transportasi, atau Industri 4.0, karena perangkat dalam domain aplikasi ini dapat menghasilkan perilaku yang berbeda dan karenanya menghasilkan kelas perangkat tambahan. Kami juga bermaksud untuk mempelajari potensi perluasan pendekatan kami untuk mencakup jenis serangan lain, misalnya untuk membuat kelas perangkat berdasarkan lalu lintas yang dihasilkannya di hadapan jenis serangan lain.

RReferensi

- [1] I. Cvitić, D. Peraković, M. Periša, dan S. Husnjak, "Tinjauan umum pendekatan deteksi lalu lintas penolakan layanan terdistribusi," *Transportasi Lalu Lintas PROMET*, jilid. 31, tidak. 4, hlm. 453–464, Agustus 2019, doi:[10.7307/ptt.v31i4.3082](https://doi.org/10.7307/ptt.v31i4.3082).
- [2] GA Jaafar, SM Abdullah, dan S. Ismail, "Tinjauan metode deteksi terbaru untuk serangan HTTP DDoS," *J. Komputer, Jaringan, Komunikasi*, jilid. 2019, hlm. 1–10, Januari 2019, doi:[Nomor telepon 10.1155/2019/1283472](https://doi.org/10.1155/2019/1283472).
- [3] Ancaman AWS Shield. (2020). *Laporan Lanskap Ancaman—Q 1 2020 AWS* Diakses: 29 Oktober 2020. [Online]. Tersedia: https://aws-shieldtr.s3.amazonaws.com/2020-Q1_AWS_Shield_TLR.pdf
- [4] I. Cvitić, M. Vujić, dan S. Husnjak, "Klasifikasi risiko keamanan di lingkungan IoT," dalam *Proses. Ann. DAAAM Int. Gejala DAAAM. Intel. Manuf. Otomatis.*, 2016, hlm. 0731–0740, doi:[10.2507/proses-hari-ke-26.102](https://doi.org/10.2507/proses-hari-ke-26.102).
- [5] I. Cvitić, D. Peraković, M. Periša, dan M. Botica, "Karakteristik lalu lintas IoT rumah pintar sebagai dasar deteksi lalu lintas DDoS," dalam *Prosiding Konferensi Int. EAI ke-3. Sistem Manufaktur Manajemen*, 2018, hlm. 1–10, doi:[Nomor telepon 10.4108/eai.6-11-2018.2279336](https://doi.org/10.4108/eai.6-11-2018.2279336).
- [6] I. Cvitić, D. Peraković, M. Periša, dan B. Gupta, "Pendekatan pembelajaran mesin ensemble untuk klasifikasi perangkat IoT di rumah pintar," *Int. J. Mach. Belajar. Cybern.*, akan diterbitkan, doi: [Nomor 10.1007/s13042-020-01241-0](https://doi.org/10.1007/s13042-020-01241-0).
- [7] PMS Sánchez, JM Valero, AH Celrá, G. Bovet, MG Pérez, dan GM Pérez, "Survei tentang sidik jari perilaku perangkat: Sumber data, teknik, skenario aplikasi, dan kumpulan data," Agustus 2020. [Online]. Tersedia: <http://arxiv.org/abs/2008.03343>.
- [8] R. Doshi, N. Aphorpe, dan N. Feamster, "Deteksi DDoS pembelajaran mesin untuk perangkat Internet of Things konsumen," dalam *Prosiding Lokakarya Keamanan Privasi IEEE (SPW)*, San Francisco, CA, AS, Mei 2018, hlm. 29–35, doi:[Nomor 10.1109/SPW.2018.00013](https://doi.org/10.1109/SPW.2018.00013).
- [9] Y. Meidandari lain-lain., "N-BaIoT—Deteksi serangan IoT Botnet berbasis jaringan menggunakan autoencoder dalam," *Komputasi Pervasif IEEE*, vol. 17, no. 3, hlm. 12–22, Juli–September 2018.
- [10] M. Özcelik, N. Chalabianloo, dan G. Gür, "Pertahanan tepi yang ditentukan perangkat lunak terhadap DDoS berbasis IoT," dalam *Prosiding Konferensi Internasional IEEE ke-17. Komputasi Inf. Teknologi (CIT)*, Helsinki, Finlandia, 2017, hlm. 308–313, doi: [Nomor 10.1109/CIT.2017.61](https://doi.org/10.1109/CIT.2017.61).
- [11] I. Cvitić, D. Peraković, M. Periša, dan M. Botica, "Pendekatan baru untuk mendeteksi lalu lintas DDoS yang dihasilkan IoT," *Jaringan Nirkabel*, jilid. 27, hlm. 1573–1586, Juni 2019, doi:[Nomor 10.1007/s11276-019-02043-1](https://doi.org/10.1007/s11276-019-02043-1).
- [12] R. Hallman, J. Bryan, G. Palavicini, J. Divita, dan J. Romero-Mariona, "IoDDoS—Serangan penolakan layanan terdistribusi di Internet—Studi kasus malware Mirai dan botnet berbasis IoT," dalam *Prosiding Konferensi Internasional ke-2. Internet Hal-hal Keamanan Data Besar*, 2017, hlm. 47–58, doi:[Nomor telepon 10.5220/0006246600470058](https://doi.org/10.5220/0006246600470058).
- [13] DH Summerville, KJ Zach, dan Y. Chen, "Deteksi anomali paket dalam yang sangat ringan untuk perangkat Internet of Things," dalam *Prosiding Konferensi Komunikasi Komputasi Performa Int. IEEE ke-34 (IPCCC)*, Nanjing, Tiongkok, Desember 2015, hlm. 1–8, doi:[Nomor 10.1109/IPCCC.2015.7410342](https://doi.org/10.1109/IPCCC.2015.7410342).
- [14] A. Saied, RE Overill, dan T. Radzik, "Deteksi serangan DDoS yang diketahui dan tidak diketahui menggunakan jaringan saraf buatan," *Komputasi saraf*, jilid. 172, hlm. 385–393, Januari 2016, doi:[10.1016/j.neucom.2015.04.101](https://doi.org/10.1016/j.neucom.2015.04.101).
- [15] R. Vishwakarma dan AK Jain, "Survei teknik serangan DDoS dan mekanisme pertahanan di jaringan IoT," *Sistem Telekomunikasi*, jilid. 73, tidak. 1, hlm. 3–25, 2020, doi:[Nomor telepon 10.1007/s11235-019-00599-z](https://doi.org/10.1007/s11235-019-00599-z).
- [16] M. Saharkhizan, A. Azmoodhe, A. Dehghantanha, K.-KR Choo, dan RM Parizi, "Sebuah ensemble jaringan saraf berulang dalam untuk mendeteksi serangan siber IoT menggunakan lalu lintas jaringan," *Jurnal IEEE tentang Internet Things.*, vol. 7, no. 9, hlm. 8852–8859, September 2020, doi:[10.1109/jiot.2020.2996425](https://doi.org/10.1109/jiot.2020.2996425).
- [17] I. Ullah dan QH Mahmoud, "Sistem deteksi aktivitas anomal berbasis aliran dua tingkat untuk jaringan IoT," *Elektronik*, jilid. 9, tidak. 3, hal. 530, Maret 2020, doi:[10.3390/elektronik9030530](https://doi.org/10.3390/elektronik9030530).
- [18] O. Salman, IH Elhajj, A. Chehab, dan A. Kayssi, "Kerangka kerja berbasis pembelajaran mesin untuk identifikasi perangkat IoT dan deteksi lalu lintas abnormal," *Trans. Telekomunikasi. Teknologi. Darurat.*, akan diterbitkan, doi: [10.1002/ett.3743](https://doi.org/10.1002/ett.3743).
- [19] E. Anthi, L. Williams, M. Slowińska, G. Theodorakopoulos, dan P. Burnap, "Sistem deteksi intrusi yang diawasi untuk perangkat IoT rumah pintar," *Jurnal IEEE tentang Internet Things.*, vol. 6, no. 5, hlm. 9042–9053, Oktober 2019, doi:[Nomor 10.1109/JIOT.2019.2926365](https://doi.org/10.1109/JIOT.2019.2926365).
- [20] D. Peraković, M. Periša, dan I. Cvitić, "Analisis dampak IoT terhadap volume serangan DDoS," dalam *Proses. Simposium ke-33 atau Novim Tehnologijama di posTel dan Telekomunikasi Saobraćaju (PosTel)*, 2015, hlm. 295–304.
- [21] N. Vlajic dan D. Zhou, "IoT sebagai lahan peluang bagi peretas DDoS," *Komputer*, jilid. 51, tidak. 7, hlm. 26–34, Juli 2018, doi: [Nomor 10.1109/MC.2018.3011046](https://doi.org/10.1109/MC.2018.3011046).

- [22] N. Chaabouni, M. Mosbah, A. Zemmari, C. Sauvignac, dan P. Faruki, "Deteksi intrusi jaringan untuk keamanan IoT berdasarkan teknik pembelajaran," *Survei Komunitas IEEE Tuts.*, vol. 21, no. 3, hlm. 2671–2701, Triwulan ke-3, 2019, doi:[Nomor telepon 10.1109/COMST.2019.2896380](https://doi.org/10.1109/COMST.2019.2896380).
- [23] A. Sivanathan dan lain-lain, "Mengklasifikasikan perangkat IoT di lingkungan pintar menggunakan karakteristik lalu lintas jaringan," *IEEE Trans. Komputer Seluler.*, jilid. 18, tidak. 8, hlm. 1745–1759, Agustus 2019, doi:[Nomor telepon 10.1109/TMC.2018.2866249](https://doi.org/10.1109/TMC.2018.2866249).
- [24] I. Cvitic, D. Perakovic, M. Perisa, dan M. Botica, *Definisi Kelas Perangkat IoT Berdasarkan Fitur Aliran Lalu Lintas Jaringan* (Inovasi EAI/Springer dalam Komunikasi dan Komputasi), L. Knapcikova, M. Balog, D. Perakovic, dan M. Perisa, Ed. Cham, Swiss: Springer, 2020, hlm. 1–17.
- [25] GitHub—Markus-Go/Bonesi: BoNeSi—Simulator Botnet DDoS Diakses: 7 Agustus 2019. [Online]. Tersedia: <https://github.com/Markus-Go/bonesi>
- [26] AH Lashkari, GD Gil, MSI Mamun, dan AA Ghorbani, "Karakterisasi lalu lintas tor menggunakan fitur berbasis waktu," dalam *Prosiding Konferensi Int. ke-3 Sistem Inf. Keamanan Privasi (ICISSP)*, 2017, hlm. 253–262, doi: doi:[Nomor telepon 10.5220/0006105602530262](https://doi.org/10.5220/0006105602530262).
- [27] N. Landwehr, M. Hall, dan E. Frank, *Pohon Model Logistik* (Catatan Kuliah dalam Kecerdasan Buatan (Catatan Kuliah dalam Ilmu Komputer 2837). New York, NY, AS: Springer, 2003, hlm. 241–252.
- [28] N. Landwehr, M. Hall, dan E. Frank, "Pohon model logistik," *Mesin. Belajar.*, vol. 59, no. 1–2, hal. 161–205, 2005.
- [29] B. Hssina, A. Merbouha, H. Ezzikouri, dan M. Erritali, "Studi perbandingan pohon keputusan ID3 dan C4.5," *Jurnal Int. Ilmu Komputer Lanjutan Terapan*, jilid. 4, tidak. 2, hlm. 13–19, 2014, doi:[10.14569/edisi khusus.2014.040203](https://doi.org/10.14569/edisi_khusus.2014.040203).
- [30] M. Hossin dan M. Sulaiman, "Tinjauan metrik evaluasi untuk evaluasi klasifikasi data," *Jurnal Int. Data Min. Pengetahuan. Manajemen. Proses*, jilid. 5, tidak. 2, hlm. 01–11, Maret 2015, doi:[10.5121/jdkp.2015.5201](https://doi.org/10.5121/jdkp.2015.5201).
- [31] Y. Sasaki, "Kebenaran dari ukuran F," *Mengajar. Guru. Guru.*, vol. 1, no. 4, hal. 1–6, 2007.
- [32] S. Lang, F. Bravo-Marquez, C. Beckham, M. Hall, dan E. Frank, "WekaDeepLearning4j: Paket pembelajaran mendalam untuk Weka berdasarkan deepLearning4j," *Sistem Berbasis Pengetahuan*, jilid. 178, hlm. 48–50, Agustus 2019, doi:[10.1016/j.knosys.2019.04.013](https://doi.org/10.1016/j.knosys.2019.04.013).
- [33] AL Buczak dan E. Guven, "Survei metode penambangan data dan pembelajaran mesin untuk deteksi intrusi keamanan siber," *Survei Komunitas IEEE Tuts.*, vol. 18, no. 2, hlm. 1153–1176, Kuartal ke-2, 2016, doi:[Nomor 10.1109/COMST.2015.2494502](https://doi.org/10.1109/COMST.2015.2494502).
- [34] SJ Stolfo, W. Fan, W. Lee, A. Prodromidis, dan PK Chan, "Pemodelan berbasis biaya untuk deteksi penipuan dan intrusi: Hasil dari proyek JAM," dalam *Prosiding Pameran Konferensi Kemampuan Bertahan Hidup Inf. DARPA (DISCEX'00)*, vol. 2. Hilton Head, SC, AS, 2000, hlm. 130–144, doi:[10.1109/DISCEX.2000.821515](https://doi.org/10.1109/DISCEX.2000.821515).
- [35] Kumpulan Data CAIDA UCSD 'Serangan DDoS 2007', CAIDA, La Jolla, CA, AS, 2007.
- [36] M. Tavallaei, E. Bagheri, W. Lu, dan AA Ghorbani, "Analisis terperinci dari kumpulan data KDD CUP 99," di *Prosiding IEEE Symp. Comput. Intell. Pertahanan Keamanan Appl.*, Ottawa, ON, Kanada, Juli 2009, hlm. 1–6, doi: doi:[10.1109/CISDA.2009.5356528](https://doi.org/10.1109/CISDA.2009.5356528).
- [37] P. Gogoi, DK Bhattacharyya, B. Borah, dan JK Kalita, "MLH-IDS: Metode deteksi intrusi hybrid multi-level," *Komputasi J.*, jilid. 57, tidak. 4, hal. 602–623, 2014, doi:[10.1093/komite/bxt044](https://doi.org/10.1093/komite/bxt044).
- [38] HH Jazi, H. Gonzalez, N. Stakhanova, dan AA Ghorbani, "Mendeteksi serangan DoS lapisan aplikasi berbasis HTTP pada server Web dengan adanya pengambilan sampel," *Komputasi. Jaringan.*, jilid. 121, hlm. 25–36, Juli 2017, doi:[10.1016/j.comnet.2017.03.018](https://doi.org/10.1016/j.comnet.2017.03.018).
- [39] I. Sharafaldin, AH Lashkari, dan AA Ghorbani, "Menuju pembuatan dataset deteksi intrusi baru dan karakterisasi lalu lintas intrusi," di dalam *Prosiding Konferensi Int. ke-4 Sistem Inf. Keamanan Privasi*, 2018, hlm. 108–116, doi: doi:[10.5220/0006639801080116](https://doi.org/10.5220/0006639801080116).
- [40] N. Koroniotis, N. Moustafa, E. Sitnikova, dan B. Turnbull, "Menuju pengembangan dataset botnet realistik di Internet of Things untuk analisis forensik jaringan: Dataset Bot-IoT," *Sistem Komputer Umum Masa Depan*, jilid. 100, hlm. 779–796, November 2019, doi:[10.1016/j.masa.depan.2019.05.041](https://doi.org/10.1016/j.masa.depan.2019.05.041).
- [41] R. Paudel, T. Muncy, dan W. Eberle, "Mendeteksi serangan DoS di perangkat IoT rumah pintar menggunakan pendekatan berbasis grafik," dalam *Prosiding Konferensi Internasional IEEE Data Besar (Big Data)*, 2019, hlm. 5249–5258, doi: doi:[10.1109/DataBesar47090.2019.9006156](https://doi.org/10.1109/DataBesar47090.2019.9006156).
- [42] F. Hussain, SG Abbas, M. Husnain, UU Fayyaz, F. Shahzad, dan GA Shah, "Deteksi serangan IoT DoS dan DDoS menggunakan ResNet," dalam *Prosiding Konferensi Multitopik Internasional IEEE ke-23 (INMIC)*, Bahawalpur, Pakistan, November 2020, hlm. 1–6, doi:[10.1109/INMIC50486.2020.9318216](https://doi.org/10.1109/INMIC50486.2020.9318216).
- [43] R. Paudel, T. Muncy, dan W. Eberle, "Mendeteksi serangan DoS di perangkat IoT rumah pintar menggunakan pendekatan berbasis grafik," dalam *Prosiding Konferensi Internasional IEEE Data Besar (Big Data)*, Los Angeles, CA, AS, Desember 2019, hlm. 5249–5258, doi:[10.1109/DataBesar47090.2019.9006156](https://doi.org/10.1109/DataBesar47090.2019.9006156).
- [44] Y. Meidan, V. Sachidananda, H. Peng, R. Sagron, Y. Elovici, dan A. Shabtai, "Pendekatan baru untuk mendeteksi perangkat IoT rentan yang terhubung di belakang NAT rumah," *Keamanan Komputer*, vol. 97, Oktober 2020, Pasal no. 101968, doi:[10.1016/j.cose.2020.101968](https://doi.org/10.1016/j.cose.2020.101968).



Ivan Cvitić menerima gelar master dari Fakultas Ilmu Transportasi dan Lalu Lintas dan gelar Ph.D. di bidang ilmu teknik dari Universitas Zagreb, Zagreb, Kroasia, masing-masing pada tahun 2013 dan 2020.

Saat ini beliau bekerja di Fakultas Ilmu Transportasi dan Lalu Lintas, Universitas Zagreb, sebagai Peneliti Pascadoktoral dan Associate di Laboratorium Keamanan dan Analisis Forensik Sistem Informasi dan Komunikasi. Beliau telah menerbitkan lebih dari 50 makalah ilmiah di jurnal internasional.

Konferensi nasional, buku ilmiah, dan jurnal ilmiah berperingkat tinggi. Bidang penelitian dan minatnya adalah keamanan siber, pembelajaran mesin terapan dan kecerdasan buatan, pemodelan anomali lalu lintas jaringan, DDoS, Internet of Things, forensik digital, dan jaringan komunikasi.

Dr. Cvitić adalah anggota dewan editorial, dewan peninjau, dan editor tamu untuk beberapa jurnal ilmiah dan konferensi internasional berperingkat tinggi.



Dragan Perakovic (Anggota, IEEE) menerima gelar master dan Ph.D. di bidang ilmu teknis dari Fakultas Ilmu Transportasi dan Lalu Lintas (FPZ), Universitas Zagreb, Zagreb, Kroasia, masing-masing pada tahun 2003 dan 2005.

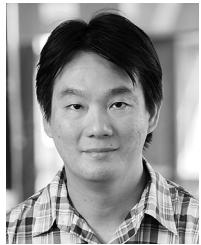
Setelah lulus, ia memulai karirnya di FPZ, di mana ia saat ini bekerja sebagai Profesor Penuh dan memegang posisi Kepala Departemen Lalu Lintas Informasi dan Komunikasi dan Kepala Departemen Sistem dan Layanan Informasi dan Komunikasi. Ia telah terlibat dalam beberapa berbagai proyek ilmiah internasional dan studi Penelitian dan Pengembangan sebagai peneliti, peneliti terkemuka, dan evaluator. Selain itu, ia telah menulis atau ikut menulis lebih dari 150 makalah ilmiah dan menjadi anggota, anggota dewan, dan editor resmi beberapa jurnal dan konferensi di bidang penelitiannya. Minat penelitiannya saat ini adalah di bidang keamanan, forensik digital, layanan komunikasi inovatif dalam sistem transportasi, kota pintar, dan industri 4.0.



Brij B. Gupta(Anggota Senior, IEEE) menerima gelar Ph.D. di bidang Informasi dan Keamanan Siber dari Institut Teknologi India Roorkee, Roorkee, India, pada tahun 2011.

Saat ini ia bekerja sebagai Asisten Profesor di Departemen Teknik Komputer, Institut Teknologi Nasional Kurukshetra, Kurukshetra, India. Ia juga bekerja sebagai Peneliti Utama di berbagai proyek Penelitian dan Pengembangan. Ia juga pernah menjadi Peneliti Tamu di Universitas Yamaguchi, Yamaguchi, Jepang, pada tahun 2015;

Deakin University, Geelong, VIC, Australia, pada tahun 2017; dan Swinburne University of Technology, Melbourne, VIC, Australia pada tahun 2018. Selain itu, ia menjadi Profesor Tamu di Temple University, Philadelphia, PA, AS, pada bulan Juni 2019, dan Staffordshire University, Stoke-on-Trent, Inggris, pada bulan Juli 2019. Ia menerbitkan lebih dari 300 makalah penelitian di Jurnal dan Konferensi Internasional yang memiliki reputasi tinggi. Minat penelitiannya meliputi keamanan informasi, keamanan siber, komputasi awan, keamanan Web, deteksi intrusi, dan Phishing.



Kim Kwang Raymond Choo(Anggota Senior, IEEE) menerima gelar Ph.D. dalam keamanan informasi dari Universitas Teknologi Queensland, Brisbane, QLD, Australia, pada tahun 2006.

Saat ini beliau memegang gelar Profesor Teknologi Awan dari Universitas Texas di San Antonio, San Antonio, TX, AS.

Prof. Choo adalah penerima Penghargaan Komite Teknis IEEE 2019 tentang Komputasi Skalabel untuk Keunggulan dalam Komputasi Skalabel (Peneliti Karier Menengah), Penghargaan UTSA College of

Kolonel Jean Piccione dan Letnan Kolonel Philip Piccione Menerima Penghargaan Penelitian Abadi untuk Dosen Tetap, Editor Asosiasi Berprestasi Tahun 2018 untuk IEEE ACCESS, Juara Kedua Wilkes Award dari British Computer Society tahun 2019, Penghargaan Highly Commended tahun 2014 oleh Badan Penasihat Kepolisian Australia Selanda Baru, Beasiswa Fulbright tahun 2009, Medali Prestasi Hari Australia tahun 2008, dan Penghargaan Wilkes dari British Computer Society tahun 2008. Ia juga menerima Penghargaan Makalah Terbaik dari IEEE Sistem JURNAL pada tahun 2021, Majalah Elektronik Konsumen IEEE untuk tahun 2020, Jurnal EURASIP tentang Komunikasi Nirkabel dan Jaringan pada tahun 2019, IEEE TrustCom 2018, dan ESORICS 2015; Masyarakat Pemrosesan Informasi Korea Jurnal Sistem Pemrosesan Informasi Penghargaan Riset Luar Biasa (Makalah Paling Banyak Dikutip) untuk tahun 2020 dan Penghargaan Makalah Survei (Emas) pada tahun 2019; Penghargaan Makalah Luar Biasa IEEE Blockchain 2019; dan Penghargaan Makalah Mahasiswa Terbaik dari Inscript 2019 dan ACISP 2005. Ia dinobatkan sebagai Pendidik Keamanan Siber Tahun Ini—APAC (Penghargaan Keunggulan Keamanan Siber diproduksi bekerja sama dengan Komunitas Keamanan Informasi di LinkedIn) pada tahun 2016, dan pada tahun 2015, ia dan timnya memenangkan Tantangan Riset Forensik Digital yang diselenggarakan oleh Universitas Erlangen-Nuremberg Jerman. Ia adalah Ketua Pendiri Komite Teknik Masyarakat Manajemen Teknologi dan Rekayasa IEEE tentang Teknologi Blockchain dan Buku Besar Terdistribusi, dan menjabat sebagai Editor Departemen IEEE Technology and Engineering Management Society, dan menjabat sebagai Editor Departemen IEEE Technology and Engineering Management Society. RANSAKSI TERHADAP Bahasa Inggris TEKNIK MMANAJEMEN, dan Editor Asosiasi IEEE TRANSAKSI TERHADAP DDAPAT DIHAPUS DAN SAMAN KOMPUTASI, dan IEEE TRANSAKSI TERHADAP BAKU eDATA Beliau merupakan ACM Distinguished Speaker dan IEEE Computer Society Distinguished Visitor dari tahun 2021 hingga 2023, dan termasuk dalam Peneliti yang Paling Banyak Dikutip Web of Science dalam bidang Lintas Bidang pada tahun 2020.