

بهینه‌سازی منابع رادیویی در یک محیط مبتنی بر سطوح هوشمند قابل تنظیم مجدد

مسعود علی‌زاده^۱، علی نوروزی^۲، بیژن عباسی آرنند^۳، نادر مکاری یامچی^۴

^۱دانشجوی کارشناسی ارشد مهندسی برق مخابرات میدان و موج، دانشگاه تربیت مدرس، masoud.alizadeh1993@gmail.com

^۲دانشجوی دکتری مهندسی برق مخابرات، دانشگاه تربیت مدرس، ali_nouruzi@modares.ac.ir

^۳دانشیار دانشکده‌ی برق و کامپیوتر دانشگاه تربیت مدرس، abbasi@modares.ac.ir

^۴دانشیار دانشکده‌ی برق و کامپیوتر دانشگاه تربیت مدرس، nader.mokari@modares.ac.ir

چکیده

یکی از کاربردهای سطوح هوشمند قابل تنظیم مجدد دست‌یابی به محیط‌های رادیویی هوشمند است. محیط‌های رادیویی هوشمند برای شبکه‌های بی‌سیم آینده این قابلیت را ایجاد می‌کنند که دارای اتصال بی‌سیم بی‌وقفه باشند و انتقال داده‌ها بدون ایجاد سیگنال‌های جدید و بلکه با بازیافت امواج رادیویی موجود صورت گیرد. در ارتباطات نسل ششم چون از فرکانس‌های بالا استفاده می‌شود، برای جلوگیری از تلفات و به‌کارگیری توان پایین از مفهوم سطوح هوشمند قابل تنظیم مجدد به کار گرفته می‌شود. در این مقاله هدف اصلی بهبود نرخ داده‌ی همه‌ی کاربران دارای تحرک در شبکه‌ی بدون سلول به کمک سطوح هوشمند قابل تنظیم مجدد است. در این مقاله با در نظر گرفتن یک سیستم چندکاربره‌ی MIMO، با تعریف یک مسئله‌ی بهینه‌سازی، تأثیر استفاده از سطوح هوشمند قابل تنظیم مجدد با المان‌های مختلف در محیط برای افزایش نرخ تبادل اطلاعات مورد بررسی قرار می‌گیرد. ظرفیت مسئله‌ی بهینه‌سازی را نسبت به زمان برای حالتی که RIS داریم و حالتی که RIS نداریم ترسیم می‌کنیم و تأثیر وجود سطوح هوشمند قابل تنظیم مجدد را به‌وضوح مشاهده می‌کنیم. نمودارهای ظرفیت متوسط هر کاربر برحسب زمان و همچنین نمودار ظرفیت متوسط برحسب تعداد کاربران رسم شده است.

کلیدواژه

شبکه‌های بی‌سیم، سطوح هوشمند قابل تنظیم مجدد، ارتباطات نسل ششم، بهینه‌سازی.

مقدمه

شود، به این سیستم شبکه‌ی بدون سلول^۲ گفته می‌شود [2]. در شبکه‌های سلولی همواره با پدیده‌ای به‌نام تداخل بین سلولی روبه‌رو هستیم، برای مقابله با تداخل بین سلولی روش‌های زیادی ارائه شده است، یکی از این روش‌ها شبکه‌های بدون سلول است که به‌عنوان یک تکنیک امیدوارکننده برای افزایش ظرفیت سیستم‌های بی‌سیم آینده در نظر گرفته شده‌اند. با این حال، برای افزایش بیشتر ظرفیت، نیاز به استقرار تعداد بیشتری ایستگاه پایه با هزینه و توان مصرفی بالا دارد. برای پرداختن به این موضوع، از تکنیک RIS که اخیراً پیشنهاد شده [3] الهام گرفته شده است، ما مفهوم شبکه‌های بدون سلول به کمک RIS را برای افزایش ظرفیت شبکه با توان مصرفی و هزینه‌ی پایین پیشنهاد می‌کنیم. سطوح بازتابی هوشمند (RIS) یک تکنیک ارتباط بی‌سیم مخرب است که قادر به ایجاد یک محیط

یک سطح هوشمند قابل تنظیم مجدد (RIS) می‌تواند خواص موج برخوردی مانند فرکانس، دامنه و فاز را تغییر دهد و سپس موج دست‌کاری شده را بدون نیاز به پردازش پیچیده‌ی سیگنال به یک مقصد مطلوب منعکس کند [1]. سطوح هوشمند قابل تنظیم مجدد (RIS^۱) اخیراً به‌عنوان یک فناوری امیدوارکننده ظاهر شده‌اند که می‌توانند با ادغام تعداد زیادی از عناصر بازتاب‌دهنده‌ی کم‌هزینه و غیرفعال، به طیف بالا و راندمان انرژی بالا برای شبکه‌های بی‌سیم آینده دست یابند.

یک منطقه‌ی تحت پوشش با تعداد زیادی نقاط دسترسی بی‌سیم در نظر گرفته می‌شود، این منطقه سلول بندی نمی‌شود و به هر کاربر به‌طور هم‌زمان توسط همه آنتن‌های نقطه دسترسی خدمات داده می‌شود.

^۲ Cell free

^۱ Reconfigurable Intelligent Surface

دیگر مانند نویسندگان در [4] و [5] با معرفی مدل مداری تغییر در شکل پرتو ایجاد می‌کنند و با استفاده از شیفت فاز گسسته و استفاده از روش Alternating Optimization به ایجاد مدلی بهینه برای انتخاب تغییر فازی که هر المان باید بر روی موج دریافتی ایجاد کند می‌پردازند.

نشانه گذاری

نشانه گذاری‌های استفاده شده در این مقاله به شرح زیر است: حروف بزرگ پر رنگ و حروف کوچک پر رنگ به ترتیب نشان‌دهنده-ی ماتریس‌ها و بردارها هستند و تمامی حروف معمولی نشان‌دهنده‌ی کمیت‌ها و اعداد اسکالر هستند. ترانهاده‌ی هرمیتی، ترانهاده و معکوس ماتریس به ترتیب با بالانویس حروف T, H, 1- نشان داده شده است. $I_{K \times K}$ نشان‌دهنده‌ی ماتریس همانی $K \times K$ است. علامت \otimes نشان‌دهنده‌ی ضرب کرنکر دو ماتریس، $det(0)$ نشان‌دهنده‌ی دترمینان یک ماتریس، علامت $vec(0)$ نشان‌دهنده‌ی برداری کردن یک ماتریس به شکل ستونی، $diag(0)$ تبدیل کننده‌ی یک بردار به یک ماتریس قطری مربعی و $\mathbf{E}\{\cdot\}$ نشان دهنده‌ی امید ریاضی می‌باشد. همچنین $[A]_{i,j}$ نشان‌دهنده‌ی درایه‌ی واقع در سطر i ام و ستون j ام در ماتریس A است. علائم و نشانه‌های دیگر در صورت نیاز در متن مقاله توضیح داده خواهد شد.

مدل سیستم

برای افزایش ظرفیت شبکه با هزینه و مصرف توان پایین، مفهوم شبکه بدون سلول به کمک RIS پیشنهاد شده است. به‌طور خاص، ما یک سناریوی پهن باند را همان‌طور که در شکل ۱ نشان داده شده است، در نظر می‌گیریم، جایی که چندین ایستگاه پایه، چندین RIS، چندین کاربر داری تحرک، پهن باند و MIMO به‌طور هم‌زمان در نظر گرفته می‌شوند. مجموعه‌های $\mathcal{M} = \{1, \dots, M\}$ و $\mathcal{U} = \{1, \dots, U\}$ و $\mathcal{N} = \{1, \dots, N\}$ و $\mathcal{B} = \{1, \dots, B\}$ و $\mathcal{R} = \{1, \dots, R\}$ و $\mathcal{K} = \{1, \dots, K\}$ و $\mathcal{P} = \{1, \dots, P\}$ به ترتیب تعداد آنتن‌های ایستگاه پایه، تعداد آنتن‌های کاربران، تعداد زیر RIS، تعداد ایستگاه‌های پایه، تعداد RISها، تعداد کاربران و تعداد زیر حامل‌ها را نشان می‌دهند. در شبکه‌ی بدون سلول به کمک RIS پیشنهاد شده، همه‌ی BSها همگام شده‌اند.

ارسالی دارای توان نرمالیزه شده هستند. در فرسو $s_{p,k}$ سمبل ارسال شده به کاربر k ام روی زیرحامل p ام را نشان می‌دهد. ما فرض می‌کنیم که سمبل‌های

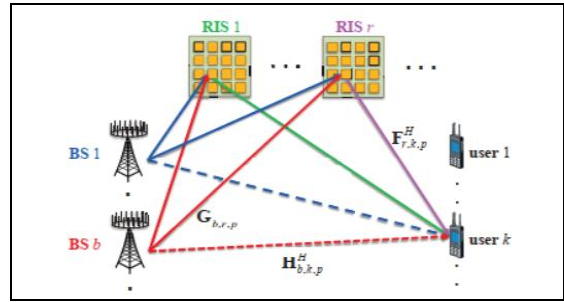
انتشار کنترل‌پذیر است. این امر با ساخت تعداد زیادی از عناصر بازتابی غیرفعال حاصل می‌شود که تغییرات فاز قابل کنترل را بر روی امواج الکترومغناطیسی وارد شده اعمال می‌کنند به‌گونه‌ای که سیگنال بازتاب شده می‌تواند برای کاربر موردنظر یا به‌طور مخرب در کاربر ناخواسته ایجاد شود تا تداخل را کاهش دهد. RIS می‌تواند برای بهبود ظرفیت کانال، کاهش قدرت انتقال، افزایش قابلیت اطمینان انتقال استفاده شود. پس، برای شبکه بدون سلول به کمک RIS پیشنهادی، ما مسئله‌ی طراحی کدگذاری شده مشترک در ایستگاه پایه و RIS را برای به حداکثر رسیدن ظرفیت شبکه فرمول بندی می‌کنیم. با توجه به غیر محدب بودن و پیچیدگی بالای مسئله‌ی فرموله شده، ما یک الگوریتم بهینه‌سازی متناوب را برای حل این مسئله‌ی چالش‌برانگیز توسعه می‌دهیم. سرانجام، نتایج شبیه‌سازی تأیید می‌کند که، در مقایسه با شبکه بدون سلول معمولی ظرفیت شبکه طرح پیشنهادی می‌تواند به‌طور قابل توجهی بهبود یابد. ازجمله کاربردها و تحقیقات انجام شده بر روی IRS می‌توان به کارهای [4] و [5] اشاره کرد، که به بررسی یک سیستم MISO³ نقطه‌به‌نقطه بهبودیافته با کمک IRS می‌پردازند که با استفاده از تکنیک SDR قصد در بهبود توان دریافتی در محل کاربر با فرض وجود اطلاعات CSI در IRS دارند. به دلیل نیاز به تخمین‌های پیچیده و بار انتقال این داده‌ها بین آنتن اصلی و IRS روش توزیع شده‌ای را ارائه می‌دهند که آنتن و IRS هرکدام به‌طور مستقل با استفاده از اطلاعات کانال به تغییر شکل الگوی تشعشعی یا تغییر فاز سیگنال برخوردی اقدام می‌کنند تا جایی که مجموع این دو به همگرایی برسند و کیفیت سیگنال دریافتی بهبود یابد. در کار [6] با استفاده از روش یادگیری عمیق سعی در تخمین فاز انعکاسی هر المان در IRS با استفاده از داده‌های ذخیره شده در پایگاه داده و توان دریافتی کاربر دارند. به این صورت که در فاز یادگیری که به‌صورت آفلاین است ابتدا شبکه عصبی عمیق DNN⁴ که وظیفه‌ی یافتن نگاشتی بین مختصات کاربر و تنظیم المان‌های IRS برای اینکه بیشترین توان در محل کاربر دریافت می‌شود، روشی برای محاسبه فاز بازتابش می‌یابد سپس در فاز آنلاین، با توجه به مختصات اندازه‌گیری شده کاربر، ضرایب طوری تنظیم می‌شوند که بیشترین توان در محل کاربر دریافت شود. از این صفحات IRS می‌توان در محیط‌های داخلی برای کاهش تداخل و سایه اجسام موجود در محل که مانع از رسیدن سیگنال مناسب به کاربر می‌شوند استفاده کرد. همچنین در موردی

³ Multi Input Single Output
⁴ Deep Neural Network

شده ابتدا توسط بردار کدگذاری شده $W_{b,p,k} \in \mathbb{C}^k$ در ایستگاه \mathbb{B} کدگذاری شده است.

شده ابتدا توسط بردار کدگذاری شده $W_{b,p,k} \in \mathbb{C}^k$ در ایستگاه \mathbb{B} کدگذاری شده است.

شده ابتدا توسط بردار کدگذاری شده $W_{b,p,k} \in \mathbb{C}^k$ در ایستگاه \mathbb{B} کدگذاری شده است.



شکل ۱. کانال‌های فرسو در شبکه‌ی بدون سلول پهن باند به کمک RIS

$$y_{k,p} = \sum_{b=1}^B y_{b,k,p} + z_{k,p} = \sum_{b=1}^B \sum_{j=1}^k (H_{b,k,p}^H + \sum_{r=1}^R F_{r,k,p}^H \theta_r^H G_{b,r,p}) W_{b,p,j} s_{p,j} + z_{k,p}$$

$$= \sum_{b=1}^B (H_{b,k,p}^H + \sum_{r=1}^R F_{r,k,p}^H \theta_r^H G_{b,r,p}) W_{b,p,k} s_{p,k} + \sum_{b=1}^B \sum_{j=1, j \neq k}^k (H_{b,k,p}^H + \sum_{r=1}^R F_{r,k,p}^H \theta_r^H G_{b,r,p}) W_{b,p,j} s_{p,j} + z_{k,p} \quad (5)$$

بر اساس سیستم مدل فوق سیگنال دریافت شده $y_{b,k,p}$ در (۵) را می توان به صورت رابطه (۶) ساده کرد:

بنابراین سمبل کدگذاری شده $X_{b,p}$ در ایستگاه \mathbb{B} روی زیر حامل \mathbb{B} را می توان به صورت زیر نوشت:

$$y_{k,p} \triangleq^{(a)} \sum_{b=1}^B \sum_{j=1}^k (H_{b,k,p}^H + F_{k,p}^H \theta_r^H G_{r,p}) W_{b,p,j} s_{p,j} + z_{k,p} \triangleq^{(b)} \sum_{j=1}^k h_{k,p}^H W_{p,j} s_{p,j} + z_{k,p} \quad (6)$$

بصورتی که در قسمت اول رابطه با $\theta = \text{diag}(\theta_1, \dots, \theta_r)$ و $G_{r,p} = [G_{1,r,p}^T, \dots, G_{B,r,p}^T]^T$ و $F_{k,p} = [F_{1,k,p}^T, \dots, F_{R,k,p}^T]^T$ با $W_{p,k} = [W_{1,p,k}^T, \dots, W_{B,p,k}^T]^T$ و $h_{k,p} = [h_{1,k,p}^T, \dots, h_{B,k,p}^T]^T$ تعریف می شوند. در این مقاله فرض بر این است که اطلاعات کانال به صورت دقیق در اختیار گیرنده قرار دارد و به عبارت دیگر اطلاعات کانال ایده آل فرض شده است.

$$X_{b,p} = \sum_{k=1}^K W_{b,p,k} s_{p,b,k} \quad (1)$$

کانال بین هر ایستگاه پایه و هر کاربر در شبکه‌ی بدون سلول به کمک RIS پیشنهاد شده از دو قسمت تشکیل شده است: لینک ایستگاه پایه- کاربر و لینک ایستگاه پایه-RIS-کاربر، بنابراین کانال معادل $h_{b,k,p}^H$ از ایستگاه پایه‌ی \mathbb{B} به کاربر \mathbb{K} روی زیر حامل \mathbb{B} را می توان به صورت زیر نوشت:

$$h_{b,k,p}^H = H_{b,k,p}^H + \sum_{r=1}^R F_{r,k,p}^H \theta_r^H G_{b,r,p} \quad (2)$$

در معادله‌ی (۲) $G_{b,r,p} \in \mathbb{C}^{N \times M}$ و $H_{b,k,p}^H \in \mathbb{C}^{U \times M}$ و $F_{r,k,p} \in \mathbb{C}^{U \times N}$ کانال فرکانسی روی زیر حامل \mathbb{B} را به ترتیب از ایستگاه پایه‌ی \mathbb{B} به کاربر \mathbb{K} و از ایستگاه پایه‌ی \mathbb{B} به r RIS $\theta_r \in \mathbb{C}^{N \times N}$ ماتریس شیفیت فاز در r RIS را نشان می دهد، که به صورت زیر نوشته می شود:

$$\theta_r \triangleq \text{diag}(\theta_{r,1}, \dots, \theta_{r,n}), \forall r \in R \quad (3)$$

مدل تحرک کاربران در سیستم مدل مور دنظر

برای تحلیل و شبیه سازی عملکرد تحرک، ما باید مدل تحرک مناسبی برای توصیف الگوی حرکتی کاربران تلفن همراه در شبکه در نظر بگیریم.

چون هدف ما به دست آوردن یک ارتباط تحلیلی و مدل برای مدیریت تحرک است، مدل وابسته برای استفاده مناسب نیست، بنابراین از مدل تصادفی استفاده می کنیم. در این راستا، مدل تحرک انفرادی تصادفی نقطه ای (RWP) بیشتر مورد استفاده قرار می گیرد.

از معروف ترین مدل های حرکتی مدل Random Way Point یا RWP می باشد. این مدل توسط Johnson و Maltz پیشنهاد شده است در این مدل هر گره به صورت تصادفی گره دیگری را به عنوان مقصد انتخاب کرده و با یک سرعت ثابت که از محدوده بازه $[V_{\max}, 0]$ به صورت تصادفی انتخاب می کند، به طرف آن مقصد شروع به حرکت می کند. پس از رسیدن گره به مقصد به اندازه یک زمان تصادفی T_{pause} توقف کرده و سپس گره دیگری را به عنوان مقصد انتخاب کرده و همان فرآیند قبلی را تکرار می کند [7].

که $\theta_{r,n} \in f$ مجموعه عملیاتی (قابل دستیابی) از ضرایب بازتاب (RC) در RIS است و به صورت رابطه‌ی (۴) تعریف می شود:

$$f \triangleq \{\theta_{r,n} \mid |\theta_{r,n}| \leq 1, \forall r \in R, \forall n \in N\} \quad (4)$$

پس از عبور از کانال معادل $h_{b,k,p}^H$ همان طور که در (۲) نشان داده شده، سیگنال ها توسط کاربران دریافت می شوند.

برای کاربران ثابت و شرایط موفقیت‌آمیز (Hand over) HO برای کاربران موبایل در حال حرکت است:

$$P^0: \max_{\theta, w} f(\theta, w) = \sum_{b=1}^B \sum_{k=1}^K \sum_{p=1}^P \Pi_k b_b^k \rho_{b,p}^k \log_2(1 + \gamma_{k,p}) \quad (8)$$

$$\begin{aligned} \text{S.T} \\ \text{C1: } & \sum_{k=1}^K \sum_{p=1}^P \rho_{b,p}^k \|w_{b,p,k}\| \leq p_{b,max}, \forall b \in B \\ \text{C2: } & 0 \leq \theta_r \leq \pi \quad \forall r \in R \\ \text{C3: } & E_R\{HSR\} \geq \tau \quad \forall k \in K \\ \text{C4: } & p_k^0 > p_{min} \quad \forall k \in K \\ \text{C5: } & \rho_{b,p}^k \in \{0,1\} \quad \forall k, b, p \\ \text{C6: } & \sum_{k=1}^K \rho_{b,p}^k \leq 1 \quad \forall p \in P, b \in B \\ \text{C7: } & b_b^k \in \{0,1\} \quad \forall k \in K, \forall b \in B \\ \text{C8: } & \sum_{b=1}^B b_b^k \leq 1 \quad \forall k \in K \end{aligned}$$

Π_k نشان‌دهنده‌ی وزن کاربر k ام است. و w به صورت زیر تعریف می‌شود:

$$W = [w_{1,1}^T, w_{1,2}^T, \dots, w_{1,k}^T, w_{2,1}^T, w_{2,2}^T, \dots, w_{n,k}^T]^T$$

(C1) محدودیت توان را نشان می‌دهد.

(C2) محدودیت شیفت فاز را نشان می‌دهد.

(C3) برای آنکه HO بین کاربران رخ دهد، HSR باید از آستانه‌ی τ بیشتر یا مساوی باشد.

(C4) محدودیت نرخ هر کاربر را نشان می‌دهد.

(C5,6) بیان می‌کند که هر زیر حامل نمی‌تواند، هم‌زمان توسط ۲ یا چند کاربر در یک BS به اشتراک گذاشته شود.

(C7,8) هر کاربر فقط توسط یک BS خدمات می‌گیرد.

حل مسئله بهینه‌سازی به کمک یادگیری تقویتی

تفاوت عمده‌ی این نوع یادگیری با دیگر روش‌ها این است که یادگیری در این روش به صورت تجربه‌ای حاصل می‌شود به این صورت که در هر مرحله از یادگیری با اجرای یک عمل در تعامل با محیط یادگیری، آن محیط را از حالتی به حالت دیگری می‌برد. برای ارزیابی چگونگی عملکرد این روش یادگیری، یک تابع کمی برای بررسی یادگیری و کارایی آن در تعامل با محیط تحت عنوان پاداش تعریف می‌شود [8]. بر اساس تجربیات به دست آمده از اجرای اعمال مختلف در حالت‌های گوناگون، یک نوع سیاست به دست می‌آید که بر اساس آن نحوه‌ی انجام عمل را در حالت‌های متعاقب بعدی را رهنمون می‌شود. با ذخیره‌ی پاداش اعمال در هر حالت خاص در یک جدول، در صورت قرارگیری مجدد محیط آموزش در یک آن حالت عملی انتخاب می‌شود که بیشتری پاداش را حاصل کند. نحوه‌ی گذار

در این مدل حرکتی، انتخاب مقصد بعدی، انتخاب سرعت و انتخاب زمان توقف در مقصد همگی به صورت تصادفی انجام می‌شود. سرعت، جهت و مقصد و زمان توقف هر گره مستقل از گره‌های دیگر است. در مورد زمان توقف هرگاه $T_{pause} = 0$ باشد اجازه حرکت داده خواهد شد. در این مدل T_{pause} و V_{max} دو پارامتر برای تعیین رفتار تحرک گره‌ها می‌باشد. هر چقدر V_{max} کم و T_{pause} بیشتر باشد توپولوژی شبکه نسبتاً پایدار است و برعکس هر چقدر سرعت زیاد و زمان توقف کم باشد توپولوژی بشدت پویا است.

در این مدل زمان مکث بین تغییرات جهت یا سرعت در نظر گرفته شده است. نقطه‌ی راه‌اندازی تصادفی در این مدل را می‌توان در ۵ مرحله تعریف کرد:

(۱) یک مقصد جدید انتخاب کنید.

(۲) سرعت با توزیع یکنواخت انتخاب کنید.

($[speed_{min} \quad speed_{max}]$)

(۳) حرکت تا رسیدن به مقصد.

(۴) در موقعیت به مدت زمان مشخص شده (زمان مکث) صبر کنید.

(۵) بازگشت به مرحله‌ی اول.

همچنین در این مدل $speed_{max}$ و زمان مکث بسیار مؤثر هستند. اگر $speed_{max}$ زیاد باشد و زمان مکث کم باشد، شبکه پویا می‌باشد و اگر $speed_{max}$ و زمان مکث زیاد باشد، شبکه به پایداری نسبی می‌رسد.

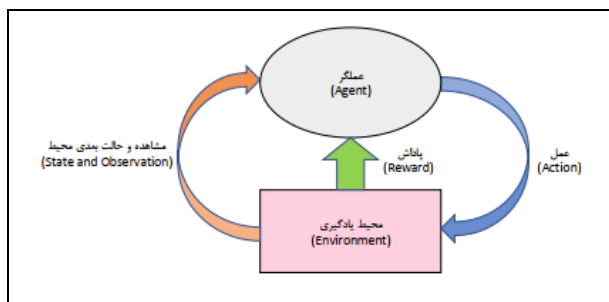
مسئله بهینه‌سازی

SINR از کاربر k ام که متصل شده به ایستگاه پایه‌ی p ام روی زیر حامل p ام به صورت زیر به دست می‌آید:

$$\gamma_{k,p} = w_{b,p,k}^H h_{k,p} \left(\sum_{j=1, j \neq k}^K h_{k,p}^H w_{p,j} (h_{k,p}^H w_{p,j})^H + [I]_{k,p} \right)^{-1} h_{k,p}^H w_{b,p,k} \quad (7)$$

یک کاربر موبایل تحت پوشش قرار دارد اگر به یکی از b تا B BS موجود متصل باشد و SINR بزرگ‌تر یا مساوی β_i داشته باشد. ماکزیمیم توان انتقال از هر BS محدودیت تخصیص زیر حامل در نظر گرفته شده است. هدف اصلی این مسئله‌ی بهینه‌سازی به حداکثر رساندن نرخ داده هم‌هی کاربران شبکه با در نظر گرفتن شرایط پوشش

در همین راستا، فرآیندهای یادگیری مبتنی بر شبکه‌ی عصبی عمیق را یادگیری عمیق می‌نامند. به‌کارگیری واژه "عمیق" برگرفته از تفاوت ساختاری شبکه‌های جدید با شبکه‌های نسل پیش از خود در تعداد لایه، ساختار و معماری اتصال لایه‌هاست [13] و [14]. یادگیری تقویتی خود به چندین روش تقسیم‌شده است که با توجه به گسسته یا پیوسته بودن مسئله از این روش‌ها استفاده می‌شود، از آنجا که مسئله‌ی بهینه‌سازی به‌کاربرده شده در این مقاله پیوسته است از روش Deep Deterministic Policy Gradient برای حل مسئله‌ی بهینه‌سازی استفاده شده است.



شکل ۲. چگونگی عملکرد یادگیری تقویتی و نحوه‌ی تعامل بین عملگر و محیط یادگیری.

نتایج شبیه‌سازی

در این مقاله ورودی‌های محیط شبیه‌سازی مسئله‌ی بهینه‌سازی مطابق جدول ۱ است.

جدول ۱: ورودی‌های مسئله‌ی بهینه‌سازی.

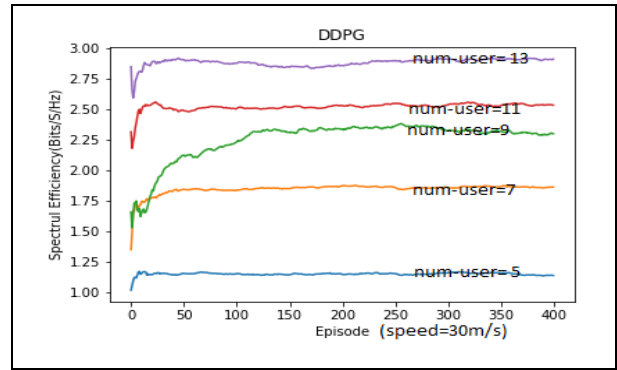
مقدار	پارامتر
2s	Time Slots
1200m	بعد X
1200m	بعد Y
10	Number sub channel
4	Number BS
4	Number RIS
40 watts	P Max
(5, 10, 20, 30)m/s	Velocity
4	Number-antenna

در حالت‌های محیط آموزش در این نوع یادگیری بر اساس یک مدل فرآیند تصمیم‌گیری مارکوف یا یک‌شبه مدل از آن است. در برخی از منابع تحقیقاتی و موارد عملی به‌کارگیری یادگیری تقویتی، بحث از یک عملگر یا چند عملگر به میان می‌آید که با انجام عمل‌های مختلف در حالت‌های مختلف، در یک رابطه‌ی دوسویه بین محیط و عملگر شرکت می‌کند و در نتیجه به تجربه یا پاداش منتهی می‌شود. در شکل (۲) نحوه‌ی عملکرد یادگیری تقویتی نمایش داده شده است. شایان ذکر است که در بسیاری از موارد بکارگیری یادگیری تقویتی به دلیل تعداد بسیار زیاد حالت‌های موجود برای هر مرحله از یادگیری و همچنین نیاز به ذخیره‌ی حالت‌ها، عمل‌ها و پاداش به‌دست‌آمده از آن‌ها در تمام Transaction (مراحل گذار)، استفاده از یک جدول و یا حافظه غیرعملی است [9] و [10]. بنابراین از یک شبکه‌ی عصبی برای تخمین پاداش‌های به‌دست‌آمده از عمل‌های مختلف در حالت‌های گوناگون استفاده می‌کنند. به این ترتیب که حالت کنونی در هر مرحله از یادگیری را به‌عنوان ورودی شبکه‌ی عصبی فرض کرده و سپس خروجی شبکه‌ی عصبی تعیین‌کننده‌ی عملی است که بیشترین پاداش را فراهم آورد [11] و [12]. اگر چه استفاده از شبکه‌ی عصبی در کاربردهای مختلف در دو دهه‌ی گذشته همواره موردتوجه بوده است، اما در سال‌های اخیر به دلیل ظهور روش‌های جدید و معرفی ساختارهای نوین برای این شبکه‌ها و از طرفی افزایش قدرت پردازشی و محاسباتی سخت‌افزارهای موجود، شاهد رشد چشم‌گیر نوعی از شبکه‌های عصبی تحت عنوان شبکه‌های عصبی عمیق بوده‌ایم.

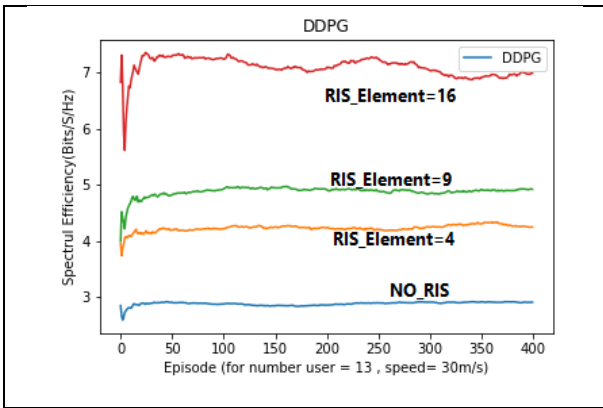
به منظور تطبیق یک روش مبتنی بر یادگیری تقویتی برای حل مسئله مورد نظر، ابتدا ماتریس ضرایب کانال به صورت یک بردار تبدیل می‌شود. سپس این بردار به عنوان یک ورودی به یک شبکه‌ی عصبی عمیق با معماری خاص و معین اعمال می‌شود. سپس در خروجی آن، مقادیر متغیرهای بهینه‌سازی بدست می‌آید. پس از محاسبه‌ی میزان بهبود تابع هدف بهینه‌سازی محاسبه می‌شود. در صورتی تغییر متغیرهای بهینه‌سازی موجب بهبود تابع هدف شود، ضرایب و وزن‌های درونی شبکه عصبی ثابت باقی می‌مانند. در غیر این صورت، در طی یک الگوریتم داخلی شبکه‌ی عصبی، که Replaying نام دارد، وزن‌های شبکه‌ی عصبی بر اساس روش بازگشت به عقب^۵، اصلاح می‌شوند.

^۵ Back Propagation

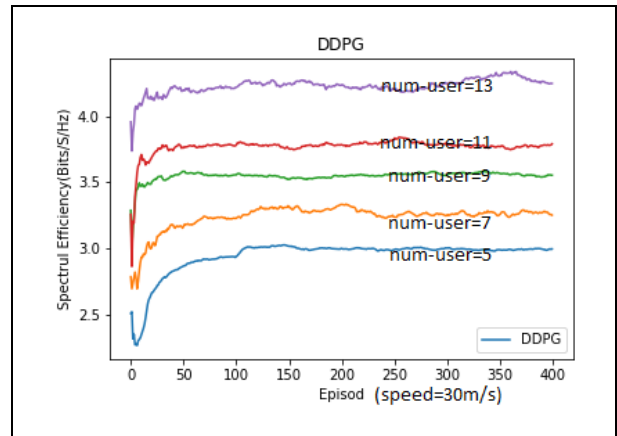
شکل ۳ یک محیط فاقد RIS برای زمانی که تعداد کاربران برابر ۵ و ۷ و ۹ و ۱۱ و ۱۳ و سرعت هرکاربر برابر 30m/s است را نشان می‌دهد. شکل‌های ۴ و ۵ و ۶ محیط‌های حاوی RIS برای زمانی که تعداد المان‌ها برابر ۴ و ۹ و ۱۶ و تعداد کاربران ۵ و ۷ و ۹ و ۱۱ و ۱۳ و سرعت هرکاربر برابر 30m/s است را نشان می‌دهد. مطابق شکل‌ها هرچه تعداد کاربران بیشتر مجموع نرخ دیتا در کلیه کاربران بیشتر می‌شود چون در این حالت multi user diversity زیاد می‌شود و ممکن است کاربر خوبی به شبکه اضافه شود و باعث می‌شود نرخ دیتا کاربران بیشتر شود.



شکل ۳. محیط فاقد RIS.

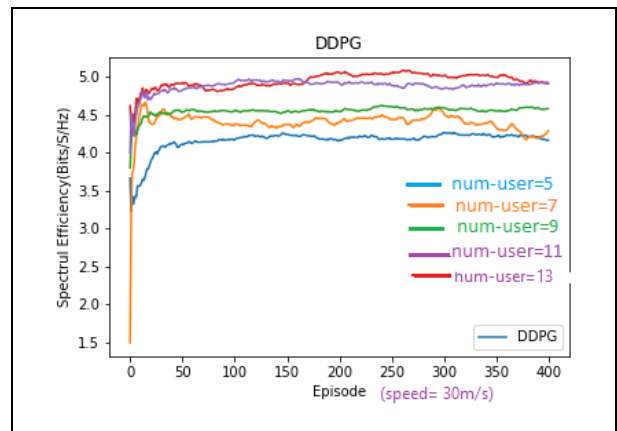


شکل ۷. مقایسه‌ی محیط حاوی RIS با المان‌های ۴ و ۹ و ۱۶ با محیط فاقد RIS از نظر نرخ داده.

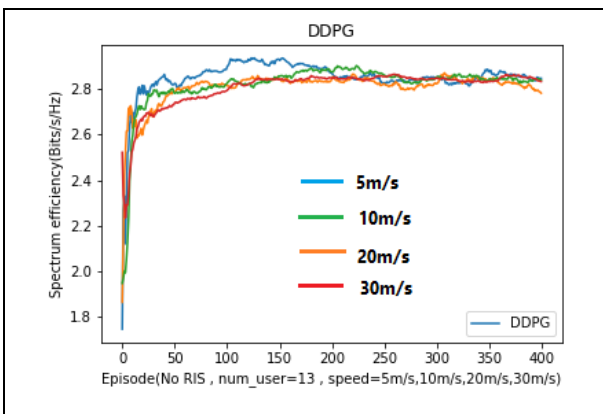


شکل ۴. محیط حاوی RIS با تعداد ۴ المان.

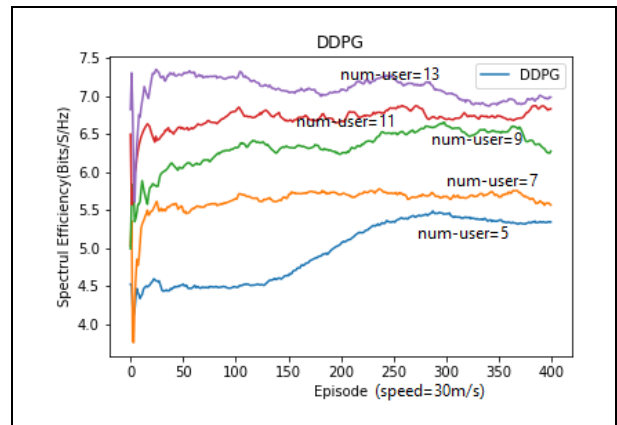
شکل ۷ یک محیط فاقد RIS را با یک محیط حاوی RIS موقعی که تعداد المان‌ها ۴ و ۹ و ۱۶ است مقایسه می‌کند. مطابق شکل محیط حاوی RIS نرخ دیتای بیشتری نسبت به محیط فاقد RIS دارد و هرچه تعداد المان‌های RIS بیشتر باشد نرخ دیتا بیشتر خواهد بود.



شکل ۵. محیط حاوی RIS با تعداد ۹ المان.



شکل ۸. محیط فاقد RIS زمانی که تعداد کاربران ۱۳ و سرعت تحرک 5m/s, 10m/s, 20m/s, 30m/s است.



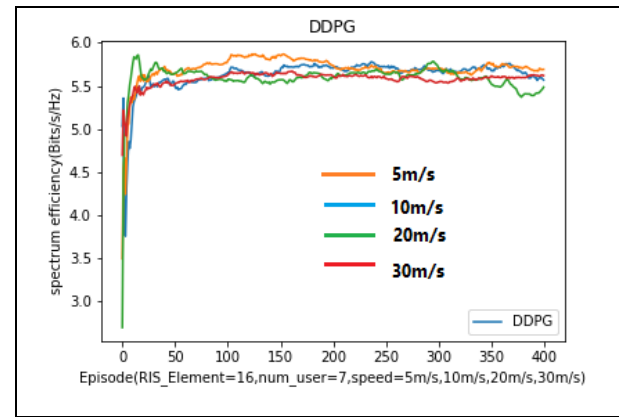
شکل ۶. محیط حاوی RIS با تعداد المان ۱۶.

روش آزمند	$ W + H + \theta $	$E*\sum_{n=1}^{N-1} h_n h_{n+1}$
	اندازه ماتریس کانال و ماتریس متغیرهای بهینه‌سازی	تعداد نرون های هر لایه در لایه‌ی بعد از آن

نتیجه‌گیری و کارهای آینده

در این پژوهش، هدف اصلی بهینه‌سازی منابع رادیویی با بکارگیری سطوح هوشمند قابل تنظیم مجدد با تنظیم شیفت فاز در المان‌های سطوح هوشمند قابل تنظیم مجدد و تخصیص توان کدگذاری شده در سمت ایستگاه‌های پایه می‌باشد و پس از آن طراحی یک الگوریتم برای بهبود ظرفیت (به حداکثر رساندن نرخ دیتا) این شبکه‌ها بود. در این مقاله ابتدا یک مدل سیستم MIMO چندکاربره در حضور سطوح هوشمند قابل تنظیم مجدد با در نظر گرفتن تحرک کاربران تعریف شد. پس از آن روابط کانال مخابراتی بی‌سیم برای استفاده از سطوح هوشمند با قابلیت پیکربندی مجدد اصلاح شد و ظرفیت این سیستم بر اساس روابط کانال اصلاح شده به دست آمد.

در ادامه یک مسئله‌ی بهینه‌سازی جهت انتخاب بهترین حالت هر یک از سطوح هوشمند با قابلیت پیکربندی مجدد طراحی شد و پس از تشریح مرتبه‌ی پیچیدگی آن، با بهره‌گیری از شهود حاکم بر الگوریتم‌های بهینه‌سازی آزمند^۷، یک الگوریتم مبتنی بر هوش مصنوعی و یادگیری ماشین برای حل آن ارائه شد. در ادامه مدل سیستم تعریف شده طبق روابط ریاضی گفته شده شبیه‌سازی شده و نتایج استفاده سطوح هوشمند قابل تنظیم مجدد با الگوریتم پیشنهادی ارزیابی شد. در این مقاله نمودارهای نرخ داده متوسط هر کاربر بر حسب زمان رسم شده و نشان داده شد که با استفاده از سطوح هوشمند با قابلیت پیکربندی مجدد با و به کمک الگوریتم پیشنهادی می‌توان ظرفیت سیستم را افزایش داد. در این پژوهش اطلاعات وضعیت کانال به صورت ایده آل فرض شده از این رو در پژوهش‌های آینده می‌توان بر فرض اینکه اطلاعات وضعیت کانال را نمی‌دانیم نیز مطالعات خود را شروع کنیم. در این پژوهش ما از الگوریتم یادگیری دقیق DDPG مبتنی بر هوش مصنوعی استفاده کردیم در مطالعات بعدی می‌توان از سایر الگوریتم‌های یادگیری



شکل ۹. محیط حاوی RIS با تعداد المان‌های ۱۶ و زمانی که تعداد کاربران ۷ و سرعت تحرک 5m/s, 10m/s, 20m/s, 30m/s است.

شکل ۸ و شکل ۹ به ترتیب یک محیط فاقد RIS و یک محیط حاوی RIS را زمانی که کاربران دارای تحرک با سرعت‌های 5m/s, 10m/s, 20m/s, 30m/s هستند را نشان می‌دهد. از این نمودارها می‌توان نتیجه گرفت که در الگوریتم پیشنهادی زمانی که سرعت کم یا زیاد می‌شود کارایی تغییر نمی‌کند. این الگوریتم سرعت‌های بالا را هم حمایت می‌کند و به همان عملکرد سرعت پایین می‌رساند.

پیچیدگی محاسباتی و داده‌ی سرباره‌ی روش یادگیری تقویتی:

برای حل هر مسئله‌ی بهینه‌سازی، نیاز به یک مجموعه از اطلاعات مانند ضرایب کانال مخابراتی و یا نقاط اولیه هستیم. برای انتقال این مجموعه‌ی داده در شبکه همواره نیازمند ارتباطات هستیم که یک سرباره‌ی اطلاعاتی^۶ نامیده می‌شود.

از سویی دیگر، بدلیل اینکه پیچیدگی محاسباتی روش‌های مبتنی بر یادگیری تقویتی به ساختار شبکه‌ی عصبی مورد استفاده در آن‌ها، از نظر معماری درونی و تعداد لایه‌ها و نرون‌ها وابسته است، در جدول ۲ روابط مربوط به پیچیدگی محاسباتی و داده‌ی سرباره نوشته شده است.

جدول ۲ پیچیدگی محاسباتی و سرباره‌ی داده‌ی روش حل

روش حل	سرباره‌ی داده	پیچیدگی محاسباتی
DDPG	$ W + H + \theta $ اندازه ماتریس کانال و ماتریس متغیرهای بهینه‌سازی	$E*\sum_{n=1}^{N-1} h_n h_{n+1}$ تعداد تکرارهای حل مسئله (E)*تعداد نرون های هر لایه در لایه‌ی بعد از آن

^۷ Greedy

^۶ Signaling Overhead

[7] Reconfigurable Intelligent Surfaces.” *2019 IEEE 20th International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)* (2019): 1-5.

[8] B. Li, H. Zhang, and H. Lu, “User mobility prediction based on Lagrange’s interpolation in ultra-dense networks,” in *2016 IEEE 27th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, 2016, pp. 1–6.

[9] D. Koller, N. Friedman, S. Džeroski, C. Sutton, A. McCallum, A. Pfeffer, P. Abbeel, M. F. Wong, D. Heckerman, C. Meek, et al. *Introduction to statistical relational learning*. MIT press, 2007

[10] K. Arulkumaran, M. P. Deisenroth, M. Brundage, and A. A. Bharath, “A brief survey of deep reinforcement learning,” arXiv preprint arXiv:1708.05866, 2017

[11] N. C. Luong, D. T. Hoang, S. Gong, D. Niyato, P. Wang, Y.C. Liang, and D. I. Kim, “Applications of deep reinforcement learning in communications and networking: A survey,” *IEEE Communications Surveys & Tutorials*, vol.21, no.4, pp.3133–3174, 2019

[12] F. R. Yu and Y. He, “Deep reinforcement learning for wireless networks, springer,” 2019

[13] L. P. Kaelbling, M. L. Littman, and A. W. Moore, “Reinforcement learning: A survey,” *Journal of artificial intelligence research*, vol.4, pp.237–285, 1996.

[14] J. Schmidhuber, “Deep learning in neural networks: An overview,” *Neural networks*, vol.61, pp.85–117, 2015

[15] Z. Zhang, P. Cui, and W. Zhu, “Deep learning on graphs: A survey,” *IEEE Transactions on Knowledge and Data Engineering*, 2020

مانند الگوریتم Actor_Critic استفاده نمود و با الگوریتم فعلی قیاس انجام شود.

مراجع

[1] M. Jung, W. Saad, M. Debbah, and C. S. Hong, “On the optimality of reconfigurable intelligent surfaces (RISs): Passive beamforming, modulation, and resource allocation,” *arXiv Prepr. arXiv1910.00968*, 2019

[2] Z. Zhang and L. Dai, “Capacity improvement in wideband reconfigurable intelligent surface-aided cell-free network,” in *2020 IEEE 21st International Workshop on Signal Processing Advances in Wireless Communications (SPAWC)*, 2020, pp. 1–5.

[3] S. Mosleh, H. Almosa, E. Perrins, and L. Liu, “Downlink resource allocation in cell-free massive MIMO systems,” in *2019 International Conference on Computing, Networking and Communications (ICNC)*, 2019, pp. 883–887.

[4] Q. Wu and R. Zhang, “Intelligent reflecting surface enhanced wireless network: Joint active and passive beamforming design,” in *Proc. IEEE GLOBECOM*, 2018, [Online] Available: <https://arxiv.org/abs/1809.01423>.

[5] Q. Wu and R. Zhang, “Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network,” [online] <https://arxiv.org/abs/1905.00152.2019>

[6] Huang, C., George C. Alexandropoulos, Chau Yuen and M’erouane Debbah. “Indoor Signal Focusing with Deep Learning Designed