نشریه علمی پدافند غیرعامل

سال دهم، شماره ۴، زمستان ۱۳۹۸، (پیاپی ۴۰): صص ۱۰۱-۹۱

شناسایی کور شبکه‌های مخابراتی در لایه سرویس

مهدی تیموری*[۱]، حمیدرضا کاکایی مطلق[۲]، جواد گرشاسبی[۳]

تاریخ دریافت: ۱۳۹۷/۰۹/۲۸

تاریخ پذیرش: ۱۳۹۸/۰۵/۱۶

چکیده

یکی از مسائل مهم و اولیه در جنگ الکترونیک و سایبری، دسترسی غیرمجاز به حداکثر مقدار ممکن اطلاعات قابل حصول از یک شبکه مخابراتی است. در این راستا، اولین قدم در این راه، مهندسی معکوس لایه‌های فیزیکی و انتقال داده و در نتیجه، رسیدن به بسته‌های ارسالی در شبکه است. بعد از آن باید لایه سرویس مهندسی معکوس گردد. در این مرحله باید علاوه بر درک معنای هر یک از بسته‌های ارسالی، ماشین حالت مورد استفاده برای ارتباط در شبکه کشف گردد. با استفاده از این کار می‌توان به محتوای اطلاعات منتقل‌شده در شبکه دست یافت. در این مقاله، طرحی جامع برای شناسایی کور شبکه‌های مخابراتی در لایه سرویس ارائه می‌گردد. برای اثبات کارایی طرح ارائه‌شده، امکان استفاده از آن برای شناسایی کور نُه استاندارد مخابراتی مختلف مورد ارزیابی قرار می‌گیرد. در انتها نیز ضمن بررسی روش‌های موجود برای شناسایی لایه سرویس، مسائل حل‌نشده در این حوزه معرفی خواهد شد.

کلیدواژه‌ها: شناسایی کور، شبکه‌های مخابراتی، لایه‌سرویس

۱- استادیار، آزمایشگاه تئوری اطلاعات و کدینگ، دانشگاه تهران، (mehditeimouri@ut.ac.ir)- نویسنده مسئول

۲- دکتری، دانشگاه جامع امام حسین(ع)

۳- دانشجوی دکتری مهندسی کامپیوتر گرایش شبکه‌های کامپیوتری، آزمایشگاه تئوری اطلاعات و کدینگ، دانشگاه تهران

۱. مقدمه

شناسایی کور شبکه‌های مخابراتی، یکی از مسائل مهم در کاربردهای نظامی و امنیتی است. شناسایی کور شبکه به معنی مهندسی معکوس پروتکل ارتباطی شبکه با استفاده از مشاهده سیگنال‌های ارسالی در شبکه (و بدون هیچ اطلاعات جانبی دیگر) است. هدف نهایی چنین فرایندی، دست‌یابی به اطلاعات ردوبدل‌شده در شبکه می‌باشد [۱].

یکی از روش‌های مرسوم برای تعریف پروتکل شبکه‌های ارتباطی، مدل‌سازی لایه‌ای ارتباطات شبکه‌ای است. هر یک از استانداردهای مخابراتی ادبیات مختص به خود را برای این لایه‌بندی به کار برده‌اند که لزوماً مشابه استانداردهای دیگر نیست. با این حال، می‌توان سه لایه کلی زیر را برای یک شبکه مخابراتی نوعی تصور نمود: ۱- لایه سرویس[۱]، ۲- لایه انتقال داده[۲] و ۳- لایه فیزیکی[۳].

وظیفه لایه سرویس، مدیریت پیام‌های ارسالی و دریافتی در قالب ماشین حالت سرویس‌ها است. در یک شبکه مخابراتی، ممکن است سرویس‌های متنوعی ارائه گردد. سرویس ثبت‌نام در شبکه، سرویس تغییر موقعیت، سرویس تماس صوتی و سرویس ارسال پیام نمونه‌هایی از این سرویس‌ها هستند. در یک نگاه کلی، ارتباط شبکه‌ای بین گره‌ها نیز نوعی سرویس است که توسط یک شبکه ارائه می‌گردد. تبدیل پیام‌های لایه سرویس به محتوای باینری، بخش‌بندی و بسته‌بندی آن‌ها و افزودن سرایندهای[۴] لازم از وظایف این لایه است. بسته به نوع شبکه، بسته‌بندی پیام‌ها می‌تواند در یک ساختار چند سطحی صورت گیرد. به‌عنوان مثال، در شبکه GSM[۵] ساختاری سه سطحی (شامل پیام‌های لایه ۱، پیام‌های لایه ۲ و پیام‌های لایه ۳) برای این بسته‌بندی پیشنهاد شده است. با مقایسه عملکرد این لایه با عملکرد لایه‌های مدل هفت لایه‌ای OSI[۶] برای شبکه، می‌توان لایه سرویس را تجمیع لایه‌های سوم تا هفتم OSI (لایه‌های شبکه[۷]، انتقال[۸]، نشست[۹]، نمایش[۱۰] و کاربرد[۱۱]) دانست.

وظیفه لایه انتقال داده انجام اعمال مرتبط با کدگذاری کانال (یعنی کدگذاری تشخیص و تصحیح خطا، جایگردان[۱۲] و درهم‌ساز[۱۳]) بر روی بسته‌های دریافتی از لایه سرویس و تحویل آن‌ها به لایه فیزیکی (و بالعکس) است. مسائلی مانند کنترل دسترسی به رسانه (MAC[۱۴]) و همزمان‌سازی فریم نیز در این لایه انجام می‌شود. با مقایسه عملکرد این لایه با عملکرد لایه‌های مدل هفت لایه‌ای OSI برای شبکه، می‌توان لایه انتقال داده را معادل لایه دوم OSI (لایه پیوند داده[۱۵]) دانست.

لایه فیزیکی نیز مسئولیت استفاده از محیط انتقال فیزیکی جهت ارسال و دریافت بسته‌های لایه انتقال را به عهده دارد. محیط انتقال فیزیکی می‌تواند طیف الکترومغناطیس، فیبر نوری و غیره باشد. عملکرد این لایه معادل لایه فیزیکی در مدل هفت‌لایه‌ای OSI است.

پس دریافت سیگنال‌های ارسالی در یک شبکه مخابراتی، اولین مرحله از شناسایی کور پروتکل مخابراتی مورد استفاده در این شبکه، شناسایی لایه فیزیکی است. این موضوع در مقالات متعددی مورد بررسی قرار گرفته است [۵-۲]. در حالت کلی، سیگنال می‌تواند در یک یا چند بُعد از فضای چند بُعدی فرکانس-زمان-فضا-کد حضور داشته باشد. در اولین قدم، باید فضای فرکانسی (باندهای فرکانسی و فرکانس‌های مرکزی مورد استفاده) شناسایی گردد. در مرحله بعد، باید پیوسته بودن و یا گسسته بودن سیگنال در حوزه زمان مشخص گردد. در صورت گسسته بودن سیگنال در حوزه زمان، بخش‌های فعال ارسال سیگنال که آن‌ها را برست‌های سیگنال[۱۶] می‌نامیم، باید از بخش‌های سکوت تفکیک شوند. در ادامه باید تشخیص داده شود که آیا از چند آنتن فرستنده برای ارسال سیگنال استفاده شده است یا خیر؟ اگر از چند آنتن فرستنده استفاده شده باشد، باید مشخص گردد که کد فضا-زمان مورد استفاده احتمالی چیست. در ادامه باید مشخص نمود که آیا از روش طیف گسترده برای ارسال استفاده می‌شود یا خیر؟ در صورتی که از این روش استفاده می‌شود باید تمام کدهای گسترش‌دهنده مورد استفاده را شناسایی نموده و با استفاده از آن‌ها کانال‌های مختلف فیزیکی را در حوزه کد تفکیک نمود. بعد از یافتن سیگنال در فضای چند بُعدی فرکانس-زمان-فضا-کد، باید نوع مدولاسیون مورد استفاده و نرخ بیت آن شناسایی گردد.

[1] Service Layer
[2] Data Transmission Layer
[3] Physical Layer
[4] Headers
[5] Global System for Mobile
[6] Open Systems Interconnection
[7] Network
[8] Transport
[9] Session
[10] Presentation
[11] Application

[12] Interleaver
[13] Scrambler
[14] Medium Access Control
[15] Data Link
[16] Signal Bursts

پس از شناسایی کامل لایه فیزیکی و اجرای فرایند دمدولاسیون، به بیت‌های ارسالی در لایه انتقال داده می‌رسیم. لذا قدم بعدی، شناسایی کور لایه انتقال داده است. در این لایه باید ضمن شناسایی و تفکیک کانال‌های داده مختلف از یکدیگر، کدگذاری کانال مورد استفاده در هر یک از آنها (شامل کدگذاری تشخیص و تصحیح خطا، جایگردان و درهم‌ساز) را شناسایی نمود. این موضوع در تحقیقات زیادی مورد بررسی قرار گرفته است [۶-۱۰].

پس از شناسایی و تفکیک کانال‌های داده و اجرای عمل کدگشایی کانال بر روی هر یک از آنها، به بسته‌های ارسالی توسط لایه سرویس می‌رسیم. برای اجرای عملیات شناسایی در لایه سرویس، باید علاوه‌بر درک معنای هر یک از بسته‌های ارسالی، ماشین حالت مورد استفاده برای ارتباط در شبکه کشف گردد. در این زمینه تحقیقات متنوعی صورت پذیرفته است. با این حال، تا جایی که نویسندگان مقاله اطلاع دارند، طرحی جامع برای شناسایی کور این لایه ارائه نشده است. در حقیقت، هر یک از تحقیقات موجود به حل یک مسئله خاص در این حوزه پرداخته‌اند و هیچ‌کدام روشی جامع برای حل کامل مسئله شناسایی لایه سرویس ارائه نداده‌اند. در این مقاله، طرحی جامع برای شناسایی کور شبکه‌های مخابراتی در لایه سرویس ارائه می‌گردد. برای اثبات کارایی طرح ارائه‌شده، امکان استفاده از آن برای شناسایی کور نُه استاندارد مخابراتی مختلف مورد ارزیابی قرار می‌گیرد.

ساختار مقاله در ادامه به این شرح است. در بخش دوم، طرح جامع برای شناسایی کور لایه سرویس ارائه می‌گردد. همچنین توانایی طرح پیشنهادی برای شناسایی کور شبکه‌های مخابراتی مختلف مورد ارزیابی قرار می‌گیرد. در بخش سوم، تحقیقات موجود در شناسایی کور لایه سرویس و ارتباط آن‌ها با طرح جامع پیشنهادی مورد بررسی قرار می‌گیرد. در ادامه و در بخش چهارم، مسائل حل‌نشده در حوزه شناسایی کور لایه سرویس معرفی می‌گردند. در انتها و در بخش پنجم نیز، مقاله جمع‌بندی و نتیجه‌گیری خواهد شد.

۲. طرح جامع برای شناسایی کور لایه سرویس

از آنجا که هزاران پروتکل شناخته‌شده و ناشناخته مخابراتی وجود دارد، امکان مطالعه و بررسی همه آن‌ها وجود ندارد. در حقیقت اگر تمام پروتکل‌های مخابراتی شناخته شده باشند، عملاً شناسایی کور معنایی ندارد. زیرا در این حالت با سیگنال‌هایی طرف هستیم که توسط یکی از این سیستم‌های شناخته‌شده تولید شده است. در چنین حالتی، مسئله تغییر یافته و هدف آن تعیین نوع سیستم مخابراتی هدف از میان تعدادی سیستم مخابراتی مشخص است. این مسئله به نوعی یک دسته‌بندی چند کلاسی (هر کلاس یک پروتکل مخابراتی است) می‌باشد که در بسیاری از موارد به راحتی و تنها با توجه به فرکانس کاری قابل انجام است.

قواعد علمی شناخته‌شده‌ای برای طراحی سیستم‌های مخابراتی وجود دارد. این قوانین موجب محدودیت در طراحی ساختار پروتکل‌های مخابراتی می‌شود؛ چرا که رعایت نکردن این قواعد علمی باعث افت عملکرد و اختلال در خدمات شبکه می‌گردد. مثلاً قبل از ارسال داده بر روی کانال باید عملیاتی بر روی داده انجام شود که به گیرنده در کنترل خطا کمک نماید. با توجه به حوزه کاربرد پروتکل، طراحان پروتکل‌های مخابراتی برای رعایت این قاعده‌ی علمی می‌توانند روش‌های مختلفی را به کار ببندند. در برخی از پروتکل‌ها صرفاً به تشخیص خطا توجه می‌شود؛ برخی پروتکل‌ها نیز تشخیص و تصحیح خطا در گیرنده را در نظر می‌گیرند. به‌عنوان مثال دیگر از قواعد علمی، وقتی چندین کاربر در یک سیستم مخابراتی به‌صورت همزمان مشغول به فعالیت هستند، باید روشی برای تسهیم منابع شبکه پیش‌بینی شده باشد. برخی پروتکل‌های تسهیم در حوزه زمان و برخی تسهیم در حوزه فرکانس را در نظر می‌گیرند. پروتکل‌های دیگری نیز وجود دارند که از روش‌های متنوع‌تری مانند تسهیم در فضای کد استفاده می‌کنند.

قواعد مخابراتی و روش‌های پیاده‌سازی آنها محدود و معین هستند. آن چیزی که موجب تنوع بسیار بالا در پروتکل‌های مخابراتی می‌گردد، ترکیب (اینکه کدام قواعد مورد استفاده قرار گیرد)، ترتیب (تقدم و تأخر قواعد چگونه باشد) و تنظیم (مشخصات دقیق روش مورد استفاده) قواعد طراحی هر پروتکل است. مهم‌ترین ویژگی یک طرح شناسایی کور این است که مبتنی بر قواعد علمی شناخته‌شده (و نه ترتیب، ترکیب و تنظیم آنها) باشد.

با توجه به توضیحات فوق، نُه پروتکل مخابراتی مهم و شناخته‌شده را که نماینده طیف وسیعی از پروتکل‌ها و اغلب قواعد پیاده‌شده مخابراتی هستند انتخاب کرده‌ایم و مسیر طراحی جامع برای شناسایی کور را به شرح زیر طی نموده‌ایم:

۱. استانداردهای انتخاب‌شده به دقت مورد بررسی و مطالعه قرارگرفته‌اند. در این راستا تلاش شده است که شیوه پردازش و ارسال اطلاعات در این استانداردها به دقت استخراج گردد.

۲. در مرحله دوم، فرض کرده‌ایم که سیگنال تولیدی توسط هر یک از این استانداردها، به یک سامانه جامع شناسایی

کور وارد شده است. بدون آن‌که فرض مشخصی در خصوص ساختار این سامانه شناسایی کور داشته باشیم، کمترین تعداد قدم‌هایی را مشخص نموده‌ایم که این سامانه شناسایی باید برای رسیدن به محتوای اطلاعاتی طی نماید. به عبارت دیگر، به‌عنوان یک ناظر بیرونی و منصف مشخص کرده‌ایم که اگر سامانه شناسایی کور قرار باشد به محتوای اطلاعاتی ردوبدل‌شده برسد، کدام قدم‌های اساسی را باید طی نماید. بدیهی است که با توجه به ویژگی هر کدام از پروتکل‌های مورد بررسی، قدم‌های خاص مشخص‌شده صرفاً مختص به آن پروتکل خواهد بود. به عبارت دیگر، قدم‌هایی که باید برای شناسایی یک پروتکل طی شود، لزوماً مشابه قدم‌هایی که برای شناسایی پروتکل دیگر برداشته می‌شود نیست.

۳. در مرحله آخر، نگاه مستقیم خود را از روی پروتکل‌های مورد بررسی برداشته و صرفاً با ملاحظه قدم‌های تعریف‌شده در مرحله قبل، ساختاری جامع برای شناسایی کور این پروتکل‌های مخابراتی ارائه می‌شود که تعداد قدم‌های مشخص‌شده در این ساختار جامع بسیار کمتر از مجموع قدم‌هایی است که در مرحله قبل مشخص شده است. کمتر بودن قدم‌های ساختار جامع پیشنهادی نسبت به مجموع قدم‌های تعریف‌شده در مرحله قبل متضمن جلوگیری از بیش‌برازش[1] در مدل پیشنهادی است. این موضوع احتمال کار کردن این مدل پیشنهادی برای پروتکل‌های مخابراتی ناشناخته را بسیار بالا خواهد برد.

به‌منظور ارائه طرح جامع پیشنهادی، نه استاندارد مختلف مخابراتی و شناسایی کور آن‌ها در لایه انتقال داده مورد بررسی دقیق قرار گرفته است. استانداردهای مورد بررسی در چند دسته قابل بررسی هستند.

۱. استانداردهای مخابرات سلولی ارتباط رادیویی و دسترسی داده شامل GSM، LTE و HSPDA

۲. استانداردهای ماهواره‌ای شامل DVB-S2

۳. استانداردهای مربوط به پیجرها شامل FLEX و POCSAG

۴. استانداردهای مربوط به ارتباط‌های رادیویی شامل TETRA و DMR

۵. استانداردهای مربوط به دسترسی بی‌سیم به شبکه‌های داده شامل WiMAX

انتخاب این استانداردها دلایل متعددی دارد.

۱. دلیل اول این است که منابع متعددی برای تشریح آن‌ها در دسترس هستند. در حقیقت بسیاری از این استانداردها به صورت رسمی منتشر شده‌اند.

۲. دلیل دوم این است که انتخاب این استانداردها به‌نحوی انجام شده است که در نتیجه بررسی این استانداردها، با بیشتر روش‌های مخابراتی مرسوم مواجه شویم.

- بیشتر پروتکل‌های مورد استفاده مبتنی بر ارسال برست هستند.
- تنوع بالایی از کدگذاری‌های کانال ساده (مانند کانولوشنی و گولی) تا کدهای کانال مدرن (مانند LDPC و کدهای ضربی) را شامل می‌شوند.
- تنوع بالایی از داده‌های مختلف (صوت، پیام کوتاه، سیگنالینگ، چندرسانه‌ای و اینترنت) را شامل می‌شوند.
- روش استفاده از چند آنتن برای ارسال و دریافت در LTE استفاده شده است.
- روش استفاده از طیف گسترده دنباله مستقیم (DSSS[2]) در HSDPA استفاده شده است.
- روش پرش فرکانسی[3] در GSM استفاده شده است.
- روش FDMA در GSM استفاده شده است.
- روش TDMA در GSM، TETRA، DMR و ... استفاده شده است.

۳. دلیل آخر این است که جزئیات بسیاری از استانداردهای مخابراتی نظامی به طور مستقیم و یا غیر مستقیم از این استانداردها الهام گرفته شده است. حتی برخی از این استانداردها کاربرد نظامی هم دارند.

[1] Overfitting

[2] Direct Sequence Spread Spectrum
[3] Frequency Hopping

- دو استاندارد TETRA و DMR خود دارای کاربرد نظامی هستند.
- استاندارد ماهواره‌ای GMR-1 که توسط موبایل ماهواره‌ای ثریا استفاده شباهت‌های ساختاری بسیار زیادی به استاندارد GSM دارد.
- پروتکل مخابراتی مورد استفاده در ماهواره‌های ارتباطی WGS شباهت ساختاری زیادی به استاندارد DVB-S2 دارد.
- ترکیبی از روش‌های DSSS، TDMA و پرش فرکانسی مورد استفاده در استانداردهای انتخاب‌شده در استاندارد ارتباط نظامی Link-16 استفاده شده است.

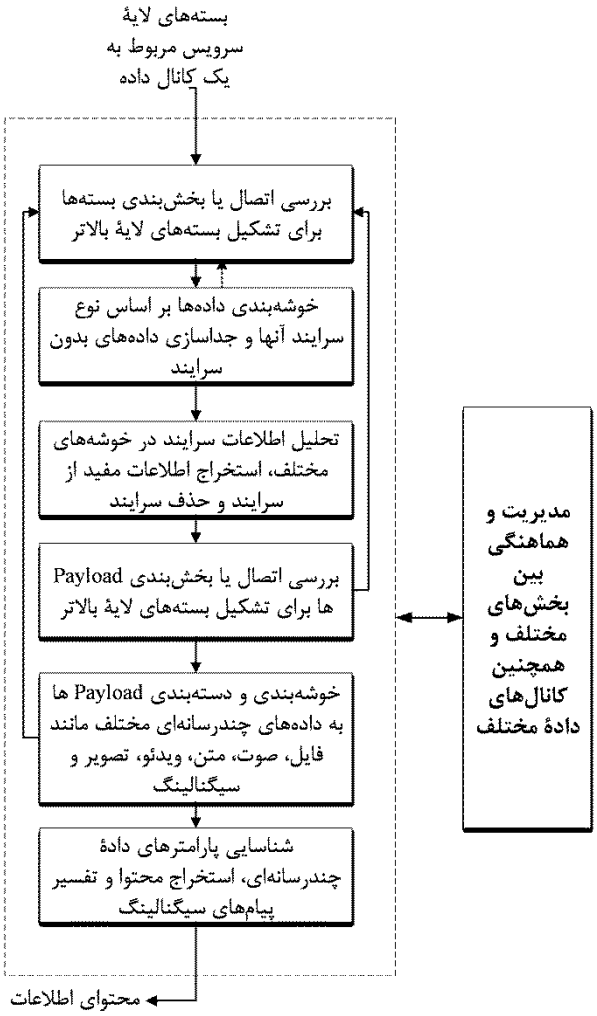
طرح کلی شناسایی در شکل (۱) نمایش داده شده است. این طرح دارای شش بلوک اصلی است:

۱. بررسی اتصال یا بخش‌بندی بسته‌ها برای تشکیل بسته‌های لایه بالاتر
۲. خوشه‌بندی داده‌ها بر اساس نوع سرایند آنها و جداسازی داده‌های بدون سرایند
۳. تحلیل اطلاعات سرایند در خوشه‌های مختلف، استخراج اطلاعات مفید از سرایند و حذف سرایند
۴. بررسی اتصال یا بخش‌بندی Payload ها برای تشکیل بسته‌های لایه بالاتر و برگشت به مرحله ۱
۵. خوشه‌بندی و دسته‌بندی Payload ها به داده‌های چندرسانه‌ای مختلف مانند فایل، صوت، متن، ویدئو، تصویر و سیگنالینگ و برگشت به مرحله ۱ برای داده‌های نوع سیگنالینگ
۶. شناسایی پارامترهای داده چندرسانه‌ای و استخراج محتوا

یک بخش مهم و کلیدی در این طرح مدیریت و هماهنگی بین بخش‌های مختلف است. این بخش علاوه‌بر هماهنگی بین سایر بخش‌ها، کمک می‌کند تا جریان اطلاعات کانال‌های داده مختلف در کنار یکدیگر مورد تحلیل و ارزیابی قرار گیرد. به‌عنوان مثال، یکی از انواع اطلاعاتی که می‌تواند برای این هدف مورد استفاده قرار گیرد زمان و ترتیب دریافت بسته‌ها در کانال‌های داده مختلف است. بخش مدیریت با دادن فرمان‌هایی به زیربخش‌های دیگر می‌تواند منجر به تحلیل‌های ارزشمندی شود. به عنوان مثال، در حالتی که سامانه مورد شنود چند کاربری است، با اعلام این موضوع به بخش تحلیل سرایند می‌توان آدرس MAC کاربران فعال را شناسایی نمود.

فرض‌های مهم این طرح کلی برای شناسایی به شرح زیر است:

- شناسایی در لایه زیرین شبکه به‌طور کاملاً دقیق انجام شده و داده ورودی این سامانه، بدون خطا و کامل است.
- ورودی سامانه بسته‌های لایه سرویس مربوط به کانال‌های داده (منطقی) مختلف است.
- خروجی سامانه محتوای اطلاعاتی مربوط به شبکه است. این اطلاعات می‌تواند شامل ساختار شبکه، تعداد و مشخصات کاربران و اطلاعات ردوبدل‌شده باشد.



شکل (۱): طرح جامع برای شناسایی کور لایه سرویس

قبل از توصیف بخش‌های مختلف، لازم است که اصطلاح‌های مورد نیاز ارائه گردند. لازم به ذکر است که برخی از این اصطلاح‌ها ممکن است در استانداردهای مختلف دارای تعریف‌های خاص‌تری باشند. با این حال، در این بخش، این اصطلاح‌ها تعمیم یافته و به صورت عمومی‌تری مورد استفاده قرار می‌گیرند.

- **داده چندبعدی:** همان‌طور که اشاره شد، اطلاعات می‌تواند در ابعاد مختلف فرکانس، زمان، فضا و کد ارسال گردد. چنین داده‌ای را (که می‌تواند از جنس سیگنال یا بیت باشد) یک داده چند بُعدی می‌نامیم. این داده

می‌تواند در قالب یک ماتریس چند بُعدی نمایش داده شود که بُعد اول آن فرکانس، بُعد دوم آن زمان، بُعد سوم آن کد و بُعد چهارم آن فضا باشد. با چنین تعریفی، عناصر یک سطر مشخص، داده‌هایی هستند که در یک فرکانس، کد و فضای مشخص و در زمان‌های مختلف ارسال می‌گردند. این داده می‌تواند پیوسته یا گسسته باشد. در حالتی که داده گسسته است، به هر بخش آن یک برست گفته می‌شود. زمانی که داده در طول زمان پیوسته باشد، به آن جریان (سیگنال یا بیت) پیوسته گفته می‌شود.

- **برست سیگنال:** یک رشته سیگنال که زمان شروع و پایان ارسال آن مشخص است، دارای هویتی مجزا نسبت به سایر سیگنال‌ها است و همچنین تمام آن در دسترس است.
- **برست بیتی:** یک رشته بیت که زمان شروع و پایان ارسال آن مشخص است، دارای هویتی مجزا نسبت به سایر بیت‌ها است و تمام آن در دسترس است. بدیهی است که سیگنال متناظر با این برست می‌تواند پیوسته باشد. همچنین یک برست سیگنال ممکن است شامل یک یا چند برست بیتی باشد.
- **کانال داده:** یک کانال داده زیرمجموعه‌ای از یک داده بیت چند بعدی است که به‌صورت مجزا از سایر داده‌ها بر روی آن کدگذاری کانال مشخصی (شامل مجموعه‌ای از کدهای تشخیص و تصحیح خطا، جایگردان‌ها و درهم‌سازها) انجام شده است. در برخی از استانداردها مانند GSM، به کانال داده، کانال منطقی گفته می‌شود. البته باید توجه داشت که در چنین استانداردهایی، ممکن است کدگذاری دو کانال منطقی مختلف یکسان باشد. تحت چنین شرایطی، یک گیرنده کور، مجموعه چنین کانال‌های منطقی را به‌عنوان یک کانال داده تشخیص خواهد داد.
- **بسته‌های اطلاعات:** با کدگشایی یک کانال داده، به بسته‌هایی می‌رسیم که آن‌ها را با عنوان کلی بسته‌های اطلاعات نام‌گذاری می‌کنیم. یک بسته اطلاعات در هر لایه شبکه، مجموعه‌ای از بیت‌ها است که دارای هویتی مجزا از سایر بسته‌ها است. این بسته می‌تواند از دو بخش سرایند و Payload تشکیل شده باشد.
- **Payload:** بخشی از بسته اطلاعاتی که هدف اصلی از ارتباط، ارسال آن به لایه بعدی است.
- **سرایند:** سرایند بخشی مشخص از ابتدای یک بسته اطلاعات است که طول آن مشخص است و محتوای اطلاعاتی آن صرفاً برای تحلیل و تفسیر صحیح Payloadها و ارتباط بین بسته‌ها مورد استفاده قرار می‌گیرد.

۲-۱. خوشه‌بندی داده‌ها بر اساس نوع سرایند آن‌ها و جداسازی داده‌های بدون سرایند

داده‌هایی که برای تحلیل وارد بخش تحلیل لایه سرویس می‌شود باید بر اساس نوع سرایند از یکدیگر تفکیک شوند. به‌عنوان مثال، در استاندارد GSM باید ابتدا اطلاعات لایه ۱ دارای سرایند از اطلاعات لایه ۱ بدون سرایند تفکیک شود. اطلاعات لایه ۱ دارای سرایند، اطلاعات کانال منطقی SACCH[1] هستند. سایر کانال‌ها مانند SDCCH[2]، FACCH[3] و TCH[4] نیز اطلاعات لایه ۱ بدون سرایند دارند. با این حال، اطلاعات لایه ۱ کانال‌های SDCCH و FAACH دارای سرایندهای لایه ۲ و ۳ هستند و اطاعات کانال TCH داده صوت خروجی کدگذار صوت می‌باشند. لذا انتظار می‌رود داده‌های لایه سرویس GSM به هفت خوشه ۱- داده‌های BCCH و CCCH، ۲- داده‌های SACCH نوع B4، ۳- داده‌های SACCH نوع A و B، ۴- داده‌های SACCH نوع Bter، ۵- داده‌های SDCCH و FACCH نوع A و B، ۶- داده‌های SDCCH و FACCH نوع Bter و ۷- داده‌های TCH بخش‌بندی شود. همچنین ممکن است در انتهای برخی پیام‌ها بیت‌های پرکننده مخصوص[5] وجود داشته باشد که اطلاعاتی در خود ندارد و باید حذف گردد.

در استاندارد FLEX همه داده‌ها در یک خوشه قرار می‌گیرند. البته ممکن است در انتهای برخی پیام‌ها بخشی به نام idle (بدون محتوای اطلاعاتی) وجود داشته باشد که اطلاعاتی در خود ندارد و باید حذف گردد. در استاندارد POCSAG، داده‌ها به دو دسته idle و غیر idle قابل طبقه‌بندی است. در این استاندارد داده دسته غیر idle (دارای محتوا) عملاً سرایندی ندارد و تنها بیت اول باعث تفکیک داده به دو نوع آدرس و پیام می‌شود. شناسایی بسته‌های اطلاعاتی بدون محتوا باید در این بخش انجام شود.

[1] Slow Associated Control Channel
[2] Stand Alone Dedicated Control Channel
[3] Fast Associated Control Channel
[4] Traffic Channel
[5] Special Filling Bits

در استاندارد WiMAX هر برست بیتی از چند MAC PDU[1] تشکیل شده است که باید در همان مرحله اول نشان‌داده‌شده در شکل (۱)، شناسایی و تفکیک گردند. در این مرحله، با انجام خوشه‌بندی داده به دو بخش MAC PDU و بیت‌های پرکننده مخصوص تفکیک می‌گردد که بخش بیت‌های پرکننده مخصوص بی‌ارزش است و باید دور ریخته شود. همچنین MAC PDU ها باید برای تحلیل به مرحله بعد بروند. در استاندارد DVB-S2 تمام داده‌ها از نوع خاص فریم باند پایه هستند که باید توسط شناساگر کور در یک خوشه قرار گیرند.

در استاندارد TETRA، دو حالت رخ می‌دهد. در حالتی که پیام SDU با اندازه بلوک MAC مطابقت دارد، بسته‌ها دارای MAC header در زیرلایه MAC بالایی و سرایند LLC[2] در زیرلایه LLC هستند. در حالتی که چنین نیست، داده بخش‌بندی می‌گردد و ابتدا فقط با سرایند زیرلایه MAC بالایی طرف هستیم. شناساگر کور باید قادر به خوشه‌بندی داده به دست‌کم دو نوع پیام باشد.

در استاندارد DMR، با خوشه‌بندی بسته‌های دریافتی در ورودی بخش شناسایی لایه سرویس، داده‌های صوت و متن به صورت داده‌های بدون سرایند تشخیص داده می‌شوند و برای تحلیل به بخش خوشه‌بندی داده Payload می‌روند. داده‌های سیگنالینگی نیز مانند GSM و TETRA، داده‌هایی با سرایند تشخیص داده می‌شوند که باید به طریق مشابه مورد تحلیل قرار گیرند.

در استاندارد LTE، بسته‌های MAC PDU برای تحلیل وارد سامانه شناسایی کور لایه سرویس می‌شوند. این بسته‌ها علاوه بر سرایند لایه MAC شامل یک یا چند MAC SDU[3] هستند. در جریان چرخه تحلیل، بسته‌های مراحل بعد نیز به ترتیب دارای سرایند لایه‌های RLC[4]، PDCP[5] و IP[6] هستند. فرایندی مشابه نیز برای استاندارد HSDPA قابل اجرا است.

۲-۲. تحلیل اطلاعات سرایند در خوشه‌های مختلف، استخراج اطلاعات مفید از سرایند و حذف سرایند

تحلیل سرایند به دو دولیل دارای اهمیت است. اول اینکه ممکن است اطلاعات مفیدی در مورد شبکه و یا کاربران در اختیار قرار دهد. دوم اینکه، ممکن است این اطلاعات برای رسیدن به اطلاعات لایه‌های بالاتر مورد نیاز باشند. خوشه‌بندی داده‌ها بر اساس نوع سرایند و تحلیل اطلاعات سرایند در بیشتر مواقع باید به‌صورت همزمان انجام شود.

در GSM اطلاعات سرایند لایه ۱ کانال SAACH توان موبایل‌ها و فاصله‌ی آنها از BTS را نمایش می‌دهد. در همین استاندارد، اطلاعات لایه ۲ نوع A می‌توانند نحوه‌ی اتصال بسته‌های لایه ۲ و رسیدن به بسته‌های لایه ۳ را مشخص نمایند. در استاندارد FLEX بخش سرایند داده اطلاعاتی است که در آن آدرس پیجرها و محل شروع پیام آنها را مشخص می‌نماید. سامانه شناسایی کور باید بتواند این بخش از بسته را از قسمت متنی تفکیک نماید. در استاندارد POCSAG، در بسته‌های ابتدایی دریافتی سرایندی وجود ندارد و شناساگر کور به مرحله بعد می‌رود. اما در مرحله بعد و پس از اتصال بسته‌های متوالی قرارگرفته پس از هر idle، به بسته‌های بزرگ‌تری می‌رسیم که پیام ارسالی به یک پیجر است. در این پیام ابتدا آدرس پیجر به عنوان سرایند قرار دارد و پس از آن پیام ارسالی قرار دارد.

در استاندارد WiMAX، MAC PDUهایی که از مرحله قبل تحویل این بخش شده‌اند، باید مورد تحلیل قرار گرفته، بخش Payload از سرایند آنها جدا شده و تحویل مرحله بعد گردد.

در استاندارد DVB-S2، با توجه به وجود CRC[7] در انتهای هر سرایند، پس از شناسایی CRC و با توجه به طول ثابت سرایند و همچنین وجود الگوهای تکراری، این بخش قابل شناسایی و تفکیک خواهد بود. پس از شناسایی سرایند با توجه به رابطه بخش DFL[8] با طول بخش داده، می‌توانیم این بخش را شناسایی کنیم و آن را از بیت‌های پرکننده مخصوص جدا کنیم. سایر اطلاعات موجود در سرایند در این مرحله از شناسایی کاربردی برای ما ندارد.

در چرخه تحلیل استاندارد LTE، هر کدام از سرایندهای لایه‌های MAC، RLC، PDCP و IP می‌توانند به اتصال بسته‌های لایه فعلی و رسیدن به بسته‌های لایه بالایی کمک نمایند. فرایندی مشابه نیز برای استاندارد HSDPA قابل اجرا است.

۲-۳. بررسی اتصال یا بخش‌بندی Payloadها و یا بسته‌ها برای تشکیل بسته‌های لایه بالاتر

برخی اوقات لازم است که چند بسته به یکدیگر متصل شوند تا در لایه‌ای بالاتر بتوان محتوای آنها را تحلیل نمود. به‌عنوان مثال، در استاندارد GSM، زمانی که طول اطلاعات لایه ۳ از حدی

[1] Protocol Data Unit
[2] Logical Link Control
[3] Service Data Unit
[4] Radio Link Control
[5] Packet Data Convergence Protocol
[6] Internet Protocol
[7] Cyclic Redundancy Check
[8] Data Field Length

بزرگ‌تر باشد، با استفاده از بخش‌بندی آن در بسته‌های کوچک تر لایه ۲ می‌توان آن را ارسال نمود. لذا گیرنده کور باید با اتصال چندین بسته لایه ۲ به یک بسته لایه ۳ برسد.

در استاندارد FLEX اطلاعات چند کاربر در قالب یک پیام ارسال می‌شود و با توجه به ساختار پیام ممکن است انجام بخش‌بندی و رسیدن به پیام هر یک از کاربران ممکن نباشد. هر چند اگر بتوان اطلاعات سرایند را تحلیل نمود، محل شروع هر پیام و آدرس پیجرها قابل استخراج است. در استاندارد POCSAG، با دیدن بسته‌های Idle می‌توان به نوعی قاب‌بندی غیرهمزمان رسید که در آن پس از هر Idle یک بسته آدرس وجود دارد و به دنبال آن پیام‌ها قرار می‌گیرند. با اتصال تمام پیام‌های بعد از یک Idle به بسته بزرگ‌تری می‌رسیم که همان پیام ارسالی به یک پیجر است.

در استاندارد WiMAX، در شروع کار و با دریافت یک برست بیتی و با جستجو برای کلمات کد CRC با طول متغیر از بین داده‌های کدشده و کدنشده می‌توان MAC PDUها را با بخش‌بندی این برست بیتی به‌دست آورد.

در استاندارد DVB-S2، بخش داده مربوط به فریم باند پایه بخشی از رشته پیوسته و یا بسته‌بندی‌شده است. در صورتی‌که داده داخل بخش داده از نوع بسته‌بندی‌شده باشد، باید CRC-8 در ابتدای هر بسته شناسایی و بسته‌ها پشت سرهم قرار گرفته و تحلیل شوند. در صورتی که داده داخل بخش داده از نوع پیوسته باشد، CRC-8 وجود نخواهد داشت و باید رشته‌ای پیوسته تحلیل گردد. باید دقت کرد که در صورت وجود داده بسته‌بندی‌شده در بخش داده، نباید CRC-8 موجود در بخش BBHeader[1] را با CRC-8 های موجود در بخش داده اشتباه بگیریم. رشته ورودی ممکن است یک رشته انتقالی باشد که همواره به‌صورت بسته‌بندی شده در بسته‌های کاربر به طول ۱۸۸ بایت قرار می‌گیرند، نمونه‌ای از این بسته‌ها بسته‌های MPEG[2] است. همچنین رشته ورودی می‌تواند یک رشته عمومی باشد که در دو حالت پیوسته و بسته‌بندی شده وجود دارد. این رشته‌ها با پارامتر طول بسته کاربر شناسایی می‌شوند، در صورتی که این پارامتر برابر صفر باشد، رشته پیوسته است، در غیر این‌صورت طول یک بسته درنهایت برابر با ۶۴۰۰۰ بیت است که بیان‌گر رشته بسته‌بندی شده خواهد بود. طول بسته‌های داده در بخش سرایند فریم باند پایه آمده است. در مرحله قبل از این اطلاعات استفاده نکردیم، اما در این مرحله می‌توان از اطلاعات آن و اطمینان از طول یک بسته داده استفاده کرد.

در استاندارد TETRA و در زمانی که پیام SDU بزرگ‌تر از اندازه بلوک MAC است، با اتصال Payloadهای زیرلایه LLC می‌توان به TL-SDU رسید. جهت رسیدن به این اتصال می‌توان از FCS[3] تعبیه‌شده در TL-SDU[4] بهره برد.

همان‌طور که قبلاً هم اشاره شد، در چرخه تحلیل استاندارد LTE، هر کدام از سرایندهای لایه‌های MAC، RLC، RDCP و IP می‌توانند به اتصال بسته‌های لایه فعلی و رسیدن به بسته‌های لایه بالایی کمک نمایند. فرایندی مشابه نیز برای استاندارد HSDPA قابل اجرا است.

۲-۴. خوشه‌بندی و دسته‌بندی Payloadها به داده‌های چندرسانه‌ای مختلف

برای رسیدن به محتوای مبادله‌شده در شبکه لازم است که داده Payload ها دسته‌بندی شود. برای مثال، در استاندارد GSM، سامانه شناسایی کور باید تشخیص دهد که داده دریافتی از کانال TCH داده صوتی است. همچنین داده‌های صوتی که با کدگذارهای مختلف کد شده‌اند باید از یکدیگر متمایز گردند و در خوشه‌های مختلف قرار گیرند. مسئله دیگر جداسازی داده‌های صوتی مربوط به نشست‌های مختلف است. همچنین باید بتوان داده سایر کانال‌های منطقی را به دو نوع داده متنی و سیگنالینگی دسته‌بندی نمود. به طریق مشابه، در استاندارد TETRA و DMR نیز باید بتوان داده‌های خام صوت، متنی و سیگنالینگی را از یکدیگر تفکیک گردد.

در استاندارد FLEX و POCSAG محتوا از نوع متنی است که باید توسط شناساگر کور شناسایی گردد. در استاندارد WiMAX، داده Payload تحویلی از بخش دو نوع می‌تواند باشد: ATM و Packet (برای انتقال داده‌هایی مانند IP و Ethernet). هر یک از این دو نوع داده باید جهت تحلیل، مجدداً به مرحله ۱ انتقال داده شوند. نکته‌ی مهمی که باید به آن اشاره شود، این است که محتوای این بسته‌ها می‌تواند از نوع صوت، تصویر و ... باشد. با این حال شناساگر کور باید تشخیص دهد که این اطلاعات در قالب یک سیگنالینگ در حال انتقال هستند و داده خام صوت، تصویر و ... نیستند. این موضوع برای استانداردهای DVB-S2، LTE و HSDPA نیز صادق است.

[1] Baseband Header
[2] Moving Picture Experts Group
[3] Frame Check Sequence
[4] SDU from the service user

۲-۵. شناسایی پارامترهای داده چندرسانه‌ای و استخراج محتوا

بعد از تعیین نوع کلی داده چندرسانه‌ای، برای رسیدن به محتوای ارسال‌شده باید پارامترهای داده چندرسانه‌ای استخراج گردد. برای مثال در GSM باید بتوان پارامترهای کدگذار صوت را شناسایی نمود و یا مشخص نمود که از چه روشی برای کدکردن پیامک استفاد شده است.

نوع پیام در استاندارد FLEX و POCSAG متنی است که باید روش کدگذاری آن توسط سامانه شناسایی کور استخراج گردد تا بتوان متن ارسالی را به‌دست آورد. در استانداردهایی مانند DMR، TETRA و GSM باید بتوان اطاعات خام سیگنالینگ را دسته‌بندی کرد و با استفاده از روش های داده‌کاوی آنها را تفسیر نمود.

یک مسئله بسیار مهم در تحلیل داده‌های چنددرسانه‌ای تمامی استانداردها، مانند GSM، TETRA، DMR و ... وجود دارد. داده‌هایی که در یک کانال منطقی خاص دریافت شده‌اند ممکن است مربوط به چند نشست مختلف باشند و صرف این‌که پارامترهای مربوط به کدگذار آنها یکسان باشد، نمی‌توان نتیجه گرفت که مربوط به یک کاربر هستند. لذا یک مسئله مهم در شناسایی داده‌های چندرسانه‌ای خوشه‌بندی داده‌های مربوط به یک کدک[1] مشخص به چند خوشه (نشست) است.

۳. تحقیقات موجود در شناسایی کور لایه سرویس

به‌منظور تعیین مسائل حل‌نشده در حوزه شناسایی کور لایه سرویس، در این بخش تحقیقات موجود در این حوزه و ارتباط آن‌ها با طرح جامع پیشنهادی شکل (۱) مورد بررسی قرار می‌گیرد. تحقیقات موجود در حوزه شناسایی کور لایه سرویس را می‌توان به عموماً بر روی دو حوزه تمرکز داشته‌اند. دسته‌ای از آن‌ها شناسایی کور پروتکل‌های مخابراتی را مد نظر قرار داده‌اند. هدف اصلی این دسته از مقالات عموماً تعیین موقعیت و طول فیلدهای موجود در بسته‌های دریافتی در شبکه‌های مخابراتی است. دسته دیگر مقالات نیز تمرکز خود را بر روی شناسایی کدهای منبع (صوت، تصویر و ...) قرار داده‌اند. در این بخش، مقالات این دو دسته را به تفکیک مورد بررسی قرار می‌دهیم.

۳-۱. شناسایی کور پروتکل‌های مخابراتی

شناسایی کور پروتکل‌های مخابراتی از جنبه‌های امنیتی مختلفی در منابع علمی مورد توجه قرار گرفته است. یکی از این جنبه‌های کاربرد در سامانه‌های تشخیص نفوذ مانند Snort و Bro است. جنبه دیگر استفاده از این موضوع در نرم‌افزارهای مدیریت ترافیک شبکه است. نرم‌افزارهای تست برنامه‌های کاربردی نیز حوزه دیگری از کاربرد این موضوع هستند.

طبق مطالعات انجام‌شده، مقالات حوزه شناسایی کور پروتکل‌های مخابراتی دو دسته هدف را دنبال می‌کنند. یک دسته تعیین موقعیت و طول فیلدهای موجود در بسته‌های دریافتی را مد نظر قرار داده‌اند. دسته دیگر نیز بر روی تفسیر پیام‌ها و استخراج ماشین حالت پیام‌ها تمرکز داشته‌اند.

تقریباً تمام تمرکز مقالات موجود در این حوزه بر روی اینترنت و پروتکل‌های بر بستر اینترنت است. برای روش شدن اهمیت موضوع، کافی است به این موضوع توجه کنیم که حدود ۴۰٪ از ترافیک اینترنت مربوط به پروتکل‌هایی است که شناخته‌شده نیستند [۱۱]. کارهای اولیه در این حوزه کارهایی نیمه‌کور هستند. به‌عنوان مثال در [۱۲]، دسته‌بندی بر روی جریان‌های داده مربوط به پنج نوع فایل WMF، BMP، JPG، PNG و TIF و پنج پروتکل شبکه‌ای DNS، RPC، TFTP، HTTP و FTP را مد نظر قرار داده است.

تفسیر پیام‌ها (یعنی استخراج نوع و محتوای پیام‌ها) و استخراج ماشین حالت یک پروتکل اینترنتی مانند Skype و یا SMB که مشخصات فنی آن‌ها منتشر نشده است، مسئله‌ای بسیار پیچیده است [۱۳]. در تأیید پیچیده بودن مهندسی معکوس همین بس که ۱۲ سال طول کشید که پروتکل SMB شرکت Microsoft مهندسی معکوس شود [۱۳]. اصولاً، فرایند مهندسی معکوس یک فرایند دستی است که با تولید ابزارهای مختلف می‌توان به آن سرعت بیشتری بخشید [۱۴]. به‌عبارت دیگر، انتظار خودکار شدن این فرایند، دست‌کم در شرایط موجود منطقی به نظر نمی‌رسد.

یکی از روش‌های پایه برای تفسیر پیام‌ها و استخراج ماشین حالت پروتکل این است که پیام‌های نشست‌های مختلف برای یافتن الگوهای ساختاری یا بایتی مشخص مورد بررسی قرار گیرد [۱۵]. هرچند در برخی موارد این روش کارایی خود را نشان داده است، اما در عمل محدودیت‌هایی جدی دارد. به‌عنوان مثال، به دلیل عدم اطلاع از پروتکل، پیام‌هایی از نوع یکسان ممکن است متفاوت دیده شوند که این مسئله منجر به تشخیص اشتباه می‌گردد.

[1] Codec

۳-۲. شناسایی کدهای منبع

از اولین کارهای مرتبط با شناسایی کدگذارهای منبع شناسایی فایل‌های کامپیوتری است. روش‌های سنتی و پیش پا افتاده برای این کار از پسوند فایل (مورد استفاده در سیستم‌های عامل مایکروسافت)، اعداد طلایی فایل (مورد استفاده در سیستم‌های عامل یونیکس) و یا سرایند/ته‌آیند فایل[1] استفاده می‌کنند [۱۶]. اما این روش‌ها محدودیت‌های زیادی دارند و عملاً در یک شبکه کامپیوتری قابل استفاده نیستند. دو دلیل عمده می‌توان برای این موضوع بیان کرد: ۱- پسوند یا اعداد طلایی فایل می‌توانند تغییر داده شوند و ۲- در یک شبکه مخابراتی، ممکن است فقط بخشی از یک فایل را دریافت کرده باشیم.

اولین مقاله‌ای که فرض داشتن ابتدای فایل را در نظر نمی‌گیرد توسط کلهون و همکاران نوشته شد [۱۷]. آن‌ها استفاده از دو روش ۱- جداساز خطی و ۲- بزرگ‌ترین زیر رشته‌های مشترک را برای شناسایی قطعات فایل پیشنهاد دادند و توانستند با دقت نزدیک به ۸۰٪ برخی از انواع فایل متداول (مانند JPG، PDF، GIF و ...) را از یکدیگر تفکیک نمایند.

جدای از تشخیص دقیق نوع یک فایل، شاید یک کار مهم در شناسایی، خوشه‌بندی داده‌های چندرسانه‌ای با توجه به نوع آن‌ها (صوت، متن، تصویر و ...) باشد. این موضوع توسط ژنگ و همکارانش مورد بررسی قرار گرفته است [۱۸]. آن‌ها برای انجام این کار پیشنهاد استفاده از یک الگوریتم C-Means احتمالی مرتبه بالا را داده‌اند.

۴. مسائل حل‌نشده

در این بخش، با توجه به تحقیقات انجام‌شده در حوزه شناسایی کور لایه سرویس و با در نظر گرفتن طرح جامع پیشنهادی شکل (۱)، مسائل باز در این حوزه معرفی خواهد شد.

۴-۱. خوشه‌بندی داده‌ها بر اساس وجود و عدم وجود سرایندها و تحلیل سرایندها

همان‌طور که اشاره شد، تقریباً تمام تمرکز مقالات موجود در این حوزه بر روی اینترنت و پروتکل‌های بر بستر اینترنت است و عموماً فرض می‌شود که برنامه‌ای در دسترس است که قابلیت برقراری ارتباط با پروتکل مورد نظر را دارد. با توجه به تحقیقات انجام‌شده، مسائل باز در این حوزه را می‌توان در موارد زیر خلاصه نمود:

- بررسی مسئله خوشه‌بندی داده‌ها بر اساس وجود و عدم وجود سرایند
- تحلیل کاملاً کور سرایند بر اساس مشاهدات تصادفی از نشست‌ها و پیام‌ها
- تفسیر کاملاً کور پیام‌ها و تحلیل کور پروتکل بر اساس مشاهدات تصادفی از نشست‌ها و پیام‌ها

۴-۲. بررسی اتصال یا بخش‌بندی بسته‌ها برای تشکیل بسته‌های لایه بالاتر

با توجه به اینکه به طور مستقیم در این حوزه تحقیقاتی انجام نشده است ، پیشنهاد برای کارهای آتی را می‌توان در موارد زیر خلاصه نمود:

- جداسازی بخش‌های مختلف متنی، صوتی و ... در یک بسته یا رشته بیت
- جستجو برای کلمات کد CRC با طول متغیر از بین داده‌های کدشده و کدنشده
- استفاده کور از اطلاعات سرایند جهت اتصال بسته‌ها

۴-۳. شناسایی و دسته‌بندی داده‌های چندرسانه‌ای

با توجه به تحقیقات موجود، پیشنهاد برای کارهای آتی را می‌توان در موارد زیر خلاصه نمود:

- اثر داده ناشناخته در دسته‌بندی
- اثر تغییر پارامترهای کدگذار در دسته‌بندی
- توسعه روش‌های شناسایی و دسته‌بندی برای حالتی که تعداد زیادی کدگذار (صوت، تصویر و یا ...) محتمل مفروض است
- خوشه‌بندی داده‌های مربوط به یک کدک مشخص به چند خوشه (خوشه‌بندی به چند نشست)
- دسته‌بندی اطاعات خام سیگنالینگ و تفسیر آن‌ها با استفاده از روش‌های داده‌کاوی
- تمرکز بیشتر بر روی قیود شبکه‌های مخابراتی و اثر آن‌ها در شناسایی

۵. نتیجه‌گیری

در این مقاله، طرحی جامع برای شناسایی کور لایه سرویس ارائه گردید. در این راستا، نه استاندارد مختلف مخابراتی و شناسایی کور آن‌ها در لایه سرویس مورد بررسی دقیق قرار گرفت. این نُه

[1] File Header/Trailer

استاندارد که شامل استانداردهای FLEX و POCSAG برای پیجرها، استانداردهای GSM، TETRA و DMR برای ارتباط تلفنی و بی‌سیم سلولی، استانداردهای LTE، HSDPA و WiMAX برای ارتباطات مبتنی بر داده و استاندارد DVB-S2 برای ارتباطات ماهواره‌ای هستند، نماینده طیف وسیعی از سیستم‌های مخابراتی یک‌سویه، دوسویه، سلولی و ماهواره‌ای و ... هستند.

در ادامه مقاله، تحقیقات موجود در زمینه شناسایی کور لایه سرویس و ارتباط آن‌ها با طرح جامع پیشنهادی مورد بررسی قرار گرفت. بر این اساس مسائل متنوعی به‌عنوان مسائل باز و حل‌نشده معرفی گردیدند.

۶. مراجع

[1] S. Kleber, L. Maile, and F. Kargl, "Survey of Protocol Reverse Engineering Algorithms: Decomposition of Tools for Static Traffic Analysis," IEEE Communications Surveys & Tutorials, 2018.

[2] K. Zhang, E. L. Xu, Z. Feng, and P. Zhang, "A Dictionary Learning Based Automatic Modulation Classification Method," IEEE Access, vol. 6, pp. 5607-5617, 2018.

[3] R. Gupta, S. Majhi, and O. A. Dobre, "Design and Implementation of a Tree-Based Blind Modulation Classification Algorithm for Multiple-Antenna Systems," IEEE Transactions on Instrumentation and Measurement, 2018.

[4] T. Li, Y. Li, Y. Chen, L. J. Cimini Jr, and H. Zhang, "Estimation of MIMO Transmit-Antenna Number Using Higher-Order Moments Based Hypothesis Testing," IEEE Wireless Communications Letters, vol. 7, pp. 258-261, 2018.

[5] O. A. Dobre, "Signal Identification for Emerging Intelligent Radios: Classical Problems and New Challenges," IEEE Instrumentation & Measurement Magazine, vol. 18, pp. 11-18, 2015.

[6] A. G. Soteh and H. K. Bizaki, "On the Analytical Solution of Rank Problem in the Convolutional Code Identification Context," IEEE Communications Letters, vol. 20, pp. 442-445, 2016.

[7] R. Swaminathan, A. Madhukumar, G. Wang, and T. S. Kee, "Blind Reconstruction of Reed-Solomon Encoder and Interleavers Over Noisy Environment," IEEE Transactions on Broadcasting, vol. 64, pp. 830-845, 2018.

[8] P. Yu, J. Li, and H. Peng, "A least square method for parameter estimation of RSC sub-codes of turbo codes," IEEE Communications Letters, vol. 18, pp. 644-647, 2014.

[9] A. Jamshidi, A. Keshavarz-Hadad, and F. Zare, "Estimation of Convolutional Interleaver Parameters in the Burst and BSC Channels," Iranian Journal of Science and Technology, Transactions of Electrical Engineering, vol. 40, pp. 93-102, 2016.

[10] S. Han and M. Zhang, "A Method for Blind Identification of a Scrambler Based on Matrix Analysis," IEEE Communications Letters, vol. 22, pp. 2198-2201, 2018.

[11] Y. Wang, X. Yun, M. Z. Shafiq, L. Wang, A. X. Liu, Z. Zhang, et al., "A Semantics Aware Approach to Automated Reverse Engineering Unknown Protocols," in 20th IEEE International Conference on Network Protocols, pp. 1-10, 2012.

[12] W. Cui, M. Peinado, K. Chen, H. J. Wang, and L. Irun-Briz, "Tupni: Automatic Reverse Engineering of Input Formats," in 15th ACM conference on Computer and communications security, pp. 391-402, 2008.

[13] Z. Lin, X. Jiang, D. Xu, and X. Zhang, "Automatic Protocol Format Reverse Engineering through Context-Aware Monitored Execution," in NDSS, pp. 1-15, 2008.

[14] P. M. Comparetti, G. Wondracek, C. Kruegel, and E. Kirda, "Prospex: Protocol Specification Extraction," in Security and Privacy, 2009 30th IEEE Symposium on, pp. 110-125, 2009.

[15] G. Wondracek, P. M. Comparetti, C. Kruegel, E. Kirda, and S. S. S. Anna, "Automatic Network Protocol Analysis," in NDSS, pp. 1-14, 2008.

[16] M. McDaniel and M. H. Heydari, "Content Based File Type Detection Algorithms," in 36th Annual Hawaii International Conference on System Sciences, p. 10, 2003.

[17] W. C. Calhoun and D. Coles, "Predicting the Types of File Fragments," digital investigation, vol. 5, pp. S14-S20, 2008.

[18] Q. Zhang, L. T. Yang, Z. Chen, and F. Xia, "A High-Order Possibilistic C-Means Algorithm for Clustering Incomplete Multimedia Data," IEEE Systems Journal, vol. 11, pp. 2160-2169, 2017.

Blind Identification of Communications Networks in Service Layer

M. Teimouri*, H. R. Kakaei Motlagh, J. Garshasbi

Abstract

One of the major challenges in electronic and cyber warfare is to achieve maximum information from the intercepted signals of a communications network. To this aim, the first step is reverse engineering of the physical and data transmission layers in order to obtain the transmitted packets. In the next step, service layer should be identified. In this step, the exact meaning of each packet along with the communications state machine should be identified. By completing these tasks, the information content exchanged through the network would be available. In this paper, a comprehensive scheme is proposed for identification of the service layer. With the intention of proving proficiency of the proposed scheme, usability of this scheme for identification of nine different standards, including GSM, LTE, HSDPA, DVB-S2, FLEX, POCSAG, TETRA, DMR, and WiMAX, is studied. Finally, open problems in the field of blind identification of service layer are introduced along with the current methods.

Key Words: *Blind Identification, Communications Networks, Service Layer*

* University of Tehran, (mehditeimouri@ut.ac.ir) - Writer-in-Charge