دانشکده مهندسی کامپیوتر

کارآموزی پروژه بهینه‌سازی تخصیص منابع و بارسپاری در شبکه محاسبات لبه‌ای

محل کارآموزی: آزمایشگاه جامع (اتاق ۱۰۴)، دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران

داریوش سیف اشرفی

استاد راهنما

جناب آقای دکتر صالح اعتمادی

آبان ۱۴۰۴

بسم الله الرحمن الرحيم

تأییدیه‌ی صحت و اصالت نتایج

باسمه تعالی

اینجانب داریوش سیف اشرفی به شماره دانشجویی **۴۰۱۵۲۱۳۳۶** دانشجوی رشته مهندسی کامپیوتر مقطع تحصیلی کارشناسی تأیید می‌نمایم که کلیه‌ی مطالب مندرج در این گزارش حاصل **۳۰۰** ساعت حضور و کار اینجانب در شرکت/کارخانه آزمایشگاه تحقیقاتی محاسبات توزیع‌شده و مقیاس‌پذیر دکتر انتظاری و بدون هرگونه دخل و تصرف است و موارد نسخهبرداریشده از آثار دیگران را با ذکر کامل مشخصات منبع ذکر کردهام. درصورت اثبات خلاف مندرجات فوق، به تشخیص دانشگاه مطابق با ضوابط و مقررات حاکم آموزشي، پژوهشي و انضباطي با اینجانب رفتار خواهد شد و حق هرگونه اعتراض درخصوص احقاق حقوق مکتسب و تشخیص و تعیین تخلف و مجازات را از خویش سلب مینمایم.

نام و نام خانوادگی: داریوش سیف اشرفی

تاریخ و امضا: آبان ۱۴۰۴

قدردانی

سپاس خداوندگار حکیم را که با لطف بی‌کران خود، آدمی را زیور عقل آراست.

در آغاز وظیفه خود می‌دانم از زحمات بی‌دریغ سرپرست کارآموزی خود، جناب آقای دکتر انتظاری، صمیمانه تشکر و قدردانی کنم که قطعاً بدون راهنمایی‌های ارزنده ایشان، این مجموعه به انجام نمی‌رسید.

از جناب آقای محسن دارچینی که در مسیر یادگیری و آمادگی برای حضور در این صنعت به اینجانب کمک کردند و در کسب دانش و تجربه در این رشته و در زمینه آشنایی با این حوزه تحقیقاتی، از هیچ کمکی دریغ نکردند نیز کمال تشکر را دارم.

داریوش سیف اشرفی

آبان ۱۴۰۴

چکیده

آزمایشگاه جامع (اتاق ۱۰۴) واقع در طبقه اول دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران، زیرنظر دکتر انتظاری، محیطی پژوهشی است که دانشجویان در مقاطع مختلف زیرنظر اساتید به کار بر روی مباحث نوین و پیشرفته در حوزه‌های مختلف علوم کامپیوتر مشغول می‌باشند. من در طول مدت ۳۰۰ ساعت کارآموزی به عنوان توسعه‌گر پروژه بهینه‌سازی تخصیص منابع و بارسپاری در شبکه محاسبات لبه‌ای در این آزمایشگاه مشغول به فعالیت بودم. در ابتدا، به یادگیری پیش‌نیازهای ضروری شامل یادگیری ماشین، یادگیری عمیق، یادگیری تقویتی و الگوریتم‌های بهینه‌سازی هیورستیک و متاهیورستیک پرداختم. سپس با مطالعه مقالات مروری و بررسی معماری‌های محاسباتی مختلف، درک جامعی از حوزه محاسبات لبه‌ای و چالش‌های آن به دست آوردم. در مرحله بعد، به مطالعه دقیق و پیاده‌سازی کامل مقاله پایه "Robust Task Offloading and Resource Allocation Under Imperfect Computing Capacity Information in Edge Intelligence Systems" پرداختم که یک رویکرد بهینه‌سازی مقاوم برای بارسپاری جزئی و تخصیص منابع تحت شرایط اطلاعات ناقص ارائه می‌دهد. پس از آن، با مطالعه بیش از ۱۵ مقاله پژوهشی در حوزه‌های یادگیری تقویتی چندعامله، پیش‌بینی مسیر، و بارسپاری وابسته به وظایف، به دنبال بهبود نتایج کار قبلی خود برآمدم. در حال حاضر، مشغول توسعه یک سیستم پیشرفته‌تر با استفاده از مدل‌های مبتنی بر Diffusion برای پیش‌بینی مقصد خودروهای خودران و ایجاد تعادل بار بین سرورهای لبه‌ای هستم. همچنین، الگوریتم‌های نوینی برای مدیریت صف وظایف و بارسپاری جزئی (Partial Offloading) در حال پیاده‌سازی می‌باشد.

واژگان کلیدی: یادگیری ماشین، یادگیری عمیق، یادگیری تقویتی، بارسپاری جزئی، تخصیص منابع، محاسبات لبه‌ای

فهرست مطالب

فهرست تصاویر ح

فهرست جداول خ

فصل ۱: معرفی واحد کارآموزی ۱

۱ـ۱ مقدمه ۱

۱ـ۲ شرح مختصر از موضوع کارآموزی و اهمیت آن ۱

۱ـ۳ چالش‌های حوزه محاسبات لبه‌ای ۳

فصل ۲: شرح فعالیت‌های انجام شده ۴

۲ـ۱ مقدمه ۴

۲ـ۲ شرح پروژه‌ها و فعالیت‌های انجام شده ۴

۲ـ۲ـ۱ بخش آموزش پیش‌نیازها ۴

۲ـ۲ـ۲ فاز اول: مطالعه مقالات مروری و آشنایی با حوزه Edge Computing ۷

۲ـ۲ـ۳ فاز دوم: مطالعه و پیاده‌سازی مقالات کاربردی ۱۱

۲ـ۲ـ۴ فاز سوم: پیاده‌سازی کامل مقاله پایه ۱۲

۲ـ۲ـ۵ فاز چهارم: مطالعه مقالات پیشرفته برای بهبود ۱۵

۲ـ۲ـ۶ فاز پنجم: توسعه و بهبود سیستم (کار جاری) ۱۹

فصل ۳: نتیجه‌گیری ۲۴

۳ـ۱ جمع‌بندی ۲۴

۳ـ۲ دستاوردها ۲۴
۳ـ۲ـ۱ دستاوردهای علمی ۲۴
۳ـ۲ـ۲ دستاوردهای فنی ۲۶
۳ـ۲ـ۳ دستاوردهای شخصی ۲۶
۳ـ۳ چالش‌ها و پیشنهادات ۲۷
۳ـ۳ـ۱ چالش‌های علمی و فنی ۲۷
۳ـ۳ـ۲ چالش‌های زیرساختی و مدیریتی ۲۷
۳ـ۴ پیشنهادات برای بهبود ۲۷
۳ـ۵ نتیجه‌گیری نهایی ۲۸

پیوست آ: جدول کامل مقالات مطالعه شده ۲۹

پیوست ب: منابع و مراجع ۳۳
ب ـ ۱ دوره‌های آموزشی ۳۳
ب ـ ۲ ابزارها و فریمورک‌ها ۳۳
ب ـ ۳ وب‌سایت‌ها و منابع آنلاین ۳۴

فهرست تصاویر

۱ـ۱ معماری پایه‌ای شبکه لبه‌ای ۲
۲ـ۱ معماری شبکه ۶لایه ۲۲
۲ـ۲ کاربرد عملی پیش‌بینی مسیر حرکت خودرو ۲۳

فهرست جداول

آ_۱ مقالات مروری (Survey Papers) ۳۰
آ_۲ مقالات پیاده‌سازی اولیه ۳۱
آ_۳ مقاله پایه (Base Paper) ۳۱
آ_۴ مقالات یادگیری تقویتی چندعامله ۳۱
آ_۵ مقالات پیش‌بینی مسیر و تعادل بار ۳۱
آ_۶ مقالات وابستگی بین وظایف ۳۲
آ_۷ مقالات بارسپاری جزئی پیشرفته ۳۲

فصل ۱

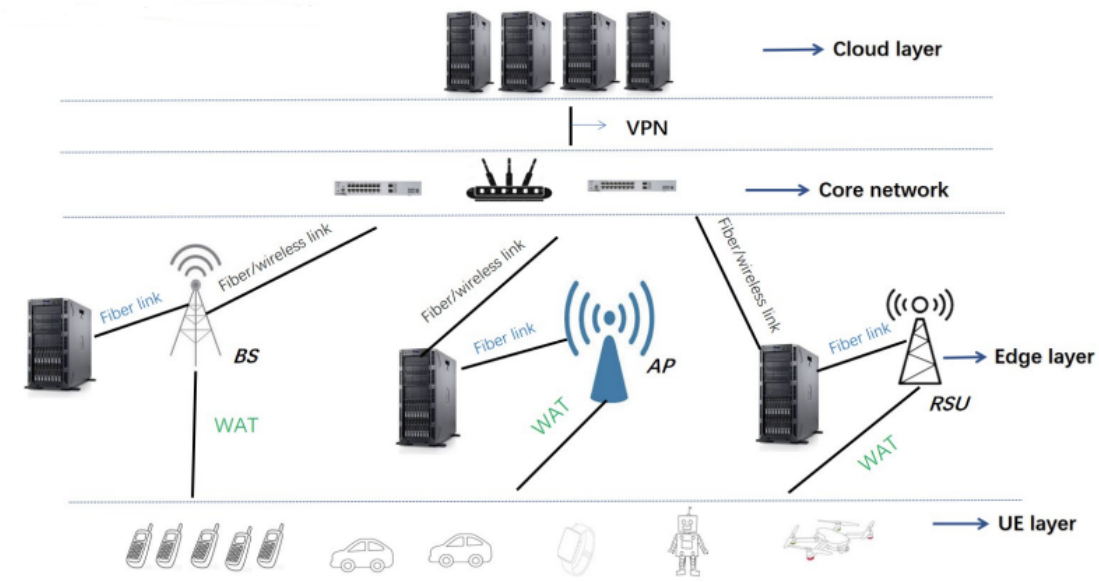
معرفی واحد کارآموزی

۱ـ۱ مقدمه

آزمایشگاه جامع (اتاق ۱۰۴) واقع در طبقه اول دانشکده مهندسی کامپیوتر دانشگاه علم و صنعت ایران، زیرنظر دکتر انتظاری، یک محیط پژوهشی فعال است که دانشجویان در مقاطع کارشناسی و کارشناسی ارشد و دکترا زیرنظر اساتید به کار بر روی مباحث نوین در حوزه‌های مختلف علوم کامپیوتر از جمله یادگیری ماشین، محاسبات لبه‌ای، شبکه‌های کامپیوتری و سیستم‌های توزیع‌شده مشغول می‌باشند. محیط این آزمایشگاه به گونه‌ای طراحی شده است که دانشجویان بتوانند در کنار یادگیری مفاهیم تئوری، به پیاده‌سازی و توسعه راه‌حل‌های عملی برای مسائل واقعی بپردازند. کار در این آزمایشگاه به صورت پروژه‌محور و با رویکرد پژوهش‌گرایانه انجام می‌شود.

۱ـ۲ شرح مختصر از موضوع کارآموزی و اهمیت آن

با گسترش استفاده از رایانش ابری (Cloud Computing) و مجازی‌سازی (Virtualization)، تعداد زیادی از دستگاه‌های اینترنت اشیا (IoT) در سطح جهان مستقر شده‌اند. این امر باعث شده است که وظایف محاسباتی سنگین از لبه (Edge) شبکه به سمت مراکز داده ابری متمرکز (Centralized Cloud Data Centers) منتقل شوند. این انتقال، ضمن افزایش مصرف انرژی (Energy Consumption) در مراکز داده ابری، تاخیر (Latency)



شکل ۱ـ۱: معماری پایه‌ای شبکه لبه‌ای

قابل‌توجهی را برای برنامه‌های حساس به زمان ایجاد می‌کند. در سال‌های اخیر، تغییر پارادایمی از رایانش ابری متمرکز به سمت محاسبات لبه موبایل (Mobile Edge Computing - MEC) رخ داده است. همانگونه که در شکل ۱ـ۱ مشاهده میکنید، در این رویکرد، وظایف محاسباتی به سرورهای لبه نزدیک به تجهیزات کاربر (User Equipment - UE) منتقل می‌شوند که منجر به: کاهش مصرف انرژی در مراکز داده ابری، کاهش تاخیر برای کاربران نهایی، استفاده بهینه از منابع در لبه شبکه و بهبود کیفیت خدمات (Quality of Service - QoS) اما با افزایش مقیاس و پیچیدگی شبکه‌های MEC، یافتن راه‌حل‌های بهینه از نظر انرژی برای بارسپاری محاسباتی (Task Offloading) و تخصیص منابع (Resource Allocation) به چالشی بزرگ تبدیل شده است.

از جمله کاربردهای عملی محاسبات لبه‌ای می‌توان به، خودروهای خودران (Autonomous Vehicles): نیاز به پردازش سریع داده‌های سنسورها، واقعیت افزوده/مجازی :(AR/VR) نیاز به تاخیر بسیار کم، اینترنت اشیای صنعتی :(IIoT) نیاز به قابلیت اطمینان بالا، شهرهای هوشمند: مدیریت ترافیک، امنیت عمومی و مراقبت‌های بهداشتی: نظارت بلادرنگ بر وضعیت بیماران اشاره کرد

۱ـ۳ چالش‌های حوزه محاسبات لبه‌ای

از چالش‌های اصلی این حوزه می‌توان به موارد زیر اشاره کرد:

- محدودیت منابع محاسباتی: سرورهای لبه‌ای دارای منابع محدودتری نسبت به مراکز داده ابری هستند.
- عدم قطعیت در اطلاعات: اطلاعات دقیق از ظرفیت محاسباتی سرورها همیشه در دسترس نیست.
- تحرک کاربران: در شبکه‌های موبایل، کاربران دائماً در حال تغییر مکان هستند.
- وابستگی بین وظایف: برخی وظایف وابسته به یکدیگر بوده و باید به ترتیب خاصی اجرا شوند.
- تعادل بار: توزیع مناسب وظایف بین سرورهای مختلف برای جلوگیری از اشباع.
- مصرف انرژی: کاهش مصرف انرژی در دستگاه‌های کاربر با باتری محدود.
- تاخیر **End-to-End**: حفظ تاخیر در حد مجاز برای برنامه‌های بلادرنگ.

حل این چالش‌ها نیازمند رویکردهای نوین و هوشمندانه است که بتوانند:

- به صورت پویا با تغییرات محیط سازگار شوند.
- تصمیمات بهینه در زمان واقعی اتخاذ کنند.
- در مقابل عدم قطعیت‌ها مقاوم باشند.
- از منابع محدود به بهترین شکل استفاده کنند.

این پژوهش با ترکیب تکنیک‌های یادگیری ماشین، بهینه‌سازی ریاضی و الگوریتم‌های هیورستیک، به دنبال ارائه راه‌حل‌های عملی برای این چالش‌ها است.

فصل ۲

شرح فعالیت‌های انجام شده

۲ـ۱ مقدمه

در این دوره کارآموزی، من به عنوان توسعه‌گر پروژه بهینه‌سازی تخصیص منابع و بارسپاری در شبکه محاسبات لبه‌ای، تلاش کردم تا علاوه بر یادگیری مفاهیم پایه و پیشرفته در حوزه‌های یادگیری ماشین و محاسبات لبه‌ای، به پیاده‌سازی و توسعه راه‌حل‌های عملی نیز بپردازم. این فرآیند شامل چندین فاز متوالی بود که هر کدام بر یکدیگر بنا شده و به تدریج پیچیدگی و عمق کار افزایش یافت. روند کار به صورت مستمر با سرپرست محترم در آزمایشگاه، دکتر انتظاری، و مربی گرامی، جناب آقای محسن دارچینی، هماهنگ و بررسی می‌شد.

۲ـ۲ شرح پروژه‌ها و فعالیت‌های انجام شده

۲ـ۲ـ۱ بخش آموزش پیش‌نیازها

برای آشنایی بیشتر با محیط پروژه و همچنین آشنایی با تکنولوژی‌های مورد نیاز برای توسعه پروژه، در ابتدای کارآموزی زمان قابل توجهی را صرف یادگیری پیش‌نیازهای ضروری کردم. این بخش شامل چهار حوزه اصلی بود:

۱. دوره یادگیری ماشین

منبع: دوره یادگیری ماشین دکتر سلیمانی و دکتر شریفی‌زارچی از دانشگاه شریف

محتوای یادگیری شده:

(آ) مفاهیم پایه یادگیری ماشین و انواع یادگیری (نظارت‌شده، بدون نظارت، تقویتی)

(ب) الگوریتم‌های رگرسیون خطی و لجستیک

(ج) ماشین بردار پشتیبان (SVM)

(د) درخت تصمیم و جنگل تصادفی (Random Forest)

(ه) خوشه‌بندی (K-Means, DBSCAN)

(و) کاهش بعد (PCA, t-SNE)

(ز) ارزیابی مدل‌ها و تکنیک‌های اعتبارسنجی متقابل (Cross-Validation)

(ح) پیش‌پردازش داده و مهندسی ویژگی (Feature Engineering)

فعالیت عملی: تمامی مطالب این دوره به طور کامل همراه با پیاده‌سازی در Python و مستندسازی جامع تهیه شدند.

۲. دوره یادگیری عمیق

منبع: دوره یادگیری عمیق دکتر سلیمانی از دانشگاه شریف

محتوای یادگیری شده:

(آ) شبکه‌های عصبی مصنوعی (ANN) و الگوریتم پس‌انتشار خطا

(ب) تکنیک‌های Regularization (Dropout, Batch Normalization)

(ج) بهینه‌سازها (SGD, Adam, RMSprop)

(د) تکنیک‌های Regularization (Dropout, Batch Normalization)

(ه) Transfer Learning و Fine-tuning

(و) شبکه‌های عصبی کانولوشنی (CNN)

(ز) شبکه‌های عصبی بازگشتی (RNN) و LSTM

(ح) معماری‌های مدرن (ResNet, Transformer)

فریم‌ورک استفاده شده: PyTorch

۳. دوره یادگیری تقویتی

منبع: دوره یادگیری تقویتی دکتر رهبان از دانشگاه شریف

محتوای یادگیری شده:

(آ) مفاهیم پایه: Agent, Environment, State, Action, Reward

(ب) فرآیندهای تصمیم‌گیری مارکوف (MDP)

(ج) الگوریتم‌های مبتنی بر ارزش (Q-Learning, SARSA)

(د) الگوریتم‌های مبتنی بر سیاست (Policy Gradient)

(ه) الگوریتم‌های Actor-Critic

(و) یادگیری تقویتی عمیق (Deep Q-Network - DQN)

(ز) الگوریتم‌های پیشرفته (PPO, DDPG)

۴. الگوریتم‌های بهینه‌سازی Heuristic و Meta-heuristic

علاوه بر یادگیری ماشین، به مطالعه و تحقیق در رابطه با برخی الگوریتم‌های بهینه‌سازی کلاسیک پرداختم که در بسیاری از مسائل تخصیص منابع کاربرد دارند:

الگوریتم‌های مطالعه شده:

(آ) الگوریتم ژنتیک (Genetic Algorithm)

(ب) بهینه‌سازی ازدحام ذرات (Particle Swarm Optimization - PSO)

(ج) الگوریتم کلونی مورچگان (Ant Colony Optimization - ACO)

(د) الگوریتم شبیه‌سازی تبرید (Simulated Annealing)

(ه) الگوریتم‌های حریصانه (Greedy Algorithms)

مستندسازی: تمامی مطالب یادگیری‌شده و کدهای پیاده‌سازی شده به طور کامل داکیومنت و در مخازن GitHub شخصی ذخیره شدند.

۲ ـ ۲ ـ ۲ فاز اول: مطالعه مقالات مروری و آشنایی با حوزه **Edge Computing**

پس از تسلط بر مفاهیم پایه، وارد فاز مطالعه تخصصی در حوزه محاسبات لبه‌ای شدم. در این فاز، چهار مقاله مروری (Survey) جامع را به طور کامل مطالعه و مستندسازی کردم:

۱. **Next Generation Task Offloading Techniques in Evolving Computing Paradigms**

محتوای اصلی: این مقاله به بررسی پارادایم‌های محاسباتی نسل جدید و تکنیک‌های بارسپاری وظایف در این پارادایم‌ها می‌پردازد. این معماری را در شکل ۲ ـ ۱ می‌توانید مشاهده کنید.

پارادایم‌های معرفی شده:

(آ) رایانش مه (Fog Computing): پردازش توزیع‌شده نزدیک به دستگاه‌های IoT

(ب) رایانش لبه (Edge Computing): پردازش وظایف سبک در گره‌های لبه شبکه

(ج) رایانش میست (Mist Computing): پردازش در نزدیک‌ترین نقطه به کاربر

(د) رایانش شبنم (Dew Computing): امکان کار بدون اتصال به اینترنت

(ه) رایانش اسمزی (Osmotic Computing): جابجایی وظایف بین پارادایم‌های مختلف

(و) رایانش ترکیبی (Hybrid Computing): ترکیب مزایای پارادایم‌های مختلف

معماری پیشنهادی: معماری محاسباتی یکپارچه شش‌لایه‌ای که نحوه تعامل پارادایم‌های مختلف را نشان می‌دهد.

دستاورد: درک جامع از تفاوت‌ها و کاربردهای هر پارادایم محاسباتی

۲. **A Comprehensive Review of Computing Paradigms, Enabling Computation Offloading and Task Execution in Vehicular Networks**

محتوای اصلی: این مقاله به طور تخصصی به بررسی بارسپاری محاسبات در شبکه‌های خودرویی (Vehicular Networks) می‌پردازد.

موضوعات کلیدی:

(آ) معماری شبکه‌های خودرویی (V2V, V2I, V2X)

(ب) چالش‌های خاص شبکه‌های خودران (تحرک بالا، تغییرات سریع توپولوژی)

(ج) تکنیک‌های بارسپاری محاسباتی در خودروهای متصل

(د) الزامات QoS برای برنامه‌های ایمنی‌محور

(ه) نقش 5G و 6G در بهبود عملکرد

دستاورد: آشنایی با چالش‌های خاص دامنه خودروهای خودران

۳. **Task Offloading in Edge and Cloud Computing: A Survey on Mathematical, Artificial Intelligence and Control Theory Solutions**

محتوای اصلی: این مقاله یک طبقه‌بندی جامع از روش‌های حل مسئله بارسپاری ارائه می‌دهد.

دسته‌بندی روش‌ها:

(آ) روش‌های ریاضی:

i. برنامه‌ریزی خطی (Linear Programming)

ii. برنامه‌ریزی غیرخطی (Nonlinear Programming)

iii. بهینه‌سازی محدب (Convex Optimization)

iv. تئوری بازی (Game Theory)

v. تئوری صف (Queueing Theory)

(ب) روش‌های هوش مصنوعی:

i. یادگیری ماشین کلاسیک

ii. یادگیری عمیق (Deep Learning)

iii. یادگیری تقویتی (Reinforcement Learning)

iv. یادگیری فدرال (Federated Learning)

(ج) تئوری کنترل:

i. کنترل لیاپانوف (Lyapunov Control)

ii. کنترل بهینه (Optimal Control)

iii. کنترل پیش‌بین مدل (Model Predictive Control)

دستاورد: درک عمیق از مزایا و معایب هر دسته از روش‌ها

۴. **Mobility-aware Computational Offloading in Mobile Edge Networks**

محتوای اصلی: این مقاله به چالش تحرک کاربران در شبکه‌های لبه موبایل می‌پردازد.

مفاهیم کلیدی:

(آ) مدل‌سازی تحرک کاربران (Mobility Models)

(ب) پیش‌بینی مسیر و مقصد (Trajectory Prediction)

(ج) Handover و مهاجرت وظایف بین سرورها

(د) تعادل بین بهینگی و هزینه مهاجرت

دستاورد: درک اهمیت در نظر گرفتن تحرک در طراحی سیستم

۵. **Living on the Edge: A Survey of Digital Twin-Assisted Task Offloading in Safety-Critical Environments**

محتوای اصلی: این مقاله به بررسی ارتباط بین فناوری دوقلوی دیجیتال (Digital Twin) و بارسپاری در محیط‌های ایمنی‌محور می‌پردازد.

مفاهیم کلیدی:

(آ) دوقلوی دیجیتال: معادل مجازی یک سیستم فیزیکی برای نظارت و شبیه‌سازی

(ب) کاربردها در محیط‌های بحرانی: صنعت، حمل‌ونقل، مراقبت‌های بهداشتی

(ج) الزامات real-time monitoring و پاسخگویی سریع به ناهنجاری‌ها

(د) مسائل امنیت سایبری و حریم خصوصی

(ه) نقش 6G در پیشرفت این فناوری

دستاورد: آشنایی با کاربردهای پیشرفته و ترکیب IoT با Digital Twin

۶. **A Survey of Energy Optimization Approaches for Computational Task Offloading and Resource Allocation in MEC Networks**

محتوای اصلی: این مقاله به بررسی جامع روش‌های بهینه‌سازی انرژی در شبکه‌های MEC می‌پردازد.

موضوعات کلیدی:

(آ) معماری شبکه‌های MEC و اجزای آن

(ب) منابع محاسباتی: CPU، حافظه، ذخیره‌سازی

(ج) منابع ارتباطی: پهنای باند، توان انتقال

(د) کاربردهای بارسپاری: پهپادها (UAV)، خودروهای خودران، IIoT، AR/VR، شهرهای هوشمند

(ه) تکنیک‌های کاهش مصرف انرژی و تاخیر

(و) Trade-off بین مصرف انرژی و تاخیر

دستاورد: درک عمیق از چالش‌های انرژی و تکنیک‌های بهینه‌سازی

تمامی مقالات فوق به دقت مطالعه شده و خلاصه‌نویسی دست‌نویس از آن‌ها تهیه شد. این خلاصه‌ها شامل: اهداف و انگیزه مقاله، روش‌شناسی و رویکرد، نتایج و یافته‌های کلیدی، نقاط قوت و ضعف و ایده‌های قابل استفاده در پروژه هستند.

۲ـ۲ـ۳ فاز دوم: مطالعه و پیاده‌سازی مقالات کاربردی

پس از کسب دانش نظری جامع، وارد فاز مطالعه مقالات کاربردی و پیاده‌سازی‌شده شدم. در این فاز، دو مقاله مهم را به طور دقیق مطالعه و کدهای پیاده‌سازی آن‌ها را بررسی کردم:

۱. **HOODIE - Hybrid Computation Offloading via Distributed Deep Reinforcement Learning in Delay-aware Cloud-Edge Continuum**

رویکرد: این مقاله از یادگیری تقویتی عمیق توزیع‌شده (Distributed DRL) برای حل مسئله بارسپاری ترکیبی در یک محیط Cloud-Edge پیوسته استفاده می‌کند.

نوآوری‌های کلیدی:

(آ) استفاده از معماری Actor-Critic توزیع‌شده

(ب) در نظر گرفتن طیف پیوسته‌ای از گزینه‌ها (محلی، لبه، ابر)

(ج) مدل‌سازی تاخیر به صورت دقیق با در نظر گرفتن تمام مؤلفه‌ها

(د) یادگیری سیاست به صورت آنلاین و سازگار با تغییرات محیط

پیاده‌سازی:

(آ) کد پایتون موجود در GitHub رسمی مقاله را دانلود و مطالعه کردم

(ب) ساختار کد را تحلیل و معماری شبکه عصبی را بررسی کردم

(ج) پارامترهای مختلف را آزمایش و تأثیر آن‌ها را مشاهده کردم

دستاورد: درک عملی از نحوه پیاده‌سازی DRL در مسائل بارسپاری

۲. **QDRL - Queue-Aware Online DRL for Computation Offloading in Industrial Internet of Things**

رویکرد: این مقاله یک الگوریتم یادگیری تقویتی آنلاین با آگاهی از وضعیت صف (Queue-Aware) برای IIoT ارائه می‌دهد.

نوآوری‌های کلیدی:

(آ) مدل‌سازی صف وظایف در سرورهای لبه‌ای

(ب) یادگیری به صورت آنلاین بدون نیاز به داده‌های قبلی

(ج) تضمین پایداری صف با استفاده از تئوری Lyapunov

(د) بهینه‌سازی مشترک تصمیم بارسپاری و تخصیص منابع

پیاده‌سازی:

(آ) بررسی دقیق کد و نحوه مدل‌سازی محیط

(ب) مطالعه نحوه ترکیب تئوری صف با یادگیری تقویتی

(ج) آزمایش الگوریتم روی سناریوهای مختلف

دستاورد: یادگیری تکنیک‌های مدیریت صف و یادگیری آنلاین

به طور کلی فعالیت‌های این فاز شامل مطالعه جزئی کد و معماری هر دو پروژه، اجرای کدها با داده‌های مختلف، تحلیل نتایج و مقایسه عملکرد، شناسایی نقاط قوت و ضعف هر رویکرد و یادداشت‌برداری از ایده‌های قابل استفاده بود.

۲ـ۲ـ۴ فاز سوم: پیاده‌سازی کامل مقاله پایه

در این فاز، که مهم‌ترین بخش کارآموزی بود، به مطالعه دقیق و پیاده‌سازی کامل مقاله زیر پرداختم:

عنوان مقاله: Robust Task Offloading and Resource Allocation Under Imperfect Computing Capacity Information in Edge Intelligence Systems

خلاصه مقاله:

این مقاله به مسئله بارسپاری کامل (Full Offloading) و تخصیص منابع تحت شرایط اطلاعات ناقص از ظرفیت محاسباتی می‌پردازد. در دنیای واقعی، مقدار دقیق و توزیع آماری ظرفیت محاسباتی سرورها همیشه از پیش قابل دسترسی نیست.

برای حل این مسئله یک روش بهینه‌سازی مقاوم (Robust Optimization) برای کمینه‌سازی ETC (Energy-Time Cost) در بدترین حالت با استراتژی بارسپاری کامل پیشنهاد شده است.

متغیرهای تصمیم: مسئله به طور مشترک موارد زیر را تعیین می‌کند:

- سیاست بارسپاری (کدام دستگاه‌ها باید بارسپاری کنند)
- تخصیص پهنای باند uplink
- تخصیص فرکانس CPU در سرور MEC

البته این یک مسئله Max-Min چالش‌انگیز است که به دسته غیرمحدب (Non-Convex) تعلق دارد. حتی اگر اطلاعات ظرفیت دقیق باشد، همچنان یک مسئله MINLP (Mixed-Integer Nonlinear Programming) است که عموماً NP-Hard می‌باشد.

روش حل:

۱. روش تولید (**Generation Method**)

- فرض می‌شود سیاست بارسپاری و نسبت تقسیم داده شده است
- دو مسئله زیرمجموعه مستقل استخراج می‌شود:
- بهینه‌سازی ETC محلی
- بهینه‌سازی ETC لبه

۲. حل مسئله **ETC** محلی

- این مسئله یک LP (Linear Programming) است
- راه حل بهینه بسته (Closed-form) آن به دست می‌آید

۳. حل مسئله **ETC** لبه

- این مسئله همچنان Max-Min پیچیده است
- با استفاده از خاصیت یگانگی و تقارن، مسئله Max-Min به یک مسئله کمینه‌سازی معادل تبدیل می‌شود
- سپس آن را به یک GP (Geometric Programming) معادل تبدیل کرده

- یک الگوریتم تکراری بر اساس SCA (Successive Convex Approximation) برای یافتن نقطه ایستا پیشنهاد می‌شود

۴. بهینه‌سازی سیاست بارسپاری

- یک الگوریتم مبتنی بر Coordinate Descent (CD) برای بهینه‌سازی سیاست بارسپاری در زمان چند مرحله‌ای طراحی می‌شود

فعالیت‌های انجام شده در این بخش:

۱. مطالعه تئوری عمیق:

- مطالعه تئوری بهینه‌سازی مقاوم
- مطالعه برنامه‌ریزی هندسی (GP)
- مطالعه روش SCA و Coordinate Descent
- درک ریاضیات پشت هر مرحله از الگوریتم

۲. پیاده‌سازی:

- پیاده‌سازی مدل سیستم و محیط شبیه‌سازی
- پیاده‌سازی مدل‌های انرژی و تاخیر
- پیاده‌سازی الگوریتم حل زیرمسئله ETC محلی
- پیاده‌سازی الگوریتم حل زیرمسئله ETC لبه با CVXPY
- پیاده‌سازی الگوریتم تکراری SCA
- پیاده‌سازی الگوریتم Coordinate Descent
- پیاده‌سازی الگوریتم کلی و ادغام تمام بخش‌ها

۳. آزمایش و ارزیابی:

- تولید داده‌های شبیه‌سازی با پارامترهای مختلف

- اجرای الگوریتم و جمع‌آوری نتایج
- مقایسه با baseline methods
- تحلیل همگرایی الگوریتم
- بررسی تأثیر پارامترهای مختلف

۴. مستندسازی:

- نوشتن مستندات کامل کد
- تهیه نمودارها و جداول نتایج
- نوشتن گزارش تحلیلی

چالش‌های پیاده‌سازی:

- پیچیدگی ریاضی مسئله و نیاز به دقت در پیاده‌سازی
- دیباگ کردن الگوریتم‌های بهینه‌سازی
- تنظیم پارامترهای الگوریتم برای همگرایی بهتر
- مدیریت محدودیت‌های مختلف در مسئله بهینه‌سازی

در نهایت به پیاده‌سازی کامل و کارآمد یک سیستم بارسپاری مقاوم که قادر است در شرایط عدم قطعیت نیز عملکرد مناسبی داشته باشد دست یافتم.

۲ـ۲ـ۵ فاز چهارم: مطالعه مقالات پیشرفته برای بهبود

پس از اتمام پیاده‌سازی مقاله پایه، به دنبال بهبود و توسعه کار قبلی بودم. در این فاز، بیش از ۱۵ مقاله پیشرفته در حوزه‌های مختلف را مطالعه کردم. این مقالات در چهار دسته اصلی قرار می‌گیرند:

۱. یادگیری تقویتی چندعامله (**Multi-Agent RL**)

(آ) **Multi-Agent Reinforcement Learning for Task Offloading in Crowd-Edge Computing**

- رویکرد: استفاده از MARL برای هماهنگی بین چندین سرور لبه
- نوآوری: الگوریتم‌های مشارکتی و رقابتی بین عامل‌ها
- کاربرد: محیط‌های شلوغ با تعداد زیادی کاربر و سرور

(ب) **Multi-Agent Deep Reinforcement Learning with Trajectory Prediction for Task Migration-Assisted Computation Offloading**

- رویکرد: ترکیب پیش‌بینی مسیر با MADRL
- نوآوری: پیش‌بینی مقصد کاربران برای مهاجرت پیشگیرانه وظایف
- کاربرد: شبکه‌های موبایل با تحرک بالا
- نکته: این مقاله الهام‌بخش کار فعلی من است. می‌توانید برای کسب اطلاعات بیشتر به شکل ۲ ـ ۲ مراجعه کنید.

۲. دسته دوم: پیش‌بینی مسیر و تعادل بار

(آ) **DNN Partitioning, Task Offloading, and Resource Allocation in Dynamic Vehicular Networks: A Lyapunov-Guided Diffusion-Based Reinforcement Learning Approach**

- رویکرد: استفاده از مدل‌های Diffusion برای پیش‌بینی
- نوآوری: ترکیب تئوری Lyapunov با مدل‌های generative
- کاربرد: شبکه‌های خودرویی پویا
- نکته: مدل Diffusion در کار فعلی من استفاده می‌شود

(ب) **Distributed Resource Allocation and Coordinated Scheduling for End-Edge-Cloud Collaborative Computing**

- رویکرد: زمان‌بندی هماهنگ در معماری سه‌لایه
- نوآوری: الگوریتم توزیع‌شده برای تخصیص منابع

- کاربرد: سیستم‌های بزرگ‌مقیاس

۳. دسته سوم: وابستگی بین وظایف (**Task Dependency**)

(آ) **Dependency-Aware Task Scheduling and Layer Loading for Mobile Edge Computing Networks**

- رویکرد: مدل‌سازی وابستگی‌ها به صورت DAG (Directed Acyclic Graph)
- نوآوری: الگوریتم زمان‌بندی با در نظر گرفتن وابستگی‌ها
- کاربرد: برنامه‌های پیچیده با وظایف وابسته

(ب) **Delay-Sensitive Dependent Tasks Offloading and Resource Allocation in VEC: A Deep Reinforcement Learning Approach**

- رویکرد: DRL برای وظایف وابسته با محدودیت تاخیر
- نوآوری: ترکیب priority scheduling با DRL

(ج) **Dependent Task Scheduling and Offloading for Minimizing Deadline Violation Ratio in Mobile Edge Computing Networks**

- رویکرد: کمینه‌سازی نقض deadline
- نوآوری: الگوریتم اولویت‌بندی پویا

۴. دسته چهارم: بارسپاری جزئی پیشرفته (**Advanced Partial Offloading**)

(آ) **Partial Offloading Strategy Based on Deep Reinforcement Learning in the Internet of Vehicles**

- رویکرد: DRL برای تعیین بهینه نسبت بارسپاری
- نوآوری: یادگیری مستقیم نسبت بهینه بدون discretization
- کاربرد: اینترنت خودرویی

(ب) **Intelligent Task Offloading and Resource Allocation in Knowledge Defined Edge Computing Networks**

- رویکرد: استفاده از Knowledge Graph
- نوآوری: مدل‌سازی دانش برای تصمیم‌گیری هوشمندتر

(ج) **RADTO: A Resource-Aware and Dynamic Task Offloading Strategy for Mobile Edge Computing**

- رویکرد: استراتژی پویا با آگاهی از منابع
- نوآوری: سازگاری real-time با تغییرات منابع

۵. سایر مقالات:

(آ) **Living on the Edge: A Survey of Digital Twin-Assisted Task Offloading**

- فناوری Digital Twin و کاربردهای آن در محیط‌های ایمنی‌محور

(ب) مقالات تکمیلی:

- تکنیک‌های مهاجرت وظایف (Task migration techniques)
- زمان‌بندی با آگاهی از انرژی (Energy-aware scheduling)
- بهینه‌سازی با محدودیت QoS (QoS-constrained optimization)
- یادگیری فدرال در رایانش لبه (Federated learning in edge computing)

برای هر مقاله خلاصه‌نویسی دست‌نویس جامع، استخراج ایده‌های کلیدی، شناسایی نقاط قوت و ضعف، تحلیل قابلیت ترکیب با کار قبلی و یادداشت‌برداری از تکنیک‌های پیاده‌سازی انجام شده است. این اتفاق سبب درک عمیق از آخرین پیشرفت‌ها در حوزه، شناسایی راهکارهای نوین برای بهبود کار و دست‌یابی به ایده‌هایی برای تحقیقات آینده شد.

۲ـ۲ـ۶ فاز پنجم: توسعه و بهبود سیستم (کار جاری)

در حال حاضر، این پروژه در فاز توسعه و بهبود سیستم پیشین قرار دارد. هدف اصلی این مرحله، افزودن قابلیت‌های نوین و ارتقاء عملکرد سامانه با تکیه بر نتایج به‌دست‌آمده از مطالعات جامع در حوزه‌ی مقالات علمی روز است. توسعه‌ی سیستم در چهار محور کلیدی دنبال می‌شود که هر یک به رفع چالش‌ها و بهبود جنبه‌ای خاص از عملکرد سامانه اختصاص دارد.

۱. پیش‌بینی مقصد با مدل‌های Diffusion

در نسخه‌ی پیشین سیستم، موقعیت کاربران به‌صورت ایستا یا با استفاده از مدل‌های ساده‌ی تحرک مدل‌سازی شده بود. با این حال، در محیط‌های واقعی، پیش‌بینی دقیق مقصد کاربران نقشی اساسی در بهبود کارایی سیستم دارد. چنین قابلیتی می‌تواند تعادل بار بین سرورهای لبه‌ای را بهبود بخشد، هزینه‌های مهاجرت وظایف را کاهش دهد، امکان تصمیم‌گیری پیشگیرانه در بارسپاری را فراهم کند و تخصیص منابع را بر پایه‌ی پیش‌بینی‌های دقیق‌تر بهینه‌سازی نماید.

در پژوهش پیشین از مدل Informer مبتنی بر معماری Transformer برای پیش‌بینی مسیر کاربران استفاده شده بود. در نسخه‌ی جدید، رویکردی نوآورانه مبتنی بر مدل‌های Diffusion-based در حال پیاده‌سازی است. این مدل‌ها توانایی مدل‌سازی عدم‌قطعیت در داده‌های زمانی، تولید چندین مسیر احتمالی برای آینده، دقت بالاتر در پیش‌بینی‌های بلندمدت و امکان شرطی‌سازی بر اساس زمینه (Context) را فراهم می‌کنند.

معماری پیشنهادی در این بخش شامل مراحل زیر است: داده‌های تاریخی مسیر به عنوان ورودی به سیستم وارد می‌شوند، سپس یک کدگذار ویژگی‌های فضایی‌ـزمانی را استخراج می‌کند. مدل Diffusion بر اساس این ویژگی‌ها توزیع احتمالی مسیرهای آینده را تولید کرده و در نهایت، کدگشا محتمل‌ترین مسیر را به‌عنوان خروجی برمی‌گزیند.

در حال حاضر، مطالعات تخصصی در زمینه‌ی مدل‌های انتشار برای داده‌های سری زمانی انجام شده، نسخه‌ی اولیه‌ی مدل پیاده‌سازی گردیده، داده‌های مسیر جمع‌آوری و پیش‌پردازش شده‌اند و آزمایش‌های اولیه برای تنظیم پارامترهای مدل در حال انجام هستند.

۲. الگوریتم مدیریت صف پیشرفته

در سیستم‌های محاسبات لبه، وظایف متعددی در صف سرورهای لبه‌ای در انتظار پردازش قرار دارند و نحوه‌ی مدیریت این صف‌ها تأثیر مستقیم بر کارایی کلی سیستم دارد. چالش‌های موجود در این حوزه شامل اولویت‌بندی وظایف با مهلت‌های زمانی متفاوت (deadline)، جلوگیری از گرسنگی وظایف کم‌اولویت (starvation)، ایجاد توازن میان عدالت و کارایی، و سازگاری با بار کاری متغیر است.

برای مقابله با این چالش‌ها، رویکرد پیشنهادی ترکیبی از چند روش مختلف را به‌کار می‌گیرد: زمان‌بندی مبتنی بر اولویت با در نظر گرفتن مهلت و اهمیت وظایف، بهره‌گیری از یادگیری تقویتی (Reinforcement Learning) برای یادگیری سیاست‌های بلندمدت، و استفاده از روش‌های انطباقی جهت سازگاری با تغییرات پویا در بار کاری.

در حال حاضر طراحی معماری الگوریتم مدیریت صف تکمیل شده و فرآیند پیاده‌سازی و انجام آزمایش‌های اولیه آغاز گردیده است. همچنین مجموعه‌ای از معیارهای کمی برای ارزیابی عملکرد الگوریتم در شرایط مختلف در نظر گرفته شده است.

۳. بارسپاری جزئی پیشرفته

در سیستم پایه، نسبت بارسپاری (Offloading Ratio) از طریق روش‌های بهینه‌سازی ریاضی تعیین می‌شد. هرچند این روش‌ها از دقت بالایی برخوردارند، اما نیازمند محاسبات سنگین در هر تکرار بوده و در محیط‌های Real-Time کارایی لازم را ندارند. علاوه بر این، این روش‌ها انعطاف‌پذیری کافی در مواجهه با الگوهای پیچیده و پویا را ندارند.

در این پروژه، رویکرد جدیدی مبتنی بر یادگیری عمیق برای یادگیری مستقیم سیاست بارسپاری جزئی پیشنهاد شده است. در این معماری، وضعیت سیستم شامل بار سرورها، ویژگی‌های وظیفه، وضعیت کانال و سطح باتری به‌عنوان ورودی به یک شبکه‌ی عصبی داده می‌شود. شبکه، نسبت بهینه‌ی بارسپاری را برای هر وظیفه تخمین می‌زند که عددی پیوسته بین ۰ (اجرای کامل محلی) و ۱ (بارسپاری کامل) است.

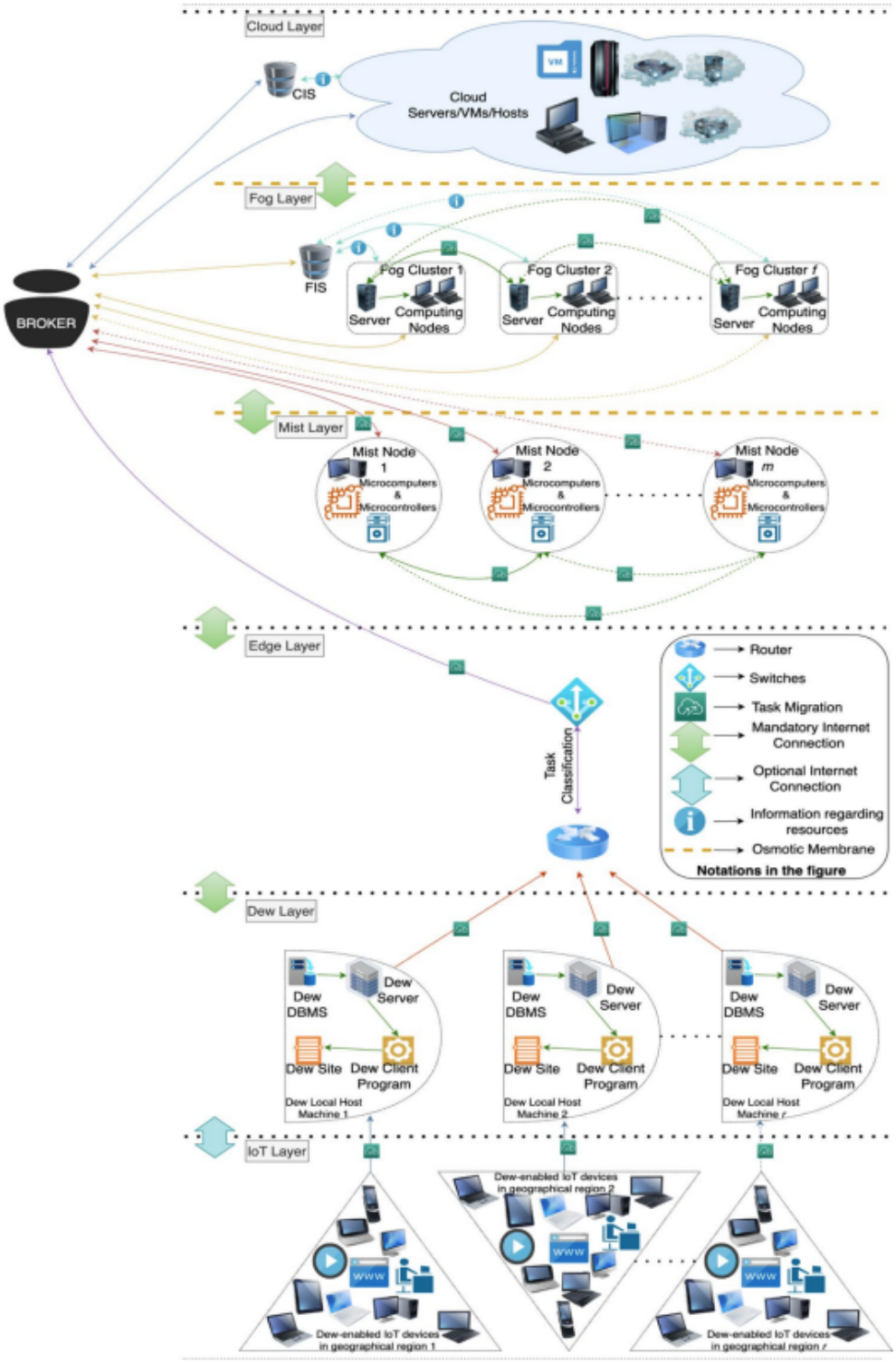
این رویکرد جدید سرعت بالایی در تصمیم‌گیری‌های لحظه‌ای دارد، با الگوهای پیچیده و غیرخطی سازگار است، قابلیت استفاده از Transfer Learning را دارد و محاسبات زمان اجرا را به میزان قابل توجهی کاهش می‌دهد. در حال حاضر مطالعات تکمیلی برای انتخاب معماری بهینه در حال انجام است و نسخه‌ی اولیه‌ی مدل در حال پیاده‌سازی می‌باشد.

۴. ادغام و آزمایش جامع سیستم

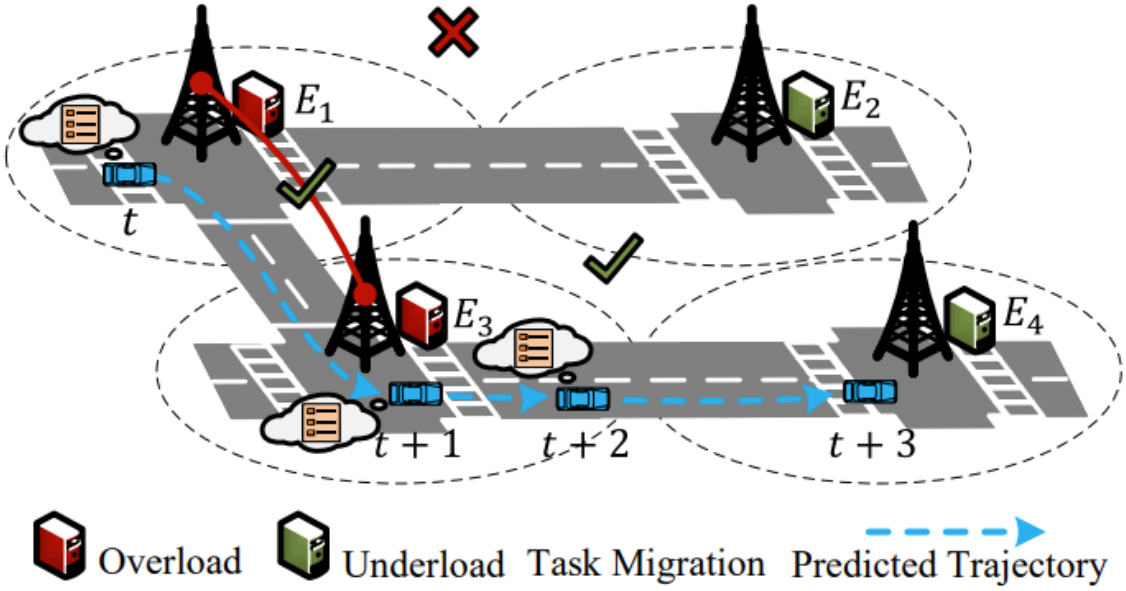
هدف نهایی پروژه، توسعه‌ی یک سامانه‌ی یکپارچه و هوشمند است که قابلیت پیش‌بینی مقصد برای تعادل بار پویا، مدیریت هوشمند صف وظایف و انجام بارسپاری جزئی بهینه را به‌صورت هم‌زمان داشته باشد. این سیستم باید در محیط‌های Real-Time کارایی بالا و پایداری مناسبی در برابر عدم‌قطعیت‌های محیطی از خود نشان دهد.

ادغام ماژول‌های مختلف سیستم چالش‌هایی از جمله هماهنگی اجزای متنوع، بهینه‌سازی سرتاسری (end-to-end)، آزمایش در سناریوهای گوناگون و مقایسه با روش‌های پیشرفته‌ی موجود را به همراه دارد.

برنامه‌ی آینده شامل تکمیل پیاده‌سازی ماژول‌ها، ادغام تدریجی آن‌ها، انجام آزمایش‌های جامع، تحلیل نتایج و در نهایت تدوین مقاله‌ی علمی بر اساس دستاوردهای حاصل خواهد بود.



شکل ۲ـ۱: معماری شبکه ۶لایه



شکل ۲ـ۲: کاربرد عملی پیش‌بینی مسیر حرکت خودرو

فصل ۳

نتیجه‌گیری

۳ـ۱ جمع‌بندی

این دوره کارآموزی، یک تجربه جامع و عمیق در حوزه محاسبات لبه‌ای، یادگیری ماشین و بهینه‌سازی برای من بود. از نظر فنی، توانستم دانش نظری گسترده‌ای در زمینه‌های مختلف کسب کنم و آن را با پیاده‌سازی‌های عملی ترکیب نمایم.

مسیر یادگیری من از پایه‌ای‌ترین مفاهیم یادگیری ماشین آغاز شد و تا پیاده‌سازی الگوریتم‌های پیچیده بهینه‌سازی مقاوم و یادگیری تقویتی عمیق پیش رفت. این مسیر به من این امکان را داد که نه تنها با ابزارها و تکنیک‌ها آشنا شوم، بلکه درک عمیقی از چرایی انتخاب هر روش و trade-offهای موجود به دست آورم.

کار در آزمایشگاه جامع و تعامل با دانشجویان و اساتید محترم، فرصتی ارزشمند برای یادگیری نحوه کار تیمی، مدیریت پروژه و ارائه نتایج بود. همچنین، مواجهه با چالش‌های واقعی پیاده‌سازی و دیباگ کردن، مهارت‌های problem-solving من را تقویت کرد.

۳ـ۲ دستاوردها

۳ـ۲ـ۱ دستاوردهای علمی

۱. تسلط بر مفاهیم پایه

در طول این دوره، تسلط جامعی بر مفاهیم اصلی حوزه یادگیری ماشین، یادگیری عمیق و یادگیری تقویتی حاصل شد.

در حوزه Machine Learning، مباحثی چون الگوریتم‌های کلاسیک، ارزیابی مدل و cross-validation به‌صورت کامل مرور و پیاده‌سازی شدند. در بخش Deep Learning، ساختار و پیاده‌سازی شبکه‌های CNN، RNN، LSTM و Transformer مورد بررسی قرار گرفت.

در زمینه Reinforcement Learning نیز، مفاهیمی چون MDP، Q-Learning، Policy Gradient، Actor-Critic، DQN و PPO به‌صورت عمیق مطالعه شدند.

همچنین، در حوزه الگوریتم‌های بهینه‌سازی، روش‌های Genetic Algorithm، PSO و Simulated Annealing پیاده‌سازی و ارزیابی شدند.

۲. تخصص در حوزه محاسبات لبه‌ای

در این دوره، درک جامعی از پارادایم‌های مختلف محاسباتی شامل Cloud، Fog، Edge، Mist، Dew و Osmotic Computing به‌دست آمد. همچنین چالش‌های بارسپاری و تخصیص منابع در سیستم‌های MEC تحلیل و مدل‌سازی شدند. تسلط بر معماری شبکه‌ها، پروتکل‌های ارتباطی و درک جریان داده‌ها از دیگر دستاوردهای این بخش بود.

۳. پیاده‌سازی عملی

پیاده‌سازی کامل مقاله "Robust Task Offloading and Resource Allocation" از مهم‌ترین بخش‌های عملی این دوره بود. در این مسیر، از فریم‌ورک‌های PyTorch، CVXPY و Scikit-learn استفاده شد. مهارت در debugging، بهینه‌سازی کد و بصری‌سازی نتایج با ابزارهای مدرن نیز تقویت گردید.

۴. مهارت‌های پژوهشی

مطالعه و تحلیل بیش از ۱۵ مقاله علمی، خلاصه‌نویسی و استخراج ایده‌های کلیدی، شناسایی خلأهای تحقیقاتی (Research Gaps) و طراحی راه‌حل‌های نوآورانه از جمله مهارت‌های پژوهشی کسب‌شده در این دوره بود.

۳ـ۲ـ۲ دستاوردهای فنی

۱. تسلط بر ابزارها

در زمینه ابزارهای فنی، تسلط بر زبان برنامه‌نویسی Python در سطح پیشرفته، فریم‌ورک یادگیری عمیق PyTorch، کتابخانه‌های بهینه‌سازی CVXPY و SciPy، کتابخانه‌های یادگیری ماشین Scikit-learn و TensorFlow، ابزارهای پردازش داده NumPy و Pandas، و ابزارهای بصری‌سازی مانند Matplotlib، Seaborn و Plotly حاصل شد. همچنین، مدیریت نسخه با Git و GitHub به‌صورت حرفه‌ای انجام شد.

۲. پیاده‌سازی الگوریتم‌های پیچیده

در این دوره، الگوریتم‌های بهینه‌سازی محدب (Convex Optimization)، الگوریتم‌های Successive Convex Approximation (SCA)، روش Coordinate Descent، الگوریتم‌های یادگیری تقویتی (DQN, PPO, Actor-Critic) و یادگیری تقویتی چندعامله (MADRL) پیاده‌سازی و تحلیل شدند.

۳. توسعه سیستم‌های پیچیده

طراحی محیط‌های شبیه‌سازی، مدیریت state و action spaceهای پیچیده، پیاده‌سازی سیستم‌های چندماژولی و یکپارچه‌سازی اجزای مختلف سیستم از دیگر دستاوردهای فنی این دوره بود.

۳ـ۲ـ۳ دستاوردهای شخصی

۱. مهارت‌های نرم

توانایی حل مسئله (problem-solving)، مدیریت زمان و اولویت‌بندی کارها، کار تیمی و انتقال مؤثر مفاهیم پیچیده از مهم‌ترین مهارت‌های نرم تقویت‌شده در این دوره بود.

۲. تفکر تحلیلی و انتقادی

تحلیل trade-offها در طراحی سیستم، ارزیابی انتقادی روش‌های مختلف و تصمیم‌گیری مبتنی بر داده‌ها به شکل قابل توجهی تقویت شدند.

۳. خودآموزی

یادگیری مستقل مفاهیم جدید، استفاده از منابع علمی معتبر و حل مشکلات فنی بدون وابستگی مستقیم از مهارت‌های ارزشمند به‌دست‌آمده بود.

۳ـ۳ چالش‌ها و پیشنهادات

۳ـ۳ـ۱ چالش‌های علمی و فنی

۱. پیچیدگی ریاضی

بسیاری از مقالات مورد مطالعه دارای فرمول‌بندی‌های پیچیده و اثبات‌های دشوار بودند. برای غلبه بر این چالش، مطالعه منابع مرجع در بهینه‌سازی محدب، مشورت با استاد راهنما و پیاده‌سازی مرحله‌به‌مرحله الگوریتم‌ها انجام شد.

۲. پیچیدگی پیاده‌سازی

پیاده‌سازی و debugging الگوریتم‌های بهینه‌سازی و یادگیری تقویتی، به‌ویژه در زمینه همگرایی و تنظیم hyperparameters، از چالش‌های جدی این دوره بود. برای رفع آن، از تکنیک‌های پیشرفته اشکال‌زدایی، بررسی دقیق کدهای مرجع و آزمایش سیستماتیک پارامترها استفاده شد.

۳ـ۳ـ۲ چالش‌های زیرساختی و مدیریتی

کمبود پیش‌زمینه علمی مناسب در برنامه درسی دانشگاه و شرایط نامناسب آزمایشگاه از چالش‌های زیرساختی بود. محدودیت فضای فیزیکی، نبود GPU و امکانات رفاهی محدود، پیشرفت پروژه را دشوار کرد. علاوه بر این، مدیریت زمان بین یادگیری، پیاده‌سازی و پژوهش نیازمند برنامه‌ریزی دقیق و پشتکار بالا بود.

۳ـ۴ پیشنهادات برای بهبود

برای دانشگاه، پیشنهاد می‌شود: افزودن دروس مرتبط با یادگیری ماشین و یادگیری عمیق، تجهیز آزمایشگاه‌ها به سرورهای دارای GPU، و برگزاری کارگاه‌های تخصصی جهت ترویج فرهنگ پژوهش. برای دانشجویان آینده، توصیه می‌شود از همان ابتدای مسیر با مبانی ریاضی و برنامه‌نویسی آشنا شده، از دوره‌های آنلاین نظیر

Coursera، edX و Stanford Online استفاده کرده و نتایج مطالعات و کدهای خود را در GitHub مستندسازی کنند. در نهایت، برای ادامه این پروژه، تمرکز بر توسعه ماژول Diffusion Model، پیاده‌سازی الگوریتم مدیریت صف پیشرفته، ادغام ماژول‌ها، و افزودن قابلیت‌هایی چون Federated Learning و پشتیبانی از سناریوهای End-Edge-Cloud پیشنهاد می‌شود.

۳ـ۵ نتیجه‌گیری نهایی

این دوره کارآموزی، تجربه‌ای فراتر از یک فعالیت آموزشی ساده بود و به سفری جامع در مسیر یادگیری و پژوهش تبدیل شد. از نظر علمی، تسلط بر مفاهیم یادگیری ماشین، یادگیری عمیق و یادگیری تقویتی و درک چالش‌های واقعی حوزه Edge Computing از مهم‌ترین دستاوردها بود. از نظر فنی، مهارت در برنامه‌نویسی، پیاده‌سازی سیستم‌های پیچیده و کار با ابزارهای مدرن به شکل چشمگیری ارتقا یافت. از نظر شخصی، توانایی حل مسئله، صبر، پشتکار و تفکر تحلیلی به‌طور قابل ملاحظه‌ای رشد یافت.

این دوره نه‌تنها موجب افزایش دانش و مهارت شد، بلکه دیدگاهی عمیق‌تر نسبت به دنیای پژوهش و فناوری ایجاد کرد و انگیزه‌ی ادامه مسیر آکادمیک در این حوزه را تقویت نمود.

پیوست آ

جدول کامل مقالات مطالعه شده

توضیح: خلاصهٔ دست‌نویس کامل تمامی مقالات فوق موجود می‌باشد.

جدول آـ۱: مقالات مروری (Survey Papers)

ردیف	عنوان مقاله	حوزه اصلی	سال	نکات کلیدی
۱	Next Generation Task Offloading Techniques in Evolving Computing Paradigms	پارادایم‌های محاسباتی	–	شش‌لایه‌ای، معماری پارادایم‌ها طبقه‌بندی
۲	A Comprehensive Review of Computing Paradigms in Vehicular Networks	شبکه‌های خودرویی	–	چالش‌های QoS، V2X خودران خودروهای در
۳	Task Offloading in Edge and Cloud Computing: A Survey on Mathematical, AI and Control Theory Solutions	حل روش‌های	–	روش‌های طبقه‌بندی کنترل و AI ریاضی،
۴	Offloading Using Traditional Optimization and ML in Federated Systems	و بهینه‌سازی ML	–	و سنتی روش‌های مقایسه ماشین یادگیری
۵	Mobility-aware Computational Offloading in Mobile Edge Networks	و تحرک مهاجرت	–	تحرک، مدل‌سازی Handover
۶	Living on the Edge: Digital Twin-Assisted Task Offloading	Digital Twin	–	ایمنی‌محور، کاربردهای سایبری امنیت
۷	A Survey of Energy Optimization Approaches in MEC Networks	بهینه‌سازی انرژی	–	MEC، معماری تکنیک‌های کاربردها، بهینه‌سازی

جدول آـ۲: مقالات پیاده‌سازی اولیه

ردیف	عنوان مقاله	روش اصلی	نوآوری / وضعیت مطالعه
۸	HOODIE: Hybrid Computation Offloading via Distributed DRL	Distributed DRL	Actor-Critic توزیع‌شده، کد بررسی + مطالعه
۹	QDRL: Queue-Aware Online DRL for IIoT	Queue-aware DRL	با صف تئوری ترکیب کد بررسی + مطالعه ،DRL

جدول آـ۳: مقاله پایه (Base Paper)

ردیف	عنوان مقاله	روش اصلی	نوآوری / وضعیت
۱۰	Robust Task Offloading Under Imperfect Computing Capacity Information	Robust Optimization	مقاوم، بهینه‌سازی پیاده‌سازی ،CD ،SCA □ کامل

جدول آـ۴: مقالات یادگیری تقویتی چندعامله

ردیف	عنوان مقاله	تکنیک اصلی	کاربرد / نکات مهم
۱۱	Multi-Agent RL for Task Offloading in Crowd-Edge Computing	MARL	شلوغ، محیط‌های مشارکتی الگوریتم‌های رقابتی و
۱۲	MADRL with Trajectory Prediction for Task Migration	MADRL + Prediction	پیش‌بینی موبایل، شبکه‌های مهاجرت برای مسیر پیشگیرانه

جدول آـ۵: مقالات پیش‌بینی مسیر و تعادل بار

ردیف	عنوان مقاله	مدل پیش‌بینی / ویژگی	استفاده در پروژه
۱۳	DNN Partitioning with Lyapunov-Guided Diffusion-Based RL	Diffusion Model، با Lyapunov ترکیب Diffusion	□ فعلی کار الهام‌بخش
۱۴	Distributed Resource Allocation for End-Edge-Cloud	Distributed Algorithm، معماری سه‌لایه	load balancing مفاهیم

جدول آـ۶: مقالات وابستگی بین وظایف

ردیف	عنوان مقاله	مدل وابستگی / الگوریتم	کاربرد
۱۵	Dependency-Aware Task Scheduling and Layer Loading	DAG + Priority Scheduling	پیچیده برنامه‌های
۱۶	Delay-Sensitive Dependent Tasks in VEC	DAG + DRL + Priority	خودرویی شبکه‌های
۱۷	Dependent Task Scheduling for Minimizing Deadline Violation	DAG + Dynamic Priority	deadline تضمین

جدول آـ۷: مقالات بارسپاری جزئی پیشرفته

ردیف	عنوان مقاله	روش تعیین نسبت / نوآوری	کاربرد
۱۸	Partial Offloading with DRL in IoV	DRL، مستقیم یادگیری نسبت	خودرویی اینترنت
۱۹	Intelligent Task Offloading in Knowledge Defined Networks	Knowledge Graph، دانش از استفاده	هوشمند تصمیم‌گیری
۲۰	RADTO: Resource-Aware Dynamic Task Offloading	Adaptive Algorithm، سازگاری Real-Time	پویا محیط‌های

پیوست ب

منابع و مراجع

ب‌ـ۱ دوره‌های آموزشی

- دوره Machine Learning ـ دکتر مهدیه سلیمانی و دکتر محمد شریفی‌زارچی، دانشگاه صنعتی شریف
- دوره Deep Learning ـ دکتر مهدیه سلیمانی، دانشگاه صنعتی شریف
- دوره Reinforcement Learning ـ دکتر محمدرضا رهبان، دانشگاه صنعتی شریف

ب‌ـ۲ ابزارها و فریمورک‌ها

- PyTorch ـ https://pytorch.org/
- CVXPY ـ https://www.cvxpy.org/
- Scikit-learn ـ https://scikit-learn.org/
- NumPy ـ https://numpy.org/
- Pandas ـ https://pandas.pydata.org/
- Matplotlib ـ https://matplotlib.org/

ب_۳ وب‌سایت‌ها و منابع آنلاین

- arXiv.org ـ برای دسترسی به جدیدترین مقالات
- gigalib.org ـ پایگاه دانلود مقاله‌های رایگان
- IEEE Xplore ـ پایگاه داده مقالات IEEE
- ACM Digital Library ـ پایگاه داده مقالات ACM
- Google Scholar ـ جستجوی مقالات علمی
- GitHub ـ اشتراک‌گذاری کد و پروژه‌ها