

## کاهش تداخل در شبکه‌های بی‌سیم ناهمگن LTE با تنظیم مقدار $\epsilon$ در الگوریتم $\epsilon$ -greedy کاهش با استفاده از تکنیک eICIC

شهبو صالحی<sup>۱</sup>، بهدیس اسلام‌نور<sup>۲</sup>

<sup>۱</sup> دانشجوی دکتری مهندسی کامپیوتر گرایش شبکه‌های کامپیوتری، دانشگاه ارومیه، shavbo.salehi@urmia.ac.ir

<sup>۲</sup> استادیار دانشکده برق و کامپیوتر، دانشگاه ارومیه، b.eslamnour@urmia.ac.ir

### چکیده

امروزه با افزایش ارتباطات، نیاز به بهره‌گیری از دستگاه‌های بی‌سیم بیشتر احساس می‌شود که موجب افزایش تعداد دستگاه‌های متصل به شبکه، حجم ترافیک و رقابت بر سر منابع مورد نیاز شده است. به همین دلیل، با هدف کاهش پردازش داده‌ها و ترافیک ایستگاه‌های ماکروسل (Macrocell)، از ایستگاه‌های پیکوسل (Picocell) استفاده می‌شود. اما توان ارسال ایستگاه‌های پیکوسل در مقایسه با ایستگاه‌های ماکروسل پایین‌تر و در نتیجه احتمال اتصال به ایستگاه‌های دارای توان ارسال بالاتر می‌یابد. به همین دلیل، تکنیک تقویت هماهنگی تداخل بین سلولی (eICIC) با دو متغیر بایاس گسترش دامنه سلولی (CRE) و نسبت زیرفریم خالی تقریبی (ABS) برای افزایش استفاده کاربران از ایستگاه‌های پیکوسل معرفی شده است. تنظیم مقادیر این دو متغیر موجب کاهش ترافیک ایستگاه ماکروسل و بهبود گذردهی و کیفیت ارتباطات کاربران می‌شود. در این مقاله از روش راهزن چند دست (MAB) برای انتخاب بهینه متغیرهای CRE و ABS تکنیک eICIC استفاده شده است. همچنین تاثیر گذردهی شبکه با بهره‌گیری از الگوریتم  $\epsilon$ -greedy کاهش با هدف محاسبه مقدار مناسب  $\epsilon$  بررسی شده است. نتایج شبیه‌سازی‌ها نشان می‌دهند که در مقایسه با روش‌های کران بالای اطمینان (UCB) و  $\epsilon$ -greedy لگاریتمی استفاده از الگوریتم  $\epsilon$ -greedy کاهش با مقدار بهینه  $\epsilon$  موجب بهبود ۴۵٪ گذردهی شبکه می‌شود.

### کلیدواژه

تداخل شبکه‌های ناهمگن، بایاس گسترش دامنه سلولی، نسبت زیرفریم خالی، تکنیک eICIC الگوریتم  $\epsilon$ -greedy کاهش.

### مقدمه

پس از معرفی شبکه‌های 4G که با نام تجاری تحول طولانی مدت<sup>۱</sup> (LTE) معرفی شدند، تعداد کاربرانی که از خدمات شبکه استفاده می‌کردند، افزایش یافت. به منظور برآورده کردن نیازهای روزافزون، پروژه مشارکت نسل سوم<sup>۲</sup> (3GPP) در نسخه ۱۰ ویژگی‌های پیشرفته LTE را با عنوان شبکه‌های تحول طولانی مدت پیشرفته<sup>۳</sup> (LTE-A) معرفی کرد. LTE-A شامل ارتباطات چند نقطه‌ای تعاونی<sup>۴</sup> (Comp)، چند ورودی چند خروجی<sup>۵</sup> (MIMO) پیشرفته و هماهنگی تداخل بین سلولی<sup>۶</sup> (ICIC) است که با هدف افزایش سرعت ارتباطات و کاهش تاخیر معرفی شدند. همچنین هزینه‌های چشمگیر تغییرات زیرساختی وسیع و به تعویق افتادن جایگزینی فراگیر شبکه‌های نسل‌های گذشته با شبکه‌های 4G موجب شد تا

معرفی و بهره‌گیری از شبکه‌های بی‌سیم در دهه‌ی ۱۹۸۰ میلادی با نسل اول ارتباطات (1G) آغاز شد و اکنون در برخی از مناطق جهان، نسل پنجم ارتباطات (5G) استفاده می‌شود، که این استفاده از فناوری ارتباطات با اتصال آغاز و به سمت یک شبکه‌ی کاملاً متصل و پیچیده سوق یافته است [1]. از شبکه موبایل آنالوگ تا LTE، هر نسل از فناوری‌ها در پاسخ به کاستی‌های موجود در نسل‌های قبل تعریف شده است، که به عنوان مثال عبور از نسل اول به دوم با ظهور ارتباطات دیجیتال و گذر از نسل سوم به چهارم با افزایش سرعت دسترسی به داده‌ها و کاهش زمان تاخیر همراه بود [2]. امروزه نسل پنجم شبکه‌های موبایل در برخی شهرها به بهره‌برداری‌های آزمایشی رسیده است، با این وجود در بسیاری از کشورهای جهان شبکه‌های نسل چهارم (4G) مورد استفاده قرار می‌گیرد و مطالعات و تحقیقات برای توسعه، گسترش استفاده و بهبود آنها در راستای نیازمندی‌های کنونی در حال تغییر و پیشبرد است.

<sup>1</sup> Long-Term Evolution

<sup>2</sup> 3rd Generation Partnership Project

<sup>3</sup> LTE-Advanced

<sup>4</sup> Cooperative Multi-Point

<sup>5</sup> Multiple Input, Multiple Output

<sup>6</sup> Inter-Cell Interference Coordination

بین سمبل<sup>۱۵</sup> (ISI)، محوشدگی و غیره مواجه هستند [۱۴]. این چالش‌ها ارائه سرویس‌های یکپارچه [۵]، کیفیت ارتباطات موبایل و رضایت کاربران [۶] را که یکی از مهم‌ترین نیازهای حوزه‌های در حال رشد ارتباطات موبایل است را تحت تاثیر قرار می‌دهد، زیرا تداخل ایستگاه‌های مجاور در سلول‌های ماکروسل بین  $-85dBm$  تا  $-95dBm$  است [20]. از سوی دیگر، ICIC یک چالش بزرگ برای استقرار HetNet، به ویژه برای کاربرانی که در لبه سلول‌ها قرار دارند، را به همراه دارد زیرا این کاربران بیشتر در برابر سیگنال‌های تداخلی سایر سلول‌ها قرار دارند. برای حل مشکل ICIC، 3GPP برخی از فناوری‌های مبتنی بر تسهیم منابع مانند مالتی‌پلکسینگ تقسیم زمانی<sup>۱۶</sup> (TDM)، مالتی‌پلکسینگ تقسیم فرکانسی<sup>۱۷</sup> (FDM) را ارائه داده‌است. به دنبال عدم توانایی روش ICIC در رفع مشکلات تداخل در شبکه‌های ناهمگن، تکنیک هماهنگی بهبود یافته تداخل بین سلولی<sup>۱۸</sup> (eICIC) در نسخه ۱۰ ارائه شد. این روش دارای دو متغیر بایاس CRE و زیرفریم تقریباً خالی<sup>۱۹</sup> (ABS) است. با تنظیم این دو متغیر، با انتقال بخشی از ترافیک به پیکوسل، میزان ترافیک ماکروسل کاهش پیدا می‌کند. تنظیم بهینه مقادیر این دو متغیر با توجه به شرایط شبکه مانند میزان ترافیک، موقعیت مکانی کاربران و تعداد ایستگاه‌های پیکوسل [۷] صورت می‌گیرد. تکنیک eICIC نیز مبتنی بر روش TDM است و به تعدادی از زیرفریم‌های ایستگاه‌های ماکروسل که توان آنها صفر و یا کاهش داده شده و احتمالاً با دستگاه‌های کاربران تداخل ایجاد می‌کند، اجازه می‌دهد که به سلول‌های کوچک‌تر اختصاص یابد. پس از معرفی نسخه ۱۰، در نسخه ۱۱ تکنیک ICIC پیشرفته‌تر<sup>۲۰</sup> (FeICIC) بر اساس عملکرد فرستنده یا گیرنده نشان داد که راه‌حل موثری که در نسخه ۱۰ به آن پرداخته شد، به دلیل تغییرات پیوسته در شرایط شبکه-های بی‌سیم، با چالش‌های جدی مواجه است. به عنوان مثال، استفاده کاربران از منابع شبکه، جهت و سرعت حرکت آنها و شرایط محیطی همواره در حال تغییر است و تداخل ایجاد شده توسط کاربران یا سایر سلول‌ها موجب کاهش کیفیت این ارتباطات می‌شود. از این رو استفاده از روش‌های مبتنی بر یادگیری تقویتی مورد توجه قرار گرفته است زیرا در این روش-ها عملکرد صحیح قطعی برای عامل تعیین نمی‌شود بلکه تنها با توجه به بازخوردها، عملکرد انتخابی عامل ارزیابی می‌شود. در نتیجه عامل با توجه به بازخوردهای دریافتی اقدامات بعدی

مفهوم شبکه‌های ناهمگن<sup>۷</sup> (HetNet) در LTE-A ارائه شود. HetNetها شامل سلول‌هایی با ابعاد و ایستگاه‌های پایه گوناگون مانند ماکروسل، فمتوسل، میکروسل و پیکوسل برای افزایش ظرفیت شبکه، بهره‌برداری بیشتر از طیف فرکانسی و همچنین ادغام فناوری‌های دسترسی مختلف هستند [۳]. با توجه به گستردگی شبکه‌های بی‌سیم، انتظار می‌رود که در شبکه‌های 5G نیز [۱۹] HetNetها استفاده شوند، زیرا استفاده از تکنولوژی‌های 5G نیازمند تغییر ایستگاه‌های پایه و دستگاه‌های سمت کاربران خواهد بود که به دلیل هزینه‌های زیرساختی، بهره‌گیری از تکنیک HetNet دور از انتظار نیست. همچنین اپراتورها امکان بهره‌برداری از فناوری‌های گوناگون را در منطقه تحت پوشش خود فراهم می‌آورند و در مقایسه با شبکه‌های همگن، سرویس‌های یکپارچه‌تر و سازگارتر را به کاربران ارائه می‌دهند. علاوه بر این، طراحی شبکه‌هایی با سلول‌های کوچک، انعطاف‌پذیری شبکه را افزایش می‌دهد [۴].

به طور کلی در شبکه‌های سلولی از ایستگاه‌های ماکروسل استفاده می‌شود، اما در مناطق شهری یا پرتراکم، اپراتورها از ایستگاه‌های پیکوسل با توان ارسالی کمتری استفاده می‌کنند تا با انتقال بخشی از ترافیک سلول‌های ماکروسل به پیکوسل ظرفیت شبکه را بهبود بخشند. اما، دستگاه‌های موبایل تمایل دارند به ایستگاه‌هایی با توان ارسالی بالاتر متصل شوند و افزودن ایستگاه پیکوسل موجب بهبود کیفیت سرویس نمی‌شود. به همین دلیل، مقدار گسترش دامنه سلولی<sup>۸</sup> (CRE) به توان ارسالی ایستگاه‌های سلول‌های کوچک اضافه می‌شود تا ترافیک بیشتری را از ایستگاه‌های ماکروسل به ایستگاه‌های کوچک‌تر تخلیه<sup>۹</sup> کند. اما بهره‌گیری از این روش تداخل بیشتری را به همراه دارد و برای رفع آن از ابتدای استانداردسازی LTE، 3GPP تمرکز بالایی بر روی ICIC داشته است. روش ICIC در نسخه ۱۰ از استانداردهای 3GPP ارائه شد که استفاده از آن در شبکه‌های همگن<sup>۱۱</sup> مناسب ارزیابی شده است، اما در معماری شبکه‌های ناهمگن سود قابل توجهی ارائه نمی‌کند، زیرا گسترش دامنه را محدود می‌کند. همچنین این روش تنها بر روی کانال‌های داده اعمال می‌شود اما توانایی کنترل کانال‌هایی که تداخل در آنها قابل توجه است را ندارد. علاوه بر این، HetNetها با چالش‌های متعددی مانند دست‌به‌دست شدگی<sup>۱۲</sup>، محدودیت منابع، انواع تداخل شامل تداخل هم‌کانال<sup>۱۳</sup> (CCI)، تداخل کانال مجاور<sup>۱۴</sup> (ACI)، تداخل

<sup>7</sup> Heterogeneous Network

<sup>8</sup> Cell Range Expansion

<sup>9</sup> Offload

<sup>10</sup> Release 9

<sup>11</sup> Homogeneous Networks

<sup>12</sup> Handoff

<sup>13</sup> Co-Channel Interference

<sup>14</sup> Adjacent Channel Interference

<sup>15</sup> Inter-Symbol Interference

<sup>16</sup> Time Division Multiplexing

<sup>17</sup> Frequency Division Multiplexing

<sup>18</sup> enhanced Inter-cell Interference Coordination

<sup>19</sup> Almost Blank Subframe

<sup>20</sup> Further enhanced ICIC

ارائه نمی‌دهد. اما DDPG با استفاده از softmax یک لایه خروجی خواهد داشت، که شبکه یاد می‌گیرد احتمال بالایی را برای عملی که پاداش را به حداکثر می‌رساند تنظیم کند. البته در مدل DDPG تنظیم متغیرهای مدل پیچیده است و این روش به زمان بسیار بیشتری نسبت به روش DQN نیاز دارد زیرا تعداد متغیرهای آن بسیار بیشتر است.

با توجه به گسترش استفاده از HetNetها، رویکردهای ارائه شده در شبکه‌ها و سیاست‌های اتخاذ شده توسط اپراتورها برای بهبود کیفیت تجربه کاربران در هنگام بهره‌گیری از شبکه‌ها، نیاز به ارائه یک روش با توانایی تصمیم‌گیری برخط برای شبکه‌های ناهمگن دارد. همچنین به منظور ارزیابی رویکرد ارائه شده، طراحی سناریو باید در شرایط نزدیک به واقعیت شبیه‌سازی شود. به همین دلیل، در این مقاله مقدار بهینه‌ی متغیرهای تکنیک eICIC در شرایط نزدیک به واقعیت با در نظر گرفتن سیستم‌های ایستا و پویا ارزیابی می‌شود. برای این منظور، عملکرد شبکه‌ها را در زمان واقعی<sup>۲۸</sup> با بهره‌گیری از مسئله‌ی راهزن چند دست<sup>۲۹</sup> (MAB) توسط الگوریتم greedy-ε کاهش ارزیابی می‌کنیم. برای بهبود گذردهی شبکه توسط الگوریتم ε-greedy کاهش، برای اولین بار مقادیر بهینه‌ی ε را در سناریوهایی مختلف با در نظر گرفتن پدیده‌های افت مسیر، تداخل و محوشدگی محاسبه و سپس تاثیر این مقادیر بهینه را در سناریوهایی مجزا شبیه‌سازی و بررسی می‌کنیم. نتایج شبیه‌سازی‌ها نشان می‌دهد که بهره‌گیری از مقادیر بهینه ε در روش ε-greedy کاهش نسبت به سایر روش‌های مسئله‌ی MAB بطور متوسط بهبود ۴۵ درصدی گذردهی HetNetها می‌شود. همچنین در سیستم شبیه‌سازی شده، رویکرد ارائه شده نه تنها CCI را کاهش می‌دهد، بلکه توان عملیاتی، کارایی و ظرفیت سیستم نیز افزایش می‌یابد و کیفیت تجربه<sup>۳۰</sup> (QoE) کاربران با توجه به اندازه‌گیری‌های نسبت سیگنال به تداخل و نویز<sup>۳۱</sup> (SINR) بهبود می‌یابد.

ساختار این مقاله به این شرح است که در بخش دوم، خلاصه‌ای از پژوهش‌های گذشته برای کاهش تداخل و بهبود گذردهی که برای HetNetها ارائه شده‌اند را بررسی می‌کنیم، سپس در بخش سوم، مشکلات و موانع موجود در شبکه‌های بی‌سیم و HetNet بررسی می‌شوند و پس از ارزیابی متغیرهای تاثیرگذار بر کیفیت ارتباطات شبکه‌های موبایل، نحوه‌ی ارزیابی آن‌ها ارائه می‌شود. همچنین به صورت خلاصه روش‌های ارزیابی QoE در این بخش بررسی می‌شود. در بخش چهارم فرموله کردن سیستم و روش ارائه شده در این مقاله برای بهبود QoE

خود را انتخاب می‌کند که به مرور موجب بهبود اقدامات انتخاب شده و افزایش پاداش عامل می‌شود.

الگوریتم‌های یادگیری تقویتی در سه دسته‌بندی عملگر تنها<sup>۲۱</sup>، منتقد تنها<sup>۲۲</sup> و عملگر-منتقد<sup>۲۳</sup> تقسیم‌بندی می‌شود. که در این میان روش عملگر-منتقد به دلیل همگرایی سریع‌تر، بیشتر از دو دسته‌بندی دیگر استفاده می‌شود [29]. اما جمع‌آوری اطلاعات مربوط به تابع ارزش<sup>۲۴</sup> بسیار زمان‌بر است. به همین دلیل شبکه Q عمیق<sup>۲۵</sup> (DQN) معرفی شد که مشکل همگرایی را رفع و با سرعت آموزش بالاتر، مشکل زمان بر بودن این روش را رفع کرده است. اما این روش نیز مقدار مورد استفاده در تابع Q را بیش از حد واقعی تخمین می‌زند. با این وجود در [30] از روش عملگر-منتقد عمیق برای مشکل کنترل توان انتقال در شبکه‌های بی‌سیم استفاده شده است. در این روش هر فرستنده به عنوان یک عامل یادگیری مجزا در نظر گرفته شده است که سطح توان انتقال خود را با مشاهده محیط به صورت محلی تعیین می‌کند. همچنین برای رفع مشکلات DQN، روش دوئل کردن شبکه Q عمیق<sup>۲۶</sup> (DDQN) معرفی شد. اما روش DDQN در مسائلی که فضای تعداد اقدامات بسیار متعدد باشد، عملکرد قابل قبولی ارائه نمی‌کند. به همین دلیل استفاده از آن در مسائل مدیریت تداخل برای سناریوهای شهری یا وسیع که کاربران بسیاری از شبکه استفاده می‌کنند و همچنین تعداد آنتن‌های استفاده شده بسیار متعدد و فشرده است، مناسب نیست.

روش DQN یک روش یادگیری مبتنی بر ارزش است، در حالی که روش گرادیان سیاست قطعی عمیق<sup>۲۷</sup> DDPG یک روش عملگر-منتقد است. شبکه DQN مقادیر Q را برای هر جفت حالت-عمل پیش‌بینی می‌کند، بنابراین فقط یک مدل واحد برای هر دو بخش در نظر گرفته می‌شود. اما DDPG یک مدل انتقادی دارد که مقدار Q را تعیین می‌کند اما از مدل عملگر برای تعیین اقدامی که باید انجام شود استفاده می‌کند. از این رو روش DDPG تلاش می‌کند مستقیماً خط مشی را بیاموزد؛ در حالی که DQN مقادیر Q را که برای تعریف خط مشی استفاده می‌شود، را به صورت یک خط مشی حریمانه پس‌یونی می‌آموزد. به همین دلیل زمانی که فضای عمل پیوسته است، DQN عملکرد قابل قبولی ندارد، زیرا سعی می‌کند برای هر عمل، مقدار حالت-عمل ایجاد کند. یکی از رویکردهای رایج برای حل مسئله، گسسته کردن فضای عمل است و ممکن است در برخی از شرایط مناسب باشد، اما راه حل ایده‌آل را

<sup>21</sup> Actor-Only

<sup>22</sup> Critic-Only

<sup>23</sup> Actor-Critic

<sup>24</sup> Value Function

<sup>25</sup> Deep-Q network

<sup>26</sup> Dueling Deep Q Networks

<sup>27</sup> Deep Deterministic Policy Gradient

<sup>28</sup> Real Time

<sup>29</sup> Multi-Armed Bandit

<sup>30</sup> Quality of Experience

<sup>31</sup> Signal-to-Interference-plus-Noise Ratio

تخصیص منابع تعلق گرفته است. همچنین، از روش‌های مدیریت تداخل eICIC و FeICIC با هدف کاهش تداخل در استفاده همزمان شبکه‌های LTE امنیت عمومی<sup>37</sup> (PS-LTE) در محیط اشتراک‌گذاری RAN استفاده شده است که موجب کاهش تداخل در شبکه می‌شود.

در سال‌های اخیر محققان بسیاری به بررسی و کنترل تداخل در HetNet با بهره‌گیری از روش eICIC پرداخته‌اند، که در برخی از تحقیقات تنها تاثیر مقدار یکی از متغیرها بررسی شده است ولی برخی دیگر از محققان مقدار هر دو متغیر را محاسبه کرده‌اند. در [9] رویکرد Q-Learning فازی برای محاسبه مقدار بهینه‌ی متغیر ABS ارائه شد. رومرو و همکارانش [8] در سال ۲۰۱۷ روش تصادفی بهینه‌سازی پاسخ سطحی<sup>38</sup> (RSM) را برای تعیین مقادیر هر دو متغیر eICIC در شبکه‌هایی با ترافیک و توزیع متفاوت کاربران مطالعه کردند، که در روش RSM مقادیر تقریبی محلی تابع هدف برای بهبود سطح کیفیت سرویس مورد نیاز استفاده می‌شود. همچنین بهره‌گیری از روش‌های مبتنی بر یادگیری نیز مورد توجه محققان بوده است. رومرو و همکارانش [7] یک سال بعد از روش مبتنی بر داده-های عظیم<sup>39</sup> برای تخمین مقادیر متغیرهای eICIC استفاده کردند و گذردهی شبکه را با دو روش مسئله‌ی MAB و الگوریتم گرادیان افزایشی تصادفی<sup>40</sup> (SGA) ارزیابی کردند. همچنین به منظور کاهش تداخل در HetNet‌ها از روش مبتنی بر داده تنظیم‌کننده<sup>41</sup> eICIC (DDeC) با استفاده از یادگیری عمیق استفاده شده است. در روش DDeC، سیستم با یک مجموعه داده آموزش داده می‌شود و سپس با استفاده از آموخته‌ها، در مورد مجموعه داده جدید تصمیم گرفته می‌شود. در روش‌های مبتنی بر یادگیری عمیق (همچنین یادگیری تقویتی عمیق) [27] خروجی سیستم با افزایش تعداد نمونه-های مجموعه داده، افزایش می‌یابد اما در شبکه‌های موبایل، به دلیل محرمانه بودن اطلاعات، مجموعه داده‌های دارای برچسب با کیفیت بالا و در مقیاس بزرگ توسط اپراتورهای شبکه ارائه نشده است که این موضوع توسعه و استفاده از یادگیری عمیق و یادگیری تقویتی عمیق را محدود کرده است. همچنین وابستگی بسیار زیاد روش‌های مبتنی بر یادگیری عمیق به داده‌های اولیه موجب شده است که در بسیاری از موارد تعداد، توزیع و نوع داده‌ها مهم‌تر از مدل مورد استفاده باشد که موجب افزایش زمان محاسبات و مصرف منابع انرژی می‌شود. به همین دلیل در این مقاله از روش MAB که یکی از مسئله‌های یادگیری تقویتی است، برای بهبود QoE شبکه‌های بی‌سیم

و انتخاب متغیرهای تکنیک eICIC بررسی خواهد شد. بخش پنجم نتایج سناریوهای شبیه‌سازی شده مختلف بررسی و تاثیر روش پیشنهادی بر HetNet مورد بررسی در سه سناریوی ایستا و یک سناریوی پویا با روش‌هایی دیگر مقایسه می‌شود.

## مروری بر کارهای گذشته

امروزه با توسعه‌ی سریع برنامه‌های کاربردی بی‌سیم و افزایش تقاضای کاربران برای دسترسی به سرویس‌های باکیفیت بالا، شبکه‌هایی با فناوری‌های نسل‌های 2G یا 3G به تنهایی قادر به ارائه سرویس‌های باکیفیت قابل قبول نیستند. به همین دلیل شبکه‌های 4G با هدف ایجاد ناهمگنی در فناوری شبکه‌های ارائه‌دهنده سرویس از طریق ادغام شبکه‌هایی با فناوری‌های متفاوت معرفی شدند. از سوی دیگر این HetNet‌ها با چالش-های متعددی از قبیل انتخاب نوع شبکه از بین شبکه‌های مختلف در دسترس، دست‌به‌دست شدگی، محدودیت منابع، محوشدگی و تداخل مواجه هستند. از این رو در سال‌های اخیر حل مشکلات این شبکه‌ها مانند کنترل میزان تداخل در شبکه-های LTE-A با بهره‌گیری از تکنیک‌های جدید و متفاوت مورد توجه بسیاری از محققان بوده است. در [۱۵] مکانیزم مدیریت منابع رادیویی<sup>32</sup> (RRM) برای گسترش HetNet‌ها با بهره‌گیری از روش استفاده مجدد کسری از فرکانس با آنتن‌های سه جهته و سه لایه<sup>33</sup> (FFR-3SL) ارائه شد. در روش FFR-3SL، استفاده صحیح از منابع رادیویی موجب کاهش اثرات تداخل CCI در LTE-A می‌شود و توزیع باندهای فرکانسی بین ایستگاه‌های فمتوسل و ماکروسل عملکرد سیستم را بهبود می‌بخشد. رحیم و همکاران در [۱۶] دو نوع تداخل ایستگاه‌های فمتوسل و ماکروسل و تداخل بین ایستگاه‌های فمتوسل را در شبکه‌های LTE بررسی کردند و برای جلوگیری از تداخل بین ایستگاه-های فمتوسل، توان ارسالی ایستگاه‌ها<sup>34</sup> ( $P_t$ ) را محدود کردند و همچنین تکنیک‌های برنامه‌ریزی پوشش‌دهی و استفاده مجدد از فرکانس مورد بحث قرار گرفت. نتایج آنها نشان می‌دهد که کنترل  $P_t$  و بهینه‌سازی توزیع ایستگاه‌های پایه موجب کاهش تداخل در شبکه‌های LTE شده است. این کاهش تداخل، احتمال قطع شدن کاربران در حال حرکت را کاهش می‌دهد و کیفیت ارتباطات را نیز بهبود می‌دهد. در [۱۷] عملکرد شبکه-های LTE با در نظر گرفتن موارد اشتراک شبکه دسترسی رادیویی<sup>35</sup> (non-RAN) برای سرویس‌هایی با ماموریت بحرانی<sup>36</sup> (MCS) (قابلیت اطمینان بالا و تأخیر کم) و سرویس‌های عادی ارزیابی شده است و به کاربران MCS اولویت بالایی برای

<sup>37</sup> Public Safety LTE

<sup>38</sup> Response Surface Methodology

<sup>39</sup> Big Data

<sup>40</sup> Stochastic Gradient Ascent

<sup>41</sup> Data-Driven eICIC Configurator

<sup>32</sup> Radio Resource Management

<sup>33</sup> Fractional Frequency Reuse with Three Sectors and Three Layers

<sup>34</sup> Transmission Power

<sup>35</sup> Radio Access Network

<sup>36</sup> Mission-Critical Service

توان مختلف دریافت می‌کنند و برای انتخاب ایستگاه پایه‌ی خود براساس قدرت مرجع سیگنال دریافت شده<sup>۴۲</sup> (RSRP) یا کیفیت سیگنال مرجع<sup>۴۳</sup> (RSRQ) تصمیم‌گیری می‌کنند [۱۲] و تمایل دارند به ایستگاه با بالاترین مقدار RSRP متصل می‌شوند. به طور کلی RSRP سیگنال ایستگاه‌هایی ماکروسل بیشتر از ایستگاه‌های پیکوسل است و کاربران شبکه‌های بی‌سیم و حتی کاربرانی که در محدوده‌ی پوشش ایستگاه‌های پیکوسل قرار گرفته‌اند، با احتمال بیشتری به ایستگاه ماکروسل متصل می‌شوند. علاوه بر این، با پیشرفت شبکه‌های بی‌سیم و معرفی HetNetها، این شبکه‌ها با چالش‌های دیگری مانند دست‌به‌دست شدگی، انواع تداخل‌های ISI، ACI، CCI، محوشدگی، افت مسیر، چند مسیری<sup>۴۴</sup> و انعکاس<sup>۴۵</sup> نیز مواجه هستند.

به طور کلی، تداخل سیگنال‌های ایستگاه‌های مختلف یکی از مهم‌ترین چالش‌های HetNetها است، زیرا موجب تضعیف یا تغییر سیگنال‌های دریافتی می‌شود که به منظور ارزیابی تاثیر این تداخل‌ها، از ویژگی‌های آماری نسبت سیگنال بر نویز و تداخل (SINR) دریافت شده کاربران استفاده می‌شود. SINR معیاری است که نسبت توان سیگنال دریافتی کاربر را با در نظر گرفتن نویز و تداخل سایر ایستگاه‌ها محاسبه می‌کند. سیگنال‌ها تنها در صورتی قابل آشکارسازی هستند که مقدار SINR دریافت شده کاربران بیش از حد آستانه تعیین شده باشد. رابطه ۱ نحوه‌ی محاسبه‌ی SINR دریافت شده کاربرانی را نشان می‌دهد که به صورت مستقیم به ایستگاه ماکروسل (M) در زیرحامل<sup>۴۶</sup>  $n$  متصل شده‌اند [۱۶]:

$$SINR_{M,n} = \frac{P_d^M G_{M,n} \omega d^{-\eta}}{\sum_{i=1, i \neq M}^{M'} P_d^{M'} G_{M',n} + \sum_{j=1}^{P'} P_d^P G_{P,n} + P_{noise}^{\Delta f}} \quad (1)$$

که  $P_d^M$  توان ارسالی ایستگاه ماکروسل است که کاربران به آن متصل هستند و  $P_d^{M'}$  توان ارسالی سایر ایستگاه‌های ماکروسل روی زیرحامل  $n$  را نشان می‌دهد که کاربر در فاصله‌ی  $d$  از آنها قرار دارد و موجب بروز تداخل بر ارتباطات می‌شوند و تعداد آنها از یک تا  $M'$  با توجه به اندازه‌ی شبکه‌ها متغیر است.  $G_{M,n}$  مقدار بهره کانال<sup>۴۷</sup> زیرحامل  $n$  را نشان می‌دهد که کاربر از طریق آن به ایستگاه ماکروسل وصل شده است و  $G_{P,n}$  و  $G_{M',n}$  به ترتیب بهره سایر کانال‌های ایستگاه‌های ماکروسل و پیکوسل روی زیرحامل  $n$  را نشان می‌دهند.  $\omega$  مقدار افت مسیر و  $\eta$

استفاده شده است. این روند تصمیم‌گیری برای دست‌یابی به تعادل بین کشف مسیرهای جدید و بهره‌برداری از دانش گذشته [۱۰] در فرآیندهای متناهی استفاده می‌شود و با استفاده نتایج خروجی‌ها، یادگیری صورت می‌گیرد [۱۱]. عامل با استفاده از روش MAB به صورت برخط تصمیم‌گیری می‌کنند و بیشینه‌کردن مقدار پاداش‌ها با در نظر گرفتن مجموعه‌ی اقدامات جاری است؛ که هر اقدام، پاداش مجزا دارد اما این پاداش‌ها در ابتدا ناشناخته و تصادفی هستند.

## مدل سیستم

امروزه در شبکه‌های سلولی از ایستگاه‌های ماکروسل استفاده می‌شود، اما در بیشتر مناطق شهری، تراکم جمعیت و تنوع سرویس‌های درخواست شده موجب شده است که اپراتورها از ایستگاه‌های پیکوسل که نسبت به ایستگاه‌های ماکروسل توان ارسالی کمتری دارند استفاده کنند. این بهره‌گیری از انواع مختلف ایستگاه‌های فرکانسی موجب ناهمگن شدن شبکه‌ها می‌شود. در شبکه‌های ناهمگن، HetNetها، به دلیل همپوشانی سلول‌های ماکروسل و پیکوسل، کاربران همچنان در غالب موارد از ایستگاه ماکروسل استفاده می‌کنند، زیرا تمایل دارند که خدمات مورد نیاز خود را از ایستگاه‌هایی با توان ارسالی بالاتر دریافت کنند. به همین دلیل نسبت کاربرانی که به ایستگاه‌های پیکوسل متصل می‌شوند، نسبت به کاربران ایستگاه‌های ماکروسل به صورت چشمگیری کاهش می‌یابد. اما با هدف کاهش بار ایستگاه‌های ماکروسل، کاهش تداخل در HetNetها و افزایش QoE کاربران از روش‌های مختلفی استفاده می‌شود که یکی از موفق‌ترین روش‌ها، تکنیک eICIC بوده است. این تکنیک دارای دو متغیر CRE و ABS است و با تنظیم کردن این دو متغیر میزان ترافیک ایستگاه‌های ماکروسل کاهش و استفاده از ایستگاه‌های پیکوسل در شبکه افزایش پیدا می‌کند.

## چالش‌های شبکه‌های LTE-A

اپراتورهای شبکه‌های بی‌سیم، طبق استانداردهایی از پیش تعریف شده از ایستگاه‌های ماکروسل استفاده می‌کنند. اما تراکم جمعیت و به دنبال آن تنوع سرویس‌های درخواستی و نیاز به سرویس‌های با کیفیت بالاتر موجب می‌شود که اپراتورها برای برآورده کردن این نیازمندی‌ها و همچنین کاهش بار ایستگاه‌های ماکروسل‌ها، در محدوده‌ی آنها از ایستگاه‌هایی مانند پیکوسل، فمتوسل یا میکروسل که در مقایسه با ماکروسل توان ارسالی کمتری دارند، استفاده می‌کنند. اما سیگنال‌های ارسال شده ماکروسل موجب بروز تداخل در سیگنال‌های ارسال شده ایستگاه‌ها می‌شود. همچنین کاربران شبکه سیگنال‌هایی با

<sup>42</sup> Received Signal Reference Power

<sup>43</sup> Reference Signal Received Quality

<sup>44</sup> Multipath

<sup>45</sup> Reflections

<sup>46</sup> Subcarrier

<sup>47</sup> Channel Gain

و به صورت تابعی از فاصله بین فرستنده و گیرنده و همچنین اثر سایه محیط اطراف تعریف می‌شود. واریانس سایه‌افکنی نرمال براساس اندازه‌گیری در طیف وسیعی از مکان‌هایی که دارای اختلاف محیطی متفاوت هستند، محاسبه می‌شود. در این مقاله به دلیل کوتاه بودن فاصله بین گیرنده‌ها و فرستنده‌ها، تاثیر سایه‌افکنی ناچیز و از اثر آن چشم‌پوشی شده است. اما اثر محوشدگی در مقیاس کوچک بررسی شده است و همچنین محوشدگی سریع با توزیع پواسون در نظر گرفته شده است، که مقادیر آن در حدود  $\pm 3dBm$  تغییر می‌کند [25].

عوامل بسیاری بر کیفیت شبکه‌های موبایل تاثیرگذار هستند، اما تعداد قابل توجهی از این عوامل دارای اثر تصادفی هستند که پیش‌بینی مقادیر دقیق آنها به دلیل وابستگی به عوامل متغیر و متحرک بیرونی دشوار است. به همین دلیل کیفیت شبکه‌ها براساس میزان QoE ارزیابی می‌شود. اما QoE یک متغیر نسبی است و نحوه‌ی محاسبه‌ی آن با در نظر گرفتن نوع شبکه، کاربران و سرویس مورد تقاضای آنها و منابع در دسترس شبکه متغیر است. به همین دلیل، در شبکه‌ها معیارهای متفاوتی معرفی و ارزیابی می‌شود که در بخش بعدی ابتدا به صورت خلاصه روش‌های محاسبه‌ی QoE معرفی و سپس روش مورد استفاده در این مقاله معرفی خواهد شد.

#### ارزیابی QoE در شبکه‌های LTE-A

در بستر شبکه‌ها، کانال‌های ارتباطی متفاوتی برای ایجاد ارتباط بین کاربران و سرویس‌دهندگان وجود دارد. همچنین کاربران از سرویس‌های متفاوتی استفاده می‌کنند که این سرویس‌ها براساس معیارهای کیفیت سرویس<sup>53</sup> (QoS) شامل تاخیر، لرزش<sup>54</sup>، توان عملیاتی و تلفات بسته ارزیابی می‌شوند. با تعریف یک رابطه میان نوع سرویس مورد تقاضای کاربر و نیازمندی‌های QoS آن سرویس، به ازای انواع مختلف سرویس‌ها و تخصیص کانال متناسب با نوع نیازمندی‌های سرویس، انتظار می‌رود که QoE کاربران بهبود یابد، زیرا جنبه‌های فنی یک سرویس توسط QoS نشان داده می‌شود. اما QoS تجربه واقعی کاربر برای سرویس استفاده شده را منعکس نمی‌کند و برای توصیفی کامل از سرویس ارائه شده ناکافی است. سرویس‌های ارائه شده توسط اپراتورها با سطح یکسان از QoS، تضمین‌کننده‌ی QoE یکسان در کاربران مختلف نیست. زیرا به غیر از ویژگی‌های فنی سیستم، عواملی دیگری مانند شبکه و ایستگاه‌های مجاور، ویژگی‌های جمعیت‌شناسی، محتوای ارائه شده و قیمت‌گذاری یک سرویس تاثیر قابل توجهی بر QoE کاربران دارند. جنبه‌های فوق توسط QoS در

مقدار توان افت مسیر را نشان می‌دهد که در محیط‌های شهری تقریباً مقدار 3 [26] در نظر گرفته می‌شود. همچنین  $P_{noise}$  توان نویز سفید<sup>48</sup> و  $\Delta f$  نیز فاصله‌ی زیر حامل‌ها را نشان می‌دهد.

علاوه بر این، سایر ایستگاه‌های ماکروسل و پیکوسل موجب تداخل بر سیگنال‌های دریافتی کاربران ایستگاه پیکوسل می‌شوند. بنابراین SINR دریافت شده در کاربر پیکوسل روی زیرحامل  $n$  به صورت زیر نشان داده می‌شود [16]:

$$SINR_{P,n} = \frac{P_d^P C_{loss}}{\sum_{i=1}^M P_d^M G_{P,M,n} + \sum_{j=1, j \neq P}^{P'} P_d^P G_{P,M,n} + P_{noise}^{\Delta f}} \quad (2)$$

و ظرفیت ایستگاه ماکروسل، در یک فضای زیرحامل خاص  $n$  از طریق SINR توسط رابطه زیر محاسبه می‌شود [15]:

$$C_{M,n} = BW \cdot \log_2^{(1+\phi SINR_{M,n})} \quad (3)$$

که  $BW$  نشان‌دهنده‌ی پهنای باند در دسترس برای هر یک از کاربرانی است که به صورت مشترک از زیرحامل  $n$  استفاده می‌کنند. همچنین در این معادله  $\phi$  یک عدد ثابت برای نشان دادن نرخ خطای بیت<sup>49</sup> (BER) است.

توان سیگنال‌های ارسال شده ایستگاه ماکروسل یا پیکوسل با دور شدن از فرستنده کاهش می‌یابد به همین دلیل تاثیر مقدار  $\omega$  به عنوان یکی دیگر از عوامل تاثیرگذار بر کیفیت سیگنال کاربران در نظر گرفته می‌شود. رابطه‌ی 4 نشان دهنده‌ی توان سیگنال دریافتی کاربری است که در فاصله  $d$  از ایستگاه ماکروسل قرار دارد [15]:

$$\omega = \frac{P_d^M \left(\frac{d}{d_0}\right)^{-\eta}}{P_{noise}} \quad (4)$$

که  $d_0$  محدوده‌ی دور<sup>50</sup> و برابر با یک متر است. علاوه بر تداخل و افت مسیر، در کانال‌های پخش رادیویی، دو نوع محوشدگی در مقیاس بزرگ و کوچک نیز متصور است. محوشدگی در مقیاس کوچک در محدوده طول موج سیگنال رخ می‌دهد و در مقایسه با محوشدگی در مقیاس بزرگ، نسبتاً تصادفی است. محوشدگی در مقیاس کوچک، عمدتاً نتیجه‌ی چندمسیری است که سیگنال اصلی در هر لحظه تحت تاثیر تداخل سازنده یا مخرب قرار می‌گیرد. محوشدگی سیگنال در مقیاس بزرگ ناشی از افت مسیر در فاصله‌های زیاد و اثر سایه<sup>51</sup> است و به خوبی با مدل نرمال<sup>52</sup> نمایش داده می‌شود. محوشدگی مقیاس بزرگ به میزان افت مسیر،  $\omega$ ، بستگی دارد

<sup>48</sup> White Noise Power

<sup>49</sup> Bit Error Rate

<sup>50</sup> Far Field

<sup>51</sup> Shadowing Effects

<sup>52</sup> Log-Normal Models

<sup>53</sup> Quality of Service

<sup>54</sup> Jitter

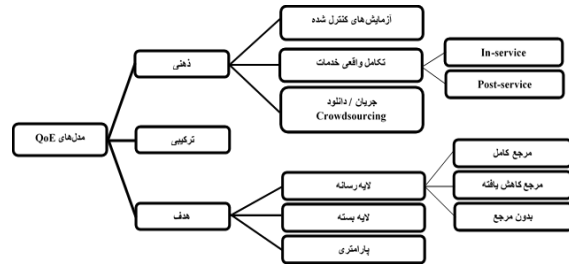
همان‌طور که پیشتر اشاره شد، اپراتورها به دلیل نیازمندی‌های مختلف کاربران و همچنین افزایش ترافیک شبکه‌ها، در برخی مناطق شهری از تعدادی پیکوسل نیز استفاده می‌کنند. اما علاوه بر تفاوت میزان RSRP ایستگاه‌های ماکروسل و پیکوسل، عوامل محیطی نیز موجب می‌شوند که کاربران از ایستگاه‌های ماکروسل استفاده نکنند. اما در این شبکه‌ها با هدف بهبود QoE و کاهش ترافیک ایستگاه‌های ماکروسل از تکنیک‌هایی برای ترغیب کاربران به منظور استفاده از ایستگاه‌های پیکوسل استفاده می‌شود که در بخش بعدی استفاده از تکنیک eICIC، برای هدف مذکور بررسی می‌شود.

### روش پیشنهادی

با ظهور HetNet‌ها، رویکردهای بسیاری برای ترغیب کاربران به منظور استفاده از ایستگاه‌های پیکوسل ارائه شده است. یکی از این رویکردها، تکنیک eICIC است که با به‌کارگیری آن در هنگام برقراری ارتباط و اتصال به ایستگاه پایه، کاربران ایستگاه‌های ماکروسل در برخی از فریم‌های زمانی به ایستگاه‌های پیکوسل متصل می‌شوند. در این تکنیک، توسط دو متغیر CRE و ABS در بازه‌های زمانی مشخصی، برخی از فریم‌ها به ایستگاه‌های پیکوسل اختصاص پیدا می‌کند و تنها سیگنال‌های ایستگاه‌های پیکوسل در دسترس قرار دارد. با این حال همچنان تعداد بیشتری از فریم‌ها به ایستگاه‌های ماکروسل اختصاص داده می‌شود و به این ترتیب ترافیک ایستگاه ماکروسل کاهش پیدا می‌کند. به همین دلیل، انتخاب درست و بهینه متغیرهای CRE و ABS موجب بهبود گذردهی شبکه و QoE کاربران می‌شود.

در HetNet‌ها با هدف بهبود کیفیت شبکه و افزایش استفاده از شبکه، تاثیر متغیرهای CRE و ABS با معیار استاندارد کارایی قطعی  $F$  که وابسته به متغیرهای هماهنگی تداخل است، ارزیابی می‌شود [7]. در این مقاله به منظور سهولت در بیان مطالب، متغیرهای بایاس CRE و نسبت ABS به ترتیب با  $\gamma$  و  $\emptyset$  نشان داده شده‌اند و هر پیکربندی  $x$  در شبکه به صورت  $x = (\gamma, \emptyset)$  فرض می‌شود. زمانی که شرایط شبکه (شدت ترافیک، تعداد ایستگاه‌های پیکوسل فعال در شبکه و غیره) ثابت است، تابع  $F(x)$  یک متغیر تصادفی است که وابسته به شرایط تصادفی شبکه از قبیل مکان کاربران، تداخل، پدیده‌های سایه‌افکنی، افت مسیر و محوشدگی تصادفی است. علاوه بر این، مقدار  $F$  وابسته به متغیرهای لایه‌ی فیزیکی مانند SINR و یا متغیرهای لایه‌های بالاتر مانند گذردهی کاربر است. در این مقاله تخمین حداکثر احتمال<sup>57</sup> (MLE) پیکربندی  $x$  در تابع  $F(x)$  با  $E[F(x)]$  نشان داده می‌شود.

نظر گرفته نمی‌شوند و این امر موجب می‌شود که اثر فاکتورهای فنی QoS روی QoE یک رابطه غیرخطی باشد. به همین دلیل در حرکت به سمت ارتباطات نسل جدید، مفهوم QoE به تصویرکشیدن درک کاربران نهایی از یک سرویس ارائه شده است [24]. همانند QoS، QoE در مکانیزم شبکه و به طور خاص در فرآیندهای تصمیم‌گیری شبکه گنجانیده شده است و در شبکه با روش‌های اندازه‌گیری مستقیم توسط انسان، اندازه‌گیری خودکار از طریق عوامل فنی و حالت ترکیبی قابل اندازه‌گیری است. در شکل ۱ طبقه‌بندی مدل‌های ارزیابی QoE را به صورت گرافیکی نشان می‌دهد.



شکل ۱: طبقه‌بندی رویکردهای مدل‌سازی QoE [21]

در روش‌های مبتنی بر آزمون ذهنی، نتایج براساس آزمایش‌های انجام‌گرفته در زندگی واقعی ارائه می‌شود. در این روش‌ها از شرکت‌کنندگان خواسته می‌شود تا ارزیابی و تجربه خود را در مورد سرویس‌هایی که استفاده کرده‌اند منعکس کنند. اما هزینه بالا، وقت‌گیر بودن و برحسب تقاضا بودن موجب می‌شود تا بازتولید این روش‌ها دشوار باشد. در نتیجه مدل‌های عینی با استفاده از معیارهای مختلف از قبیل سیگنال‌های مرجع، مدل حالت، مدل بازه زمانی و غیره معرفی شدند [21]. در [22] نشان داده شده است که ویژگی‌های آماری SINR می‌تواند معیار قابل قبولی برای ارزیابی مدل‌های عینی شبکه‌های بی‌سیم باشد.

در این مقاله برای ارزیابی QoE کاربران از مقدار توان عملیاتی استفاده شده است. براساس رابطه‌ی شانون گسترش‌یافته<sup>55</sup>، در محاسبه‌ی توان عملیاتی با در نظر گرفتن ضریب تصحیح، می‌توان از مقدار SINR استفاده کرد. به همین دلیل [23]:

$$C_{M,n} = \psi \cdot BW \cdot \log_2^{(1+\phi \text{SINR}_{M,n})} \quad (5)$$

که ضریب  $\psi$  با توجه به نوع سیستم تغییر می‌کند. همچنین یک سیستم تسهیم فضایی<sup>56</sup> MIMO  $3 \times 3$  که  $N_t = 3$  و  $N_r = 3$  به ترتیب تعداد آنتن‌های فرستنده و گیرنده را نشان می‌دهند در نظر گرفته شده است. مقدار توان عملیاتی با توجه به اندازه‌ی ضریب  $\psi$  افزایش می‌یابد، که  $\psi = N_r = N_t$ . در این مقاله  $C_{M,n}$  برای ارزیابی QoE در نظر گرفته شده است، زیرا مقدار تاخیر نیز مرتبط با مقدار  $C_{M,n}$  است.

<sup>57</sup> Maximum Likelihood Estimation

<sup>55</sup> Extended Shannon

<sup>56</sup> Spatial Multiplexing System

## فرموله کردن مسئله براساس مسئله MAB

تغییر است به همین دلیل روش‌های مبتنی بر یادگیری در بهبود کیفیت این شبکه‌ها موثر هستند. این روند تصمیم‌گیری برای دست‌یابی به تعادل بین کشف مسیرهای جدید و بهره‌برداری از دانش گذشته [۱۰] در فرآیندهای متناهی استفاده می‌شود و با استفاده از نتایج خروجی‌ها، یادگیری صورت می‌گیرد [۱۱]. هدف مسئله MAB بیشینه کردن مقدار پاداش‌ها با در نظر گرفتن مجموعه اقدامات جاری است که هر اقدام پاداش مجزا دارد اما این پاداش‌ها در ابتدا ناشناخته و تصادفی هستند. مسئله MAB مانند سایر الگوریتم‌های مبتنی بر یادگیری به صورت برخط تصمیم‌گیری می‌کند، به همین دلیل برای ارزیابی مشکل مدیریت تداخل، مسئله یادگیری به صورت برخط فرموله می‌شود.

در مسئله MAB فرض می‌کنیم که  $A = \{1, 2, \dots, a\}$  مجموعه بازوهای در نظر گرفته شده باشد و زمان به تعدادی گام قابل شمارش به صورت  $k = 1, 2, 3, \dots$  تقسیم می‌شود، که دنباله دوره‌های تصمیم‌گیری را نشان می‌دهد. در هر گام  $k$  تصمیم‌گیرنده یکی از بازوهای  $A$  را انتخاب و پاداش  $X_k^a \in [0, 1]$  را بدست می‌آورد. اثر اقدام انتخاب شده به صورت  $y_k = F(x_k)$  نشان داده می‌شود. به همین ترتیب، مقدار  $x_{k+1}$  براساس دانش بدست آمده از نمونه‌های قبلی  $x_0, \dots, x_{k-1}, x_k$  محاسبه می‌شود. نمونه‌های انتخاب شده ی قبلی  $x_0, \dots, x_{k-1}, x_k$  نیز از مجموعه  $\rho$  انتخاب می‌شوند که  $\rho = \{x = (\gamma, \emptyset) : 0 \leq \gamma \leq 1, 0 \leq \emptyset \leq \bar{\rho}\}$  بیشترین مقدار  $\square$  را نشان می‌دهد [۷]. مقدار امید ریاضی متغیر  $F(x)$  در گام  $k$  ام این الگوریتم به صورت زیر تعریف می‌شود [۷]:

$$kE[F(x^*)] - \sum_{i=1}^k E[F(x_i)], \quad (6)$$

که

$$x^* = \arg \max_{x \in \rho} E[F(x)],$$

رابطه‌ی فوق نشان می‌دهد که این الگوریتم در شرایط عدم قطعیت به چه صورت در مورد انتخاب متغیرهای تکنیک eICIC تصمیم‌گیری می‌کند.

همان‌گونه که پیشتر بیان شد، ایستگاه‌های پیکوسل برای افزایش استفاده مجدد از فرکانس‌ها در شبکه‌ها معرفی شده‌اند، اما  $P_t(\text{macro}) > P_t(\text{pico})$  موجب می‌شود که کاربران به ایستگاه‌های پیکوسل متصل نشوند و تعداد بیشتری از کاربران به ایستگاه ماکروسل وصل شوند. برای جلوگیری از این مشکل مقادیر متغیرهای eICIC در شرایط مختلف شبکه توسط روش MAB انتخاب و به موجب آن، کاربر در شرایطی که مقدار RSRP پیکوسل پایین‌تر از مقدار RSRP ماکروسل باشد، به ایستگاه پیکوسل متصل شود.

در سال‌های اخیر، بررسی و مطالعه‌ی روش‌های مبتنی بر یادگیری عمیق [27] و یادگیری تقویتی [28] به منظور کاهش تداخل و بهبود QoE کاربران شبکه‌های موبایل به صورت چشمگیری افزایش پیدا کرده است. در روش یادگیری عمیق، سیستم با یک مجموعه داده، آموزش یافته، و سپس با استفاده از آن آموزه‌ها، در مورد مجموعه داده جدید تصمیم می‌گیرد. در حالی که در یادگیری تقویتی مقادیر پارامترها براساس بازخوردهای پیوسته برای به حداکثر رساندن خروجی شبکه تنظیم می‌شوند و یادگیری به صورت پویا ادامه می‌یابد. بهره‌گیری از روش‌هایی مانند یادگیری عمیق و یادگیری تقویتی عمیق در حوزه‌ی شبکه‌های موبایل با چالش‌هایی جدی مواجه است. به عنوان مثال خروجی روش‌های مبتنی بر یادگیری عمیق (همچنین یادگیری تقویتی عمیق) [27] با افزایش تعداد نمونه‌های مجموعه داده بهبود می‌یابد اما برخلاف سایر زمینه‌های تحقیقاتی مانند بینایی ماشین<sup>58</sup> (MV) و پردازش زبان طبیعی<sup>59</sup> (NLP)، به دلیل حفظ محرمانگی اطلاعات توسط اپراتورهای شبکه، مجموعه داده‌های دارای برچسب با کیفیت بالا و در مقیاس بزرگ برای شبکه‌های موبایل ارائه نشده است. اگرچه این تصمیم از نظر حفظ حریم خصوصی کاربران منطقی است، اما توسعه سازوکارهای یادگیری عمیق و یادگیری تقویتی عمیق را محدود کرده است. داده‌های شبکه‌های موبایل جمع‌آوری شده توسط حسگرها و تجهیزات شبکه نیز اغلب در معرض اتلاف<sup>60</sup>، افزونگی و برچسب‌گذاری نامناسب قرار دارند، بنابراین نمی‌توان از آنها برای آموزش شبکه استفاده کرد زیرا الگوریتم‌های یادگیری عمیق در برابر فعالیتهای خصمانه بسیار آسیب‌پذیر هستند. همچنین وابستگی بسیار زیاد روش‌های مبتنی بر یادگیری عمیق به داده‌های اولیه موجب شده است که در بسیاری از موارد تعداد داده‌ها مهم‌تر از مدل مورد استفاده باشد. با افزایش تعداد داده‌ها میزان محاسبات به صورت چشمگیری افزایش می‌یابد و محاسبات موازی و پردازنده‌های قوی مورد نیاز خواهد بود. این پردازنده‌های قوی به منبع انرژی نیاز دارند، و در بسیاری از موارد محدودیت منبع انرژی و عدم دسترسی به منابع وجود دارد. این روش‌ها به پارامترهای بسیاری نیاز دارند که تعیین این مقادیر موجب پیچیده‌تر شدن استفاده از این روش‌ها شده است. به همین دلیل در این مقاله از روش MAB که یکی از مسئله‌های یادگیری تقویتی است، برای بهبود QoE شبکه‌های بی‌سیم استفاده شده است، زیرا شرایط شبکه به صورت پیوسته در حال

<sup>58</sup> Machine vision<sup>59</sup> Natural Language Processing<sup>60</sup> Loss

همگرا شود.	
<p><b>مزایا:</b></p> <p>✓ این الگوریتم همواره به سمت مقدار حداکثر جهانی همگرا می‌شود.</p> <p><b>معایب:</b></p> <p>✓ به نمونه‌های بسیاری برای همگرایی نیاز دارد تا نتایج به سمت حداکثر جهانی همگرا شود.</p> <p>✓ بهره‌وری این الگوریتم کمتر از بهره‌وری الگوریتم <math>\epsilon</math>-greedy است، بنابراین پاداش تجمعی کمتری برای کاربرها خواهد داشت.</p> <p>✓ پیاده‌سازی آن دشوارتر است.</p>	<p><b>نمونه برداری و فلسفه‌ی خوش‌بینی تامپسون</b></p>
<p><b>مزایا:</b></p> <p>✓ عملکردهای موفق بیشترین محبوبیت را دارند در حالی که عملکردهای ضعیف‌تر از کمترین محبوبیت برخوردار خواهند بود.</p> <p>✓ هنگامی که داده‌های تجمعی به اندازه کافی وجود داشته باشد، الگوریتم تقریباً همیشه از آن استفاده می‌کند.</p> <p>✓ به تغییرات با بهترین عملکرد سریع واکنش نشان می‌دهد.</p> <p><b>معایب:</b></p> <p>✓ با توجه به نحوه اجرای الگوریتم، درصد توزیع یک تغییر ممکن است به ۰ برسد.</p> <p>✓ عملکرد کلی آن کمتر از روش نمونه برداری تامپسون است.</p> <p>✓ پیاده‌سازی آن دشوار است.</p>	<p><b>کران بالای اطمینان</b></p>

علاوه بر مزایا و معایب هر کدام از روش‌ها، پیچیدگی زمانی نیز اهمیت قابل توجهی دارد که در جدول ۲ مقادیر مربوط به پیچیدگی زمانی این روش‌ها با دیگر مقایسه شده است.

جدول ۲: پیچیدگی محاسباتی الگوریتم‌های MAB

پیچیدگی محاسباتی	نوع الگوریتم
$O(TK)$	$\epsilon$ -greedy
$O(\sqrt{TK} \ln T)$	نمونه برداری و فلسفه‌ی خوش‌بینی تامپسون
$O(2TKe^{-2a/25})$	کران بالای اطمینان

که مقادیر  $T$  و  $K$  به ترتیب تعداد دورها و تعداد اقدامات را نشان می‌دهد. همچنین  $0 < a \leq \frac{25T-K}{36H_1}$  که  $H_1$  معیاری از سختی برای مسئله‌ی مورد نظر است.

در این مقاله با توجه به بررسی‌های حاصل از نقاط ضعف و قوت روش‌های ارائه شده و همچنین پیچیدگی زمانی که در جدول ۱ و جدول ۲ ارائه شده است، الگوریتم  $\epsilon$ -greedy انتخاب شده است. این الگوریتم دارای چند زیر الگوریتم مانند  $\epsilon$  ثابت، وفقی و کاهش‌ی است که در این مقاله از الگوریتم  $\epsilon$ -greedy کاهش‌ی استفاده شده است. یکی از مسائل مهم در بکارگیری الگوریتم  $\epsilon$ -greedy و همچنین نمونه کاهش‌ی آن، انتخاب مقادیر بهینه  $\epsilon$  است. در این مقاله مقادیر بهینه الگوریتم  $\epsilon$ -greedy کاهش‌ی

رویکرد MAB از استراتژی‌های متعددی برای یادگیری و انتخاب اقدام بهینه استفاده می‌کند که در این مقاله از الگوریتم  $\epsilon$ -greedy کاهش‌ی استفاده شده است. همچنین در بخش بعدی برخی از استراتژی‌های این الگوریتم معرفی و نحوه‌ی ارزیابی  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهش‌ی بررسی خواهد شد. در فصل سنجش عملکرد نیز با ارزیابی تاثیر مقادیر محاسبه شده، با بهره‌گیری از مسئله‌ی MAB مقادیر متغیرهای تکنیک eICIC برای ارزیابی تاثیر مقدار بهینه‌ی  $\epsilon$  در کاهش تداخل و بهبود گذرده‌ی بررسی خواهد شد.

### روش تعیین متغیرهای تکنیک eICIC

مسئله‌ی یادگیری تقویتی MAB بدون در اختیار داشتن دانش اولیه و همچنین بازخورد ابتدایی تصمیم‌گیری می‌کند، به همین دلیل رویکرد تصمیم‌گیری براساس تعادل بین کشف مسیرهای جدید و بهره‌برداری از دانش گذشته است. به طور کلی، هدف روش MAB بیشینه‌کردن مقدار پاداش براساس اقدامات متفاوت انجام شده است. برای این منظور، الگوریتم‌های مختلفی مانند الگوریتم شاخص ژتون‌ها<sup>۶۱</sup>،  $\epsilon$ -greedy، کران بالای اطمینان<sup>۶۲</sup> (UCB)، ارزش دانش و پاداش تخمینی<sup>۶۳</sup> (POKER)، نمونه برداری و فلسفه‌ی خوش‌بینی تامپسون<sup>۶۴</sup> و بهترین میانگین نمونه‌های تجربی<sup>۶۵</sup> (BESA) پیشنهاد شده است [۱۳]. در این مقاله از الگوریتم  $\epsilon$ -greedy برای تعیین مقادیر متغیرهای تکنیک eICIC استفاده شده است. که در جدول ۱ مزایا و معایب روش پیشنهادی در این مقاله و همچنین هر کدام از این الگوریتم‌های استفاده شده بررسی می‌شوند. با توجه به نتایج ارائه شده در این جدول، الگوریتم‌هایی مانند نمونه برداری و فلسفه‌ی خوش‌بینی تامپسون و کران بالای اطمینان دارای عملکرد ضعیف‌تری نسبت به الگوریتم  $\epsilon$ -greedy هستند.

جدول ۱: مزایا و معایب الگوریتم‌ها

الگوریتم	مزایا و معایب
$\epsilon$ -greedy	<p><b>مزایا:</b></p> <p>✓ بهره‌وری این الگوریتم در مقایسه با سایر روش‌ها بیشتر است زیرا حریم‌انه‌ترین الگوریتم موجود است.</p> <p>✓ در مقایسه با سایر الگوریتم‌ها (مانند: نمونه‌گیری تامپسون) با تعداد نمونه‌های کمتر به نتایج بهتر یا یکسان همگرا می‌شود.</p> <p>✓ پیاده‌سازی آن ساده است.</p> <p><b>معایب:</b></p> <p>✓ الگوریتم ممکن است به مقدار حداکثر محلی</p>

<sup>61</sup> Gittins index

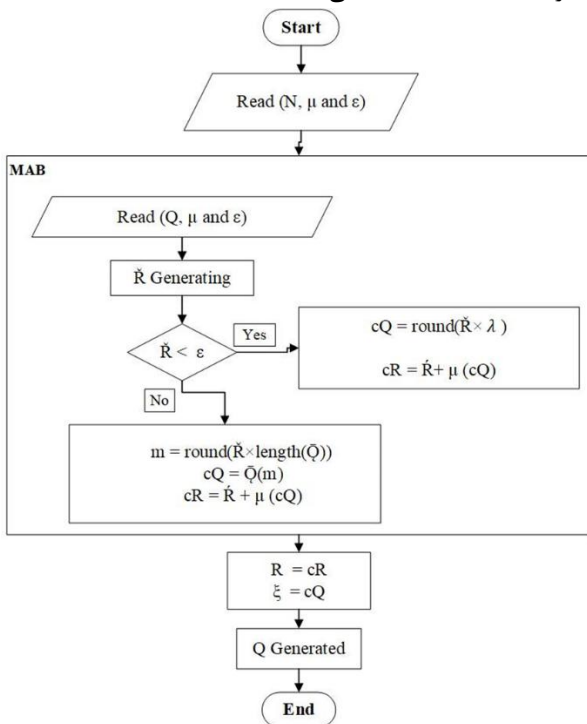
<sup>62</sup> Upper Confidence Bound

<sup>63</sup> Price of Knowledge and Estimated Reward

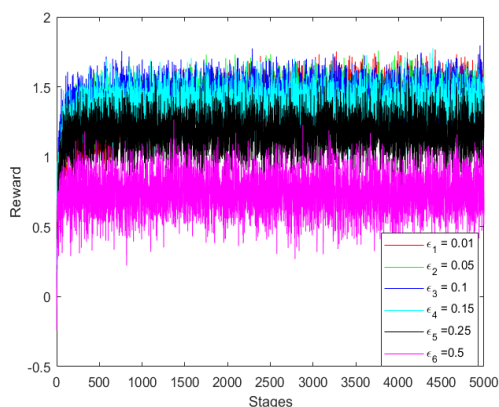
<sup>64</sup> Thompson Sampling and Optimism

<sup>65</sup> Best Empirical Sampled Average

در شکل ۲ مقادیر پاداش به ازای مقادیر مختلف  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهش نشان داده شده است. میانگین این مقادیر برای  $\epsilon$ های در نظر گرفته شده به ترتیب ۱،۳۷۸۲، ۱،۳۷۳۳، ۱،۴۱۱۲، ۱،۳۵۲۱، ۱،۱۶۸۳ و ۰،۷۴۲۴ است که بیشترین میزان پاداش به ازای مقدار  $\epsilon = 0.1$  است. همچنین در شکل ۳ اقدام بهینه در مسئله‌ی MAB به ازای مقادیر مختلف  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهش نشان داده شده است. میانگین مقادیر به دست آمده به ازای مقادیر متفاوت  $\epsilon$  به ترتیب ۰،۶۸۰۸، ۰،۸۰۰۳، ۰،۷۸۹۳، ۰،۸۲۲۹، ۰،۷۰۲۴ و ۰،۵۳۸۲ است. براساس این مقادیر می‌توان نتیجه گرفت که مقدار بهینه به ازای  $\epsilon = 0.15$  بدست می‌آید.



فلوچارت ۱: فلوچارت الگوریتم  $\epsilon$ -greedy کاهشی



شکل ۲: مقادیر پاداش به ازای مقادیر مختلف  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهشی

مطابق فلوچارت ۱ محاسبه می‌شود. در بخش بعدی تاثیر بکارگیری مقادیر بهینه این الگوریتم در سناریوهای متفاوت به منظور ارزیابی تاثیر متغیرهای eICIC در شبکه بررسی می‌شود. در الگوریتم  $\epsilon$ -greedy، مقدار  $\epsilon$  به صورت ایستا انتخاب می‌شود و در انتهای هر گام، خروجی با بالاترین کارایی با احتمال  $\epsilon$  انتخاب و با احتمال  $1 - \epsilon$  پذیرفته نخواهد شد. در الگوریتم  $\epsilon$ -greedy احتمال انتخاب اقدام بعدی به صورت  $\epsilon = \frac{1}{k}$  است و  $k$  نیز شماره مرحله‌ی فعلی است. ورودی‌های الگوریتم  $\epsilon$ -greedy کاهش، شامل  $k$  (تعداد گام)،  $\epsilon$  و  $\mu$  (مقدار پاداش واقعی هر راهزن<sup>۶۶</sup>) است. با محاسبه مقادیر اقدام انتخاب شده در هر گام ( $\xi$ ) و بردار  $R$  که نشان‌دهنده‌ی پاداش هر مرحله است، می‌توان مقدار  $\square$  یا پاداش هر راهزن را تخمین زد. این الگوریتم برای تخمین  $\square$  از روش MAB استفاده می‌کند که در این روش مقادیر  $\square$ ،  $\epsilon$  و  $\mu$  به عنوان ورودی دریافت و مقادیر  $cR$  و  $cQ$  به عنوان خروجی الگوریتم تولید می‌شوند. همچنین  $\lambda$  نشان‌دهنده‌ی تعداد راهزنان و با طول  $\square$  برابر است،  $\hat{R}$  یک عدد تصادفی با توزیع نرمال،  $\hat{R}$  یک عدد تصادفی با توزیع یکنواخت در بازه‌ی  $(0,1)$  و  $\bar{Q}$  بیشترین مقدار  $\square$  است. با تغییر مقادیر  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهش، مقدار  $\square$  تغییر می‌کند و مقدار بهینه  $\square$  با هدف محاسبه‌ی بهترین مقادیر متغیرهای تکنیک eICIC محاسبه می‌شود. در این مقاله برای شبیه‌سازی الگوریتم  $\epsilon$ -greedy کاهش، مقادیر  $\epsilon$  از مجموعه‌ی  $0.1$ ،  $0.15$ ،  $0.25$ ،  $0.5$  و  $1$  انتخاب شده است.

در شبکه‌های موبایل، میزان تبادل اطلاعات و به طور کلی بسته‌های منتقل شده شبکه در حال رشد است، به همین دلیل در شکل ۲ و شکل ۳ تاثیر تعداد مختلف بسته‌های ارسالی به ازای مقادیر متفاوت  $\epsilon$  نشان داده شده است. در الگوریتم‌های مبتنی بر یادگیری، با گذشت زمان تعداد نمونه‌ها افزایش می‌یابد و سیستم به حالت بهینه نزدیک‌تر می‌شود. همان‌طور که در شکل ۲ و شکل ۳ نشان داده شده است، در ابتدای فرآیند یادگیری، سیستم در هر گام به صورت تصادفی اقدام مرحله‌ی بعدی را انتخاب می‌کند. به همین دلیل به ازای تمامی مقادیر  $\epsilon$  در گام اول، مقدار پاداش برابر با صفر است اما با افزایش تعداد گام‌ها، پاداش سیستم افزایش می‌یابد و سیستم عمل بهینه را انتخاب می‌کند. میزان یادگیری سیستم‌ها توسط برخی عوامل محدود می‌شود، به طوری که امکان رسیدن به حالت بهینه‌ای که پاداش سیستم را به مقدار بیشینه ممکن برساند وجود ندارد. اما همانگونه که در شکل ۲ و شکل ۳ نشان داده شده است، پاداش سیستم به طور متوسط پس از ۱۰۰۰ گام همگرا می‌شود.

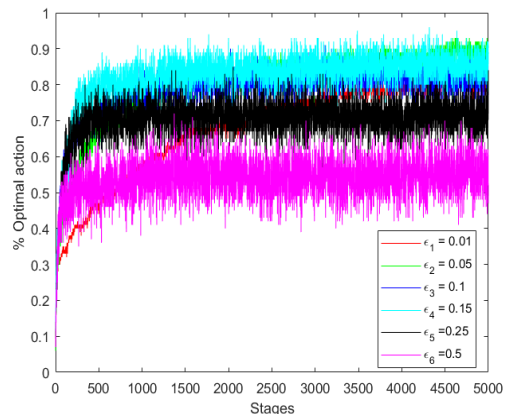
<sup>66</sup> Actual Reward for Each Bandit

جدول ۳: توصیف سناریوهای مقاله

توصیف سناریو	نوع سناریو
یک شبکه با ۱۹ سلول و هر سلول با یک ایستگاه ماکروسل و دو ایستگاه پیکوسل	سناریوی اول
یک شبکه با ۱۹ سلول و هر سلول با یک ایستگاه ماکروسل و چهار ایستگاه پیکوسل	سناریوی دوم
یک شبکه با ۱۹ سلول و هر سلول با یک ایستگاه ماکروسل و ۱۲ ایستگاه پیکوسل	سناریوی سوم
یک شبکه با ۱۹ سلول و هر سلول با یک ایستگاه ماکروسل اما، تعداد ایستگاه‌های پیکوسل در مرحله ۲۵۰۰ از سناریو ۱ به سناریو ۲ تغییر می‌کند.	سناریوی پویا

### نتایج

در این بخش، تاثیر روش ارائه شده بر شبکه‌ی سلولی با چهار سناریو مجزا که در جدول ۳ نشان داده شده است، بررسی می‌شود. در شکل ۵، شکل ۶ و شکل ۷ به ترتیب مقدار  $E[F(x)]$  سیستم برای الگوریتم‌های مختلف در سناریوهای اول، دوم و سوم نشان داده شده است. با توجه به نتایج بدست آمده، مقدار  $E[F(x)]$  در الگوریتم  $\epsilon$ -greedy کاهش می‌یابد در سناریوهای مشابه از مقادیر خروجی سایر الگوریتم‌های  $\epsilon$ -greedy، روش نمونه‌گیری تامسون و UCB نرمال کمتر است. به طور کلی با تغییر تعداد ایستگاه‌های پیکوسل، خروجی شبکه‌ها تغییر خواهد کرد، اما الگوریتم UCB نرمال به صورت چشمگیری، ضعیف‌تر از سایر الگوریتم‌ها عمل کرده است. این در حالی است که الگوریتم  $\epsilon$ -greedy کاهش می‌یابد در مقایسه با سایر الگوریتم‌های ارائه شده، با در نظر گرفتن شرایط مختلف شبکه، مقادیر بهینه را برای پارامترهای ABS و CRE انتخاب می‌کند. به طور کلی الگوریتم‌های مبتنی بر یادگیری براساس دانش گذشته‌ی خود تصمیم‌گیری می‌کنند و در هر مرحله از دانش به دست آمده از مراحل قبل استفاده می‌کنند. اما مسئله‌ی MAB با بهره‌گیری از گام‌های گذشته و با انتخاب مسیرهای جدید و بهره‌برداری از دانش، برای گام بعدی تصمیم‌گیری می‌کند. برخی از روش‌ها به مجموعه‌ای از داده‌ها برای تصمیم‌گیری نیاز دارند که موجب افزایش حجم داده‌های مورد پردازش و طولانی شدن زمان بهره‌برداری از دانش و تصمیم‌گیری برای گام بعدی می‌شود. اما الگوریتم  $\epsilon$ -greedy کاهش می‌یابد با افزایش تعداد نمونه‌های بسیار کم به یک مقدار بهینه همگرا می‌شود و با افزایش تعداد نمونه‌ها، افزایش دانش سیستم موجب کاهش تداخل در شبکه می‌شود.

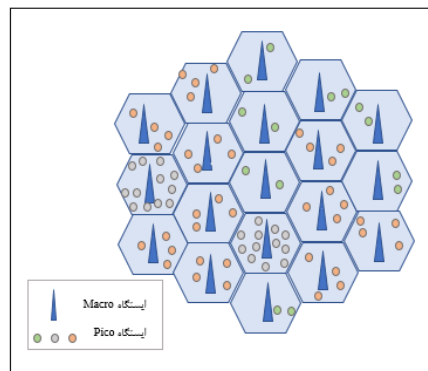


شکل ۳: مقادیر بهینه در مسئله‌ی MAB به ازای مقادیر مختلف  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهش می‌یابد

### سنجش عملکرد

### تنظیمات شبیه‌سازی

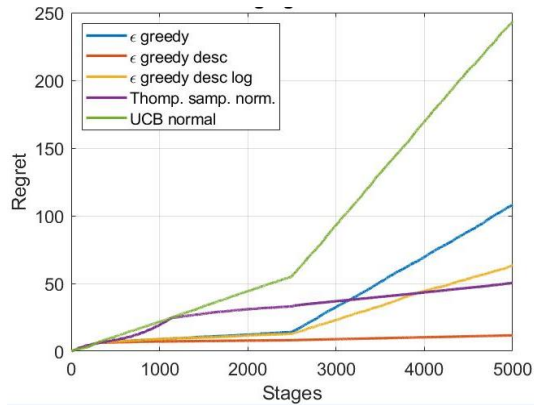
در این بخش تاثیر مقدار بهینه  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهش می‌یابد بر مقدار Regret در سه سناریوی مختلف در شبکه‌های HetNet بررسی می‌شود. سناریوهای در نظر گرفته شده در این مقاله، برگرفته از سناریوهای ارائه شده در جدول ۳ است. در شکل ۴ حالت کلی سیستم نشان داده شده است، که در این شبکه، سیستم در نظر گرفته شده با ۱۹ سلول پوشش داده شده است و در هر یک از سلول‌ها یک ایستگاه ماکروسل و در هر یک از سه سناریوی ایستا مجزا، در هر سلول علاوه بر یک ایستگاه ماکروسل، به ترتیب ۲، ۴ و ۱۲ ایستگاه پیکوسل نیز در نظر گرفته شده است. همچنین در سناریوی پویای در نظر گرفته شده نیز تا گام ۲۵۰۰، از ۲ ایستگاه پیکوسل و پس از آن ۴ ایستگاه پیکوسل در نظر گرفته شده است. آنتن تمامی ویژگی سناریوهای مختلف بیان شده است. آنتن تمامی ایستگاه‌های ماکروسل و پیکوسل به صورت تمام جهته<sup>۶۷</sup> سیگنال ارسال می‌کنند. همچنین پهنای باند در نظر گرفته شده برای شبکه‌های LTE برابر با 10Mbps و توان ارسالی ایستگاه‌های ماکروسل و پیکوسل به ترتیب 46 dBm و 30 dBm است.



شکل ۴: سناریوی HetNet

<sup>67</sup> Omnidirectional

در شکل ۸ مقدار  $E[F(x)]$  در سناریوی پویا بررسی شده است. در این سناریو مقدار  $E[F(x)]$  در الگوریتم  $\epsilon$ -greedy کاهش کمتر از مقدار بدست‌آمده در سایر الگوریتم‌های مفروض است. به طور کلی هر چقدر میزان  $E[F(x)]$  کاهش پیدا کند، الگوریتم از نظر کنترل متغیرهای تکنیک eICIC موفق‌تر عمل کرده است.

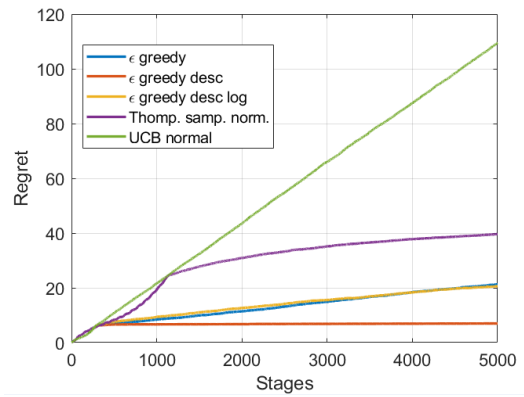


شکل ۸: ارزیابی عددی شبکه در شرایط شبکه متغیر

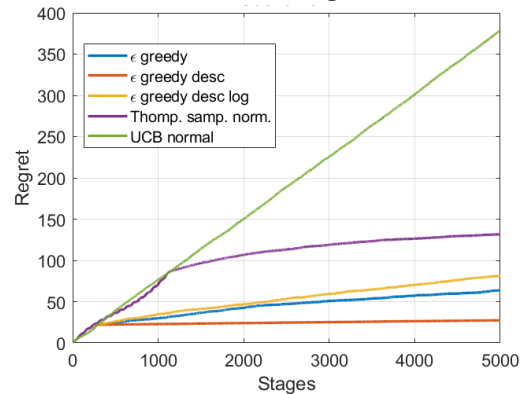
با مقایسه‌ی نتایج بدست آمده در این سناریوها، میزان تداخل در شبکه با افزایش تعداد ایستگاه‌های پیکوسل و همچنین تعداد مراحل، افزایش می‌یابد، اما میزان تاثیرپذیری الگوریتم‌ها متفاوت است. با مقایسه‌ی خروجی‌های الگوریتم UCB، با مقدار افزایش تعداد گام‌ها در شکل ۵، شکل ۶ و شکل ۷، مقدار Regret در شبکه با افزایش تعداد ایستگاه‌های پیکوسل افزایش می‌یابد. به طور مثال خروجی الگوریتم UCB برای ۵۰۰۰ گام در سه سناریوی ایستا به ترتیب ۱۰۷، ۳۷۵ و ۲۰۰۰ است که این نتایج نشان می‌دهند که دو برابر شدن تعداد ایستگاه‌های پیکوسل باعث دوبرابر شدن میزان تداخل نمی‌شود بلکه تاثیر آن بسیار چشم‌گیر است. به همین دلیل انتخاب الگوریتم بهینه، که به صورت نمایی تحت تاثیر افزایش تعداد ایستگاه‌های پیکوسل نباشد بر کیفیت ارتباطات بسیار حائز اهمیت است. البته تاثیرپذیری افزایش تعداد ایستگاه‌های پیکوسل بر کیفیت الگوریتم‌های نشان داده شده قابل توجه است اما نسبت تاثیرپذیری آنها متفاوت است.

### نتیجه‌گیری

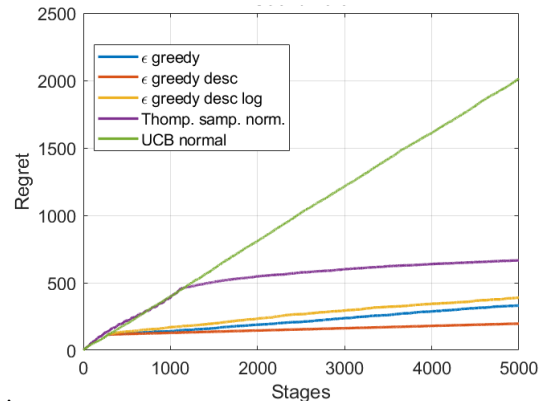
امروزه گسترش ارتباطات و بهبود کیفیت در برقراری و تبادل اطلاعات به یکی از مهم‌ترین اهداف شبکه‌هایی بی‌سیم تبدیل شده است و این پیشرفت‌ها موجب افزایش تقاضا برای دستیابی به سرویس‌های متنوع شده است. این افزایش تقاضا موجب تغییر و تحولاتی گسترده در دسترسی به منابع شبکه‌های مخابراتی، تکامل فناوری‌های بی‌سیم و ارائه‌ی مفاهیم جدیدی برای کاهش تداخل در شبکه و بهبود QoE شده است. در این مقاله با هدف کاهش تداخل در شبکه‌های ناهمگن، الگوریتم eICIC ارائه شد که مقدار بهینه متغیرهای آن توسط الگوریتم



شکل ۵: ارزیابی عددی شبکه در سناریوی اول



شکل ۶: ارزیابی عددی شبکه در سناریوی دوم



شکل ۷: ارزیابی عددی شبکه در سناریوی سوم

مقایسه‌ی خروجی‌های الگوریتم‌های مختلف در شکل ۵، شکل ۶ و شکل ۷ نشان می‌دهد که در سناریوی سوم، با افزایش تعداد ایستگاه‌های پیکوسل، تداخل به صورت چشم‌گیری نسبت به سناریوهای اول و دوم افزایش می‌یابد. زیرا در این سناریو ایجاد هماهنگی بین تعداد بیشتری از ایستگاه‌های ماکروسل و پیکوسل مورد نیاز است و انتخاب مقادیر بهینه‌ی پارامترها با محدودیت‌های بیشتری مواجه است. بنابراین تعداد ایستگاه‌هایی که در یک سلول مورد استفاده قرار می‌گیرند، موجب افزایش تعداد منابع در دسترس برای کاربران می‌شود و اما میزان تداخل را نیز افزایش می‌دهند. به همین دلیل، انتخاب تعداد بهینه‌ی ایستگاه‌های پیکوسل در شبکه، نیاز به مصالحه بین تعداد ایستگاه‌ها و تداخل قابل تحمل در شبکه می‌شود.

- parameters in small cell networks,” IEEE Transactions on Wireless Communications, 2017, 16(10), pp. 6635-6647.
- [9] A. Daeinabi, and K. Sandrasegaran, “A fuzzy Q-learning approach for enhanced intercell interference coordination in LTE-Advanced heterogeneous networks”. Proceedings, in The 20th Asia-Pacific Conference on Communication (APCC2014), 2014, pp. 139-144, IEEE.
- [10] M. Ghavamzadeh, S. Mannor, J. Pineau and A. Tamar, “Bayesian reinforcement learning: A survey,” Foundations and Trends in Machine Learning, 2015, 8(5-6), pp. 359-483.
- [11] N. Taheri Javan, M. Sabaei and V. Hakami, “IEEE 802.15.4.e TSCH-Based Scheduling for Throughput Optimization: A Combinatorial Multi-Armed Bandit Approach,” in IEEE Sensors Journal, 2020, 20(1), pp. 525-537.
- [12] B. Soret, H. Wang, K.I. Pedersen and C. Rosa, “Multicell cooperation for LTE-advanced heterogeneous network scenarios,” IEEE Wireless Communications Magazine, 2013, 20(1), pp. 27-34.
- [13] G. Burtini, J. Loepky and R. Lawrence, “A survey of online experiment design with the stochastic multi-armed bandit,” arXiv preprint arXiv:1510.00757, 2015.
- [14] N. Kato, Md. F. Zubair, M. Bomin, T. Fengxiao, A. Osamu, I. Takeru and M. Kimihiro, “The deep learning vision for heterogeneous network traffic control: Proposal, challenges, and future perspective,” IEEE wireless communications, 2016, 24(3), pp. 146-153.
- [15] S. A. Khan, A. Kavak, s. Aldirmaz Çolak and K. Küçük, “A Novel Fractional Frequency Reuse Scheme for Interference Management in LTE-A HetNets,” IEEE Access, vol. 7, pp. 109662-109672, 2019 .
- [16] R. Raheem, A. Lasebae, M. Aiash and J. Loo, “Interference Management for Co-channel Mobile Femtocells Technology in LTE Networks,” 2016 12th International Conference on Intelligent Environments (IE), London, 2016, pp. 80-87.
- [17] W. Chen, I. Ahmad and K. Chang, “Co-channel interference management using eICIC/FeICIC with coordinated scheduling for the coexistence of PS-LTE and LTE-R networks,” EURASIP Journal on Wireless Communications and Networking 2017, vol. 34.
- [18] O. Besbes, Y. Gur and A. Zeevi, “Optimal Exploration–Exploitation in a Multi-armed Bandit Problem with Non-stationary Rewards,” Stochastic Systems, 2019, 9.4, pp 319-337.
- $\epsilon$ -greedy کاهشی و روش حل مسئله‌ی MAB محاسبه شد. نتایج شبیه‌سازی‌ها نشان داد که مقادیر متفاوت  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهشی بر مقدار پاداش الگوریتم موثر است. به همین دلیل ابتدا مقدار بهینه‌ی  $\epsilon$  محاسبه و با بهره‌گیری از آن، مقادیر متغیرهای الگوریتم eICIC با هدف کمینه کردن میزان تداخل در HetNet محاسبه شدند. نتایج به دست آمده از مقایسه الگوریتم  $\epsilon$ -greedy کاهشی با سایر روش‌های الگوریتم  $\epsilon$ -greedy، نمونه برداری و فلسفه‌ی خوش‌بینی تامپسون و UCB نشان می‌دهد که مقدار بهینه‌ی  $\epsilon$  در الگوریتم  $\epsilon$ -greedy کاهشی موجب افزایش سرعت همگرایی الگوریتم، افزایش پاداش شبکه و گذردهی شبکه می‌شود. همچنین بهبود گذردهی شبکه بهبود QoE کاربران را در پی خواهد داشت.

## مراجع

- [1] Zh. Meisu, Y. Yang, H. Yao, X. Fu, O. A. Dobre and O. Postolache, “5G and IoT: Towards a new era of communications and measurements,” IEEE Instrumentation & Measurement Magazine, 2019, 22(6), pp. 18-26.
- [2] GSMA Intelligence, “Understanding 5G: perspectives on future technological advancements in mobile,” White paper, 2014.
- [3] H.B. Valiveti and P.T. Rao, “EHSD: an exemplary handover scheme during D2D communication based on decentralization of SDN,” Wireless Personal Communications, 94(4), pp. 2393-2416, June 2017.
- [4] J.S.B. Orange, A.G. Armada, B. Evans, A. Galis and H. Karl, “White paper for research beyond 5G”, 2015, Accessed 25.09.2020, url:<https://networld2020.eu/wpcontent/uploads/2015/11/B5G-Vision-for-Researchv-1.0-for-public-consultation.pdf>
- [5] S. Sasi, R.D. Daruwala, T. Palav and P. Mule, “A hybrid vertical handoff decision algorithm for seamless mobility in heterogeneous wireless networks,” Proceedings, in 2016 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET), 2016, pp. 1921-1923, IEEE.
- [6] M. Jiang, D. Xenakis, S. Costanzo, N. Passas and Mahmoodi, T, “Radio Resource Sharing as a service in 5G: A software-defined networking approach,” Computer Communications, 2017, vol. 107, pp. 13-29.
- [7] J.A. Ayala-Romero, J.J. Alcaraz and J. Vales-Alonso, “Data-driven configuration of interference coordination parameters in HetNets,” IEEE Transactions on Vehicular Technology, 2018, 67(6), pp. 5174-5187.
- [8] J.A. Ayala-Romero, J.J. Alcaraz, J. Vales-Alonso and E. Egea-López, “Online optimization of interference coordination

- [25] M. Samimi, G. MacCartney, Sh. Sun, T. Rappaport, "28 GHz Millimeter-Wave Ultrawideband Small-Scale Fading Models in Wireless Channels," In 2016 IEEE 83rd Vehicular Technology Conference (VTC Spring), pp. 1-6. IEEE, 2016.
- [26] M. Jorge, R. Abrishambaf, T. Gomes, P. Gonçalves, J. Cabral, A. Tavares and J. Monteiro, "Path loss exponent analysis in wireless sensor networks: Experimental evaluation," In 2013 11th IEEE international conference on industrial informatics (INDIN), pp. 54-58. IEEE, 2013.
- [27] C. Zhang, P. Patras and H. Haddadi, "Deep Learning in Mobile and Wireless Networking: A Survey," in IEEE Communications Surveys & Tutorials, 2019, 21(3), pp. 2224-2287, thirdquarter.
- [28] S. Chinchali, P. Hu, T. Chu, M. Sharma, M. Bansal, R. Misra, M. Pavone, S. Katti, "Cellular Network Traffic Scheduling With Deep Reinforcement Learning," in Thirty-Second AAAI Conference on Artificial Intelligence, 2018, pp. 766-774.
- [29] Konda, V.R. and Tsitsiklis, J.N. "Onactor-critic algorithms", SIAM journal on Control and Optimization, 2003, 42(4), pp.1143-1166.
- [30] Nasir YS, Guo D. "Deep actor-critic learning for distributed power control in wireless mobile networks". In 2020 54th Asilomar Conference on Signals, Systems, and Computers 2020 Nov 1 (pp. 398-402). IEEE.
- [19] S. Yunqing, J. Cao, M. Ma, H. Li, B. Niu and F. Li, "Privacy-preserving device discovery and authentication scheme for D2D communication in 3GPP 5G HetNet," In 2019 International Conference on Computing, Networking and Communications (ICNC), pp. 425-431. IEEE, 2019.
- [20] A. I. Sulyman, A. T. Nassar, M. K. Samimi, G. R. Maccartney, T. S. Rappaport and A. Alsanie, "Radio propagation path loss models for 5G cellular networks in the 28 GHz and 38 GHz millimeter-wave bands," in IEEE Communications Magazine, 2014. vol. 52, no. 9, pp. 78-86.
- [21] D. Tsolkas, E. Liotou, N. Passas, L. Merakos, "A survey on parametric QoE estimation for popular services," Journal of Network and Computer Applications, 2017, Vol. 77, PP. 1-17.
- [22] X. Yan, J. Xu, Y. Zhu, J. Wang, Y. Yang and C. Wang, "Downlink Average Rate and SINR Distribution in Cellular Networks," in IEEE Transactions on Communications, vol. 64, no. 2, pp. 847-862, Feb. 2016,
- [23] A.M. Saleh, N. Parveen and Md Forkan Uddin, "Optimal scheduling of coordinated multipoint transmissions in cellular networks," International Journal of Communication Systems 31, no. 1, 2018, e3431.
- [24] D. Tsolkas, E. Liotou, N. Passas, L. Merakos, "A Survey on Parametric QoE Estimation for Popular Services," 2017, ELSEVIER.