مقاله پژوهشی

مدیریت انرژی خانه‌های هوشمند مبتنی بر یادگیری تقویتی چندعامله با لحاظ مشارکت همزمان در بازارهای انرژی و انعطاف‌پذیری

پیمان مادح خاکسار و سید حامد دلخوش

چکیده: توسعه تولیدات تجدیدپذیر و بکارگیری مدیریت تقاضا در شبکه‌های پایین‌دست برق، به‌عنوان راه‌حل‌هایی کلیدی جهت رفع معضلات پیش‌روی شبکه‌های قدرت آینده مورد توجه است. همچنین گسترش روندهای تمرکززدایی و دیجیتالی‌شدن، جایگاه مدیریت انرژی بهینه مبتنی بر رویکردهای نوین محاسباتی را در خانه‌های هوشمند (متشکل از تجهیزات تولید، مصرف و ذخیره‌سازی) برجسته نموده است. این مقاله، یک سیستم مدیریت انرژی برای خانه هوشمندی متشکل از وسایل (سه نوع توان ثابت، با قابلیت جابجایی زمانی و توان متغیر) و تولید پراکنده تجدیدپذیر (پنل خورشیدی) پیشنهاد داده است. این سیستم، تولید و قیمت بازار را از ابزار پیش‌بینی مبتنی بر شبکه عصبی دریافت کرده و برنامه بهینه روزانه را با استفاده از یادگیری تقویتی چندعامله (روش حل یادگیری Q) به دست می‌آورد. مهم‌ترین وجه نوآورانه مدل ارائه‌شده امکان مشارکت همزمان خانه در بازارهای انرژی و انعطاف‌پذیری است که از منظر کسب سود می‌توانند همسو یا خلاف سو باشند. در این مدل، عوامل پیچیده‌کننده رضایت کاربر (آسایش در بهره‌مندی از وسایل) و الزامات محدودکننده بالادستی (قطع تولید موردنیاز بهره‌بردار) نیز لحاظ شده است. صحت و اثربخشی روش پیشنهادی، از طریق پیاده‌سازی در نرم‌افزار پایتون، شبیه‌سازی سناریوهای متنوع و تحلیل حساسیت نسبت به پارامترهای کلیدی نشان داده شده است.

کلیدواژه: **مدیریت انرژی، خانه‌های هوشمند، بازار برق، انعطاف‌پذیری، یادگیری تقویتی چندعامله.**

فهرست علائم

اندیس‌ها، تعدادها و مجموعه‌ها

i, Ω^{FA}	اندیس و مجموعه وسایل خانگی توان ثابت
j, Ω^{PS}	اندیس و مجموعه وسایل خانگی توان متغیر
k, Ω^{PV}	اندیس و مجموعه پنل‌های خورشیدی
m, Ω^{TS}	اندیس و مجموعه وسایل خانگی با قابلیت جابجایی
n, Ω^{EV}	اندیس و مجموعه خودروهای برقی
t, T	اندیس و تعداد برش زمانی (مثلاً روزانه ۲۴ برش ۱ ساعتی)

پارامترها

$\alpha^{f}_{k,t}$	سطح تنظیمی الزام قطع تولید هر k در زمان t (بین ۰ و ۱)
$P^{PV}_{k,t}, \lambda^{G}_{t}$	مقادیر پیش‌بینی‌شده قیمت بازار انرژی (\$/kW) و توان تولیدی پنل خورشیدی k (kW) در زمان t
θ, γ	نرخ‌های یادگیری و تخفیف الگوریتم حل (بین ۰ و ۱)
σ	شمارنده تکرار الگوریتم حل
τ	آستانه تعیین‌شده برای شرط همگرایی الگوریتم حل (\$)
$\lambda^{up}_{t}, \lambda^{dn}_{t}$	قیمت تعیین‌شده بازارهای روبه‌پایین و روبه‌بالای بازارهای انعطاف‌پذیری در زمان t (\$/kW)
DC^{EV}_{n}, DC^{PS}_{j}	ضرایب نارضایتی جابجایی توانی هر j و n (\$/kW^2)
DC^{PV}_{k}	ضریب نارضایتی تخطی از الزام بالادستی برای هر k (\$)
DC^{TS}_{m}	ضریب نارضایتی جابجایی زمانی هر m (\$/h^2)
$P^{d,FA}_{i,t}$	توان مصرفی وسیله خانگی توان ثابت i در زمان t (kW)
$P^{\max,PS}_{j}, P^{\min,PS}_{j}$	مقادیر کمینه و بیشینه توان عملیاتی هر j (kW)
$P^{\max,EV}_{n}, P^{\min,EV}_{n}$	مقادیر کمینه و بیشینه توان عملیاتی هر n (kW)
$P^{us,EV}_{i,t}, P^{us,PS}_{i,t}$	مقادیر توان دلخواه کاربر برای هر j و n در زمان t (kW)
$[t^{\min}_{k}, t^{\max}_{k}]$	بازه معنادار (طول دوره روز) تولید هر k
$[t^{ini}_{m}, t^{end}_{m}]$	بازه مجاز (زمان شروع و پایان قابل قبول) کارکرد هر m
$[t^{arr}_{n}, t^{dep}_{n}]$	بازه عملیاتی (زمان رسیدن و ترک منزل) برنامه‌ریزی هر n

متغیرها

$\alpha_{k,t}$	مقدار منتخب سطح تولید برای هر در زمان t (بین ۰ و ۱)
r_t, s_t, a_t	اقدام، حالت و پاداش هر عامل در زمان t
$P^{d,EV}_{i,t}, P^{d,PS}_{i,t}$	مقادیر توان منتخب برای هر j و n در زمان t (kW)
Q	شاخصه اصلی تصمیم‌گیری بهینه در الگوریتم حل؛ برابر با مجموع پاداش‌های همه عوامل در هر دوره زمانی (\$)
$r^{FA}_{i,t}$	پاداش وسیله خانگی توان ثابت i در زمان t (\$)
$r^{PS}_{j,t}$	پاداش وسیله خانگی توان متغیر j در زمان t (\$)
$r^{PV}_{k,t}$	پاداش پنل خورشیدی k در زمان t (\$)
$r^{TS}_{m,t}$	پاداش وسیله خانگی با قابلیت جابجایی m در زمان t (\$)
$r^{EV}_{n,t}$	پاداش خودروی برقی n در زمان t (\$)
t^{s}_{m}	زمان منتخب کارکرد برای هر m
$u_{m,t}$	وضعیت انتخابی روشن بودن وسیله m در زمان t (باینری)

این مقاله در تاریخ ۱۱ شهریور ماه ۱۴۰۴ دریافت و در تاریخ ۱۳ آبان ماه ۱۴۰۴ بازنگری شد.

پیمان مادح خاکسار، گروه قدرت، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران، (email: p_khaksar@modares.ac.ir).

سید حامد دلخوش (نویسنده مسئول)، گروه قدرت، دانشکده مهندسی برق و کامپیوتر، دانشگاه تربیت مدرس، تهران، ایران، (email: h.delkhosh@modares.ac.ir).

۱- مقدمه

تغییرات بنیادین ساختاری و عملکردی اخیر در شبکه‌های برق در قالب پنج کلان‌روند کلیدی مردم‌سالاری، کربن‌زدایی، دیجیتالی‌شدن، تمرکززدایی و مقررات‌زدایی در حال شکل‌گیری است [۱] که خانه‌های

هوشمند در همه موارد جایگاه ویژه‌ای دارند. نقش این خانه‌ها در جهت اهداف کلان مردم‌سالاری در حوزه انرژی، خصوصاً مشتری-محور شدن شبکه‌های آینده، انکارنشدنی است [۲]. مهم‌ترین روند جهانی در راستای کربن‌زدایی هم استفاده از انرژی‌های تجدیدپذیر است که در این راستا، پنل‌های خورشیدی پشت‌بامی در خانه‌های هوشمند سهم بسزایی از توسعه را به خود اختصاص داده‌اند [۳]. دیجیتالی‌شدن که به دنبال تقویت زیرساخت فنی شبکه بوده، بر گسترش اینترنت اشیاء و کنتورهای هوشمند خانه‌ها استوار است [۴]. تمرکززدایی و توزیع‌کردن تصمیم‌گیری نیز اهمیت مدیریت بهینه انرژی در سطوح مختلف (از شبکه‌های توزیع فعال تا خانه‌های هوشمند) را برجسته کرده است [۵]. مقررات‌زدایی نیز خرید و فروش در سطح شبکه‌های محلی را با سازوکارهای گوناگون تسهیل نموده که در آن خانه‌های هوشمند (تجمیع بار، تولید و ذخیره‌سازی) از مهم‌ترین بازیگران هستند [۶]. بر اساس این اهمیت، مقاله حاضر بر سیستم‌های مدیریت انرژی خانگی[1] (HEMS) متمرکز شده است.

با افزایش تقاضای انرژی و گسترش منابع تولید تجدیدپذیر، بهره‌برداری پایدار شبکه‌های برق با چالش‌های متعددی روبه‌رو شده است [۷]. بهره‌گیری از پاسخگویی بار از مهم‌ترین راهکارهای حل معضلات بهره‌برداری در شبکه‌های توزیع است [۸]. با توسعه تجهیزات تولید و ذخیره‌سازی خصوصاً در خانه‌های هوشمند، گسترش بازارهای انرژی محلی نیز با توجه به مقررات‌زدایی سرعت بیشتری گرفته است [۹]. توسعه بازارهای انعطاف‌پذیری محلی نیز مورد توجه محققان قرار گرفته تا بتوانند چالش‌های فنی را با یک ساختار اقتصادی رقابتی و منصفانه مدیریت کنند [۱۰]. با توجه به فرصت‌های هم‌افزا یا متعارض این بازارها، مقاله حاضر به مسئله مشارکت همزمان خانه‌های هوشمند در بازارهای محلی انرژی و انعطاف‌پذیری پرداخته است.

مطالعات اندکی به مشارکت اقتصادی خانه‌های هوشمند در بازارهای انعطاف‌پذیری پرداخته‌اند؛ درحالی‌که مکرراً به پتانسیل زیاد منابع موجود در این خانه‌ها برای خدمات انعطاف‌پذیری اشاره شده است [۱۱]. به عنوان نمونه، در [۱۲] ظرفیت انعطاف‌پذیری خانه‌های هوشمند برای حل مشکلات سیستم‌های انتقال و توزیع به‌کار گرفته شده و در [۱۳] از مدیریت انرژی ساختمان‌های مسکونی جهت تسکین عدم‌قطعیت‌های موجود در تصمیم‌گیری‌های بهره‌برداران استفاده شده است. از سویی دیگر، انعطاف‌پذیری در مطالعات بالادستی HEMS (مدیریت انرژی جامعه-محور) مفصلاً بررسی شده است. به عنوان نمونه، [۱۴] به مدیریت انرژی چندهدفه چند ریزشبکه متصل با توجه به مسئله انعطاف‌پذیری پرداخته و [۱۵] روشی سلسله‌مراتبی برای مدیریت ذخیره‌ساز مشترک جوامع انرژی جهت تامین انعطاف‌پذیری توسعه داده است. اغلب این مطالعات بر نیازهای جامعه و شبکه‌های بالادستی متمرکزاند و جزئیات کلیدی مربوط به نیازها و ویژگی‌های فردی خانه‌های هوشمند (مثل راحتی کاربران و پیچیدگی‌های الزامات بالادستی) را نادیده گرفته‌اند.

علاوه بر چارچوب مفهومی مسئله مدیریت انرژی، رویکرد حل مسئله نیز حائز اهمیت است. روش سنتی حل چنین مسائلی، انواع فرم‌های بهینه‌سازی (خطی، غیرخطی، چندهدفه، چندسطحی و ...) است [۱۶]. برای مثال، [۱۷] مسئله HEMS خانه‌های روستایی را با توجه به ویژگی‌های دسترسی محدود به منابع و زیرساخت‌ها، مبتنی بر بهینه‌سازی انجام داده است. مرجع [۱۸] کنترل پیش‌بین را در مسئله HEMS به فرم برنامه‌ریزی خطی آمیخته با عدد صحیح ادغام کرده است. همه مطالعاتی که مسئله HEMS را مقید بر انعطاف‌پذیری کرده‌اند نیز از بهینه‌سازی بهره گرفته‌اند. به عنوان نمونه، [۱۹] یک بهینه‌سازی مقاوم برای HEMS ارائه نموده که انعطاف‌پذیری لازم جهت اتخاذ تصمیمات بی‌درنگ را برای مدیریت عدم‌قطعیت تولید و تقاضا دارا است. مرجع [۲۰] نیز یک HEMS چندهدفه ارائه کرده که محدودیت‌های انعطاف‌پذیری خانگی را در تعیین تعادل بین هزینه، رضایت و خودکفایی در نظر می‌گیرد. استفاده از رویکرد مدل-محور (بهینه‌سازی) برای HEMS نیازمند اطلاعات دقیق برای مدل‌سازی است که در کاربردهای زمان‌واقعی و محیط‌های پویا با چالش، از جمله مفروضات به‌روزرسانی و بار پردازشی بالا، روبه‌رو است. مضافاً اینکه مطالعات مقید بر انعطاف‌پذیری نیز فقط اغلب روی کاهش هزینه تمرکز دارند و به امکان تصمیم‌گیری برای مشارکت در بازارهای محصولات مختلف انعطاف‌پذیری کمتر توجه کرده‌اند.

روش‌های پیشرفته یادگیری ماشین، خصوصاً یادگیری تقویتی[2] (RL)، در سال‌های اخیر کاربردهای بسیاری در حوزه شبکه‌های هوشمند پیدا کرده است [۲۱]. به عنوان مطالعاتی نزدیک به مقاله حاضر، می‌توان به حوزه مدیریت انرژی محلی اشاره کرد که البته تمرکز کافی بر خانه‌های هوشمند، مدیریت دقیق تجهیزات آن و لحاظ مسئله انعطاف‌پذیری خانه‌ها ندارند. مرجع [۲۲] یک چارچوب مدل-آزاد برای مدیریت هاب‌های انرژی مبتنی بر تولید خورشیدی معرفی کرده که از RL عمیق برای تصمیم‌گیری خودکار و بهینه در شرایط مختلف بهره‌برداری استفاده کرده است. در [۲۳]، روش RL عمیق برای مدیریت هماهنگ سیستم‌های انرژی یکپارچه استفاده شده که در آن تدابیر انعطاف‌پذیری تامین و تقاضای انرژی برای انطباق با شرایط متغیر بازار برق و شبکه هوشمند لحاظ شده است. یادگیری تقویتی چندعاملی[3] (MARL) نیز کاربردهای بسیاری در مسائلی همچون تعیین استراتژی بهینه قیمت‌گذاری در جامعه انرژی از نگاه بازیگران میانی (مثل مدیر جامعه غیرانتفاعی یا تجمیع‌کننده سود-محور) [۲۴] پیدا کرده است. چنین مطالعاتی قابلیت‌های متمایز رویکرد مدل-آزاد (RL) را در مقایسه با رویکرد مدل- محور (بهینه‌سازی) مشخص می‌کنند. نظر به امکان یادگیری و انطباق در محیط‌های پویا و نامعین بعلاوه قابلیت تصمیم‌گیری ترتیبی و بلندمدت، رویکرد RL در حوزه HEMS نیز کاربرد پیدا کرده است [۲۵]. برخی از این مطالعات (چشم‌پوشی از انعطاف‌پذیر) در ادامه مرور شده است.

برخی از مطالعات، مدل‌های نسبتاً ساده مبتنی بر RL برای HEMS ارائه کرده‌اند که بر مکمل‌های کلیدی بهبوددهنده عملکرد متمرکزند. مرجع [۲۶] یک HEMS مبتنی بر MARL پیشنهاد داده که از تحلیل داده‌های بی‌درنگ برای تحلیل رفتار مصرف بهره می‌گیرد. در [۲۷] یک چارچوب بهینه‌سازی بر اساس DRL برای HEMS ارائه شده که پارامترهای دارای عدم‌قطعیت (رفتار کاربران، الگوهای مصرف و تولید تجدیدپذیر) را در نظر می‌گیرد. بخشی دیگری از مطالعات به موضوعات همچون بارهای گرمایشی و ذخیره‌سازهای حرارتی در HEMS پرداخته‌اند. مرجع [۲۸] به‌طور خاص بر سیستم‌هایی با پمپ‌های حرارتی مدولار و پنل خورشیدی تمرکز کرده تا HEMS بتواند ضمن سازگاری با تغییرات محیطی، وابستگی به شبکه و هزینه‌ها را کاهش دهد. در [۲۹] نیز مدل MARL پیشنهادی بر HEMS مسئله آسایش حرارتی کاربران را با تمرکز بر سیستم تهویه و هدف کاهش هزینه‌ها دنبال می‌کند.

1. Home Energy Management Systems

2. Reinforcement Learning

3. Multi-Agent Reinforcement Learning

جدول ۱: خلاصه مطالعات مرورشده مبتنی بر RL برای مسئله HEMS و مقایسه مفهومی با مقاله حاضر.

مرجع	مدیریت		تجهیزات خانه					رضایت انواع بار	الزام تولید	انعطاف پذیری	ویژگی خاص
	روش	دوره	بار ثابت	بار قابل جابجایی	بار توان متغیر	تولید پراکنده	ذخیره‌ساز				
[۲۶]	MARL	روزانه	✓	✓	✓	خورشیدی تولید ثابت	خودرو برقی به‌عنوان بار	✓	×	×	استفاده از شبکه عصبی برای پیش‌بینی‌ها
[۲۷]	DRL	روزانه	✓	✓	✓	خورشیدی تولید ثابت	شیمیائی و خودرو برقی	✓	×	×	مدیریت عدم‌قطعیت با پیش‌بینی ماشینی
[۲۸]	DRL	روزانه	✓	×	×	خورشیدی تولید ثابت	شیمیائی و حرارتی	✓	×	×	پمپ‌های حرارتی در ترکیب با خورشیدی
[۲۹]	MARL	روزانه	✓	✓	✓	خورشیدی تولید ثابت	شیمیائی	✓	×	×	آسایش حرارتی از طریق سیستم تهویه
[۳۰]	DRL	چندروزه	✓	✓	✓	×	×	✓	×	×	روش بازیگر- منتقد و لحاظ نایقینی قیمت
[۳۱]	DRL	روزانه	✓	✓	✓	خورشیدی تولید ثابت	شیمیائی و خودرو برقی	✓	×	×	بهینه‌سازی سیاست نزدیکی
[۳۲]	DRL	چندروزه	✓	×	×	خورشیدی تولید ثابت	شیمیائی	×	×	×	ترکیب با سیستم خبره قانون- محور
[۳۳]	MARL	روزانه	✓	✓	✓	خورشیدی تولید ثابت	شیمیائی و خودرو برقی	✓	×	×	ترکیب با بهینه‌سازی مبتنی بر نقطه داخلی
[۳۴]	DRL	روزانه	✓	✓	✓	خورشیدی تولید ثابت	شیمیائی و خودرو برقی	✓	×	×	لحاظ قیود بهره‌بردار (ظرفیت ترانسفورماتور)
[۳۵]	DRL	روزانه	✓	✓	✓	×	×	✓	×	×	بازخوردهای انسانی و فعالیته‌ای ساکنان
مقاله حاضر	MARL	روزانه	✓	✓	✓	خورشیدی با قطع تولید	خودرو برقی به‌عنوان بار	✓	✓	✓	مشارکت همزمان در انرژی و انعطاف‌پذیری

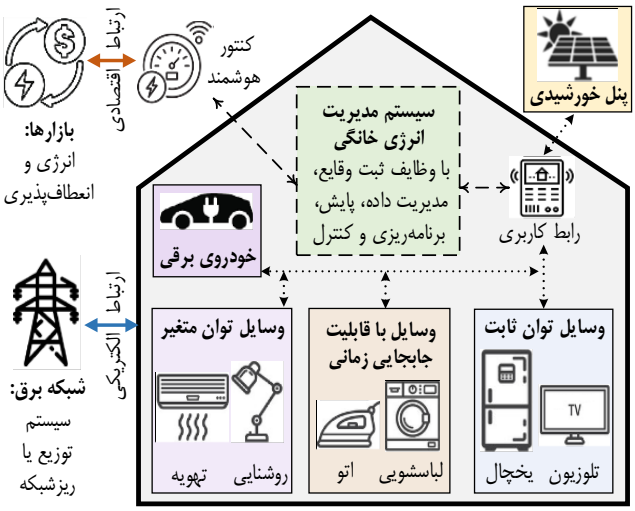
بهبود عملکرد RL با استفاده از رویکردهای حل خاص و روش‌های ترکیبی نیز مورد توجه ویژه بوده است. با توجه به پیچیدگی مدل‌ها و چالش مقیاس‌پذیری، برخی از مطالعات بر معماری‌های مختلف RL، مثل رویکرد بازیگر-منتقد ضمن مدیریت عدم‌قطعیت قیمت بازار [۳۰] و بهینه‌سازی سیاست نزدیکی جهت تطبیق بی‌درنگ با شرایط [۳۱]، برای حل مسئله تکیه کرده‌اند. روش‌های ترکیبی، همچون ادغام با سیستم‌های خبره مبتنی بر قواعد برای پیش‌بینی رفتار [۳۲] و استفاده از بهینه‌سازی سیاست بر اساس روش نقطه داخلی در کنار MARL [۳۳]، نیز در HEMSهای مدل- آزاد پیشنهادی در ادبیات تحقیق بکار رفته است. برخی از مطالعات دیگر نیز سعی کرده‌اند که مسائل خاص خانه و شبکه را در مسئله HEMS ادغام کنند. به‌عنوان نمونه، [۳۴] به‌طور خاص بر ادغام نیازهای بهره‌برداران شبکه (حفظ تعادل توان و قیود بهره‌برداری، خصوصاً بیشینه ظرفیت ترانسفورماتورهای بالادست) در مسئله HEMS تمرکز داشته است. مرجع [۳۵] نیز بازخوردهای انسانی و فعالیت‌های ساکنان خانه را در تصمیمات HEMS مبتنی بر RL گنجانده تا به‌طور پویا و بی‌درنگ توصیه‌هایی منطبق با اولویت‌های شناسایی‌شده ارائه دهد.

ابعاد مختلف مطالعات مرورشده HEMS بر پایه RL در جدول ۱ خلاصه شده و با مقاله حاضر به صورت مفهومی مورد مقایسه قرار گرفته است. مهم‌ترین نقد وارد به این مطالعات، غفلت از محصول انعطاف‌پذیری و عدم امکان مشارکت مستقیم خانه‌های هوشمند در بازار مربوطه است. مطالعاتی که به انعطاف‌پذیری خانه‌های هوشمند توجه کرده‌اند نیز، اولاً مبتنی بر RL نبوده و ثانیاً به آن به‌عنوان یک ویژگی جانبی پرداخته‌اند. خلأ مهم دیگر، صرفنظر از تولید قابل‌تغییر (اغلب متمرکز بر خورشیدی با تولید ثابت) و الزامات پیچیده‌ساز بالادستی مرتبط با آن (مثل سطح قطع تولید و جریمه برای آن) است. همچنین، مدلی جامع که انواع تجهیزات خانه را در بر بگیرد و نارضایتی کاربران را در مسئله بگنجاند، کمتر مورد توجه بوده است.

مقاله حاضر یک مدل HEMS مبتنی بر RL ارائه کرده که برای اولین بار بر مشارکت در انعطاف‌پذیری متمرکز است. پیاده‌سازی مبتنی بر MARL و به‌کارگیری شبکه‌های عصبی برای پیش‌بینی ورودی‌ها، امکان مدیریت انرژی مدل-آزاد را در محیط پویا خانه هوشمند ایجاد نموده است. همچنین بین سود اقتصادی (افزایش درآمد و کاهش هزینه) و عوامل پیچیده‌کننده (هزینه نارضایتی کاربران و جریمه تخطی از الزامات) تعادل برقرار می‌کند. بر این اساس، نوآوری‌ها اصلی مقاله حاضر عبارتند از:

- مدیریت انرژی روزانه بهینه خانه‌های هوشمند مبتنی بر یادگیری تقویتی چندعامله با لحاظ مشارکت همزمان در بازارهای انعطاف‌پذیری (روبه‌بالا و روبه‌پایین) و انرژی؛
- ادغام عامل پیچیده‌کننده نارضایتی کاربر ضمن امکان انتخاب اقتصادی محل مشارکت برای بارهای کنترل‌پذیر؛ و
- ادغام عامل پیچیده‌کننده الزام شبکه بالادست ضمن امکان انتخاب اقتصادی محل مشارکت برای تولید خورشیدی.

بخش دوم ابعاد مختلف مدل پیشنهادی، شامل چارچوب خانه هوشمند، توصیف مدل اجزاء (عامل‌ها) و رویکرد حل مسئله، را شرح می‌دهد. بخش سوم به نتایج شبیه‌سازی اختصاص دارد که علاوه بر توصیف ورودی‌ها، تحلیل‌های مهم را در قالب چهار مورد مطالعاتی متنوع ارائه و تشریح



شکل ۱: چارچوب مفهومی خانه هوشمند شامل اجزاء و ارتباطات.

می‌کند و در نهایت بخش چهارم به جمع‌بندی پرداخته است.

۲- مدل پیشنهادی

در این بخش ابتدا چارچوب و ساختار مدل پیشنهادی تبیین شده است. سپس، شرح جزئیات مدل مورداستفاده برای تجهیزات خانه هوشمند تشریح گردیده است. در انتها نیز رویکرد حل مدل توضیح داده شده است.

۲-۱ چارچوب و ساختار

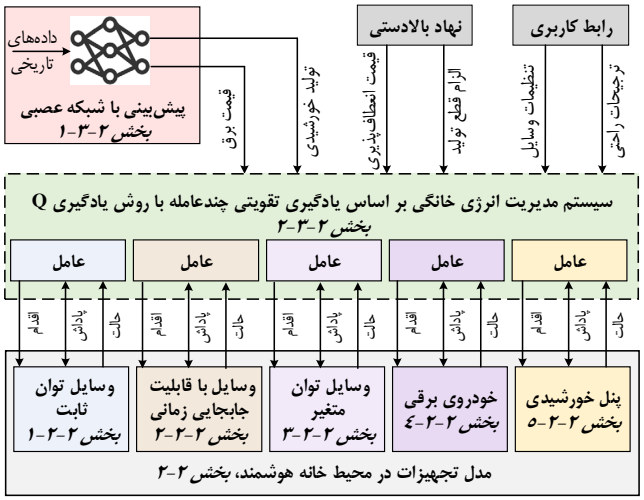
۲-۱-۱ چارچوب خانه هوشمند

شکل ۱ به فرمی مفهومی چارچوب اجزاء و ارتباطات خانه هوشمند مورد بررسی را به تصویر کشیده است. اجزاء کلیدی درون این خانه (مشخص‌شده با خاکستری)، وسائل توان ثابت (رنگ آبی) (مثل تلوزیون و یخچال)، وسائل با قابلیت جابجایی زمانی (رنگ نارنجی) (مثل اتو و لباسشویی)، وسائل توان متغیر (رنگ بنفش کمرنگ) (مثل روشنایی و تهویه)، خودروی برقی (رنگ بنفش پررنگ)، پنل خورشیدی (رنگ زرد) و HEMS (رنگ سبز) هستند. دلیل انتخاب پس‌زمینه هم‌رنگ برای وسائل توان متغیر و خودروی برقی این است که از وجه ذخیره‌سازی (تخلیه توان) این خودروها صرفنظر شده و از لحاظ مدل شبیه به بار توان متغیر هستند.

خانه هوشمند دو نوع ارتباط الکتریکی (با شبکه برق، سیستم توزیع یا ریزشبکه) و اقتصادی (با بازارهای انرژی و انعطاف‌پذیری) دارد. سیستم HEMS که وظایف کلیدی ثبت وقایع، مدیریت داده، پایش، برنامه‌ریزی و کنترل را بر عهده دارد نیز از طریق رابط کاربری با تجهیزات خانه ارتباط دوطرفه دارد. بعلاوه وظیفه گرفتن سیگنال‌های قیمتی از بازارها و مشارکت در آن‌ها نیز از طریق کنتور هوشمند بر عهده HEMS است.

۲-۱-۲ چارچوب اقتصادی مسئله

کلان‌روندهای مقررات‌زدایی و مردم‌سالاری سبب گسترش بازارهای محلی انرژی در شبکه‌های پایین‌دست شده است. این بازارها با بازیگرانی متنوع (بارها، مولدها، تجمیع‌کننده‌ها و...) و در سطوح مختلف (محله‌های کوچک و بزرگ تا سطح کل شبکه توزیع) اجرا می‌شوند. خانه‌های هوشمند با تجمیع سه وجه بار (تعیین زمان و اندازه مصرف)، تولید (فروش محصول با سود حداکثری) و ذخیره‌سازی (تعیین رژیم شارژ و دشارژ)، بازیگرانی کلیدی در این بازارها هستند. مهم‌ترین پیچیدگی مسئله تعیین برنامه بهینه مشارکت در این بازار از نگاه یک خانه، پیش‌بینی قیمت



شکل ۲: ساختار مسئله پیشنهادی شامل همه زیربخش‌ها.

دینامیک انرژی است که در مقاله حاضر فرض شده به کمک یک شبکه عصبی (یا ابزارهای مشابه) در دسترس HEMS است.

در کنار تبادل انرژی، چالش‌های فنی شبکه‌های پایین‌دست (کل توزیع تا ریزشبکه‌ها) نیز باید به کمک محصولات انعطاف‌پذیری رفع شود. در این راستا، قابلیت‌های خانه‌های هوشمند به کمک بارهای منعطف و تولید قابل‌قطع‌شان بسیار راهگشا است. این محصولات یا در قالب ساختاری رقابتی (مثلاً بازارهای محلی انعطاف‌پذیری با قیمتی دینامیک مشابه انرژی) تامین می‌شود یا مبتنی بر قیمت‌های تنظیم‌شده توسط نهاد بالادستی (اغلب بهره‌بردار شبکه) مبادله می‌شود. از نگاه مقاله حاضر مهم است که HEMS از قیمت محصولات انعطاف‌پذیری (تغییر توان در دوجهت بالا و پایین) مطلع باشد؛ حال یا مبتنی بر ابزارهای پیش‌بینی مشابه انرژی یا با دانستن قیمت تنظیم‌شده بالادستی.

۲-۱-۳ ساختار حل مسئله

شکل ۲ سعی کرده تا شفافیت مدل پیشنهادی را از منظر ورودی‌های مهم، زیربخش‌های اصلی و کلیات مسئله MARL افزایش دهد. این مدل بر پیش‌بینی قیمت برق (بازار انرژی) و تولید خورشیدی بر شبکه عصبی مبتنی بر داده‌های تاریخی (بخش ۲-۳-۱) متکی است. بعلاوه، فرض شده که شبکه بالادست برای رفع معضلات بهره‌برداری‌اش الزام قطع تولید را تعیین کرده که تخطی از آن مشمول جریمه خواهد شد. بعلاوه قیمت محصولات انعطاف‌پذیری رو به بالا و رو به پایین را هم اعلام می‌کند. در فرم‌های مدرن‌تر (مثلاً بازار محلی انعطاف‌پذیری) این قیمت‌ها هم باید با روشی (مثل شبکه عصبی) پیش‌بینی شوند. همچنین HEMS تنظیمات وسایل (مثل مقادیر دلخواه توان و بازه‌های مجاز کارکرد) و ترجیحات راحتی ساکنان (مثل ضرایب نارضایتی) را از طریق رابط کاربری خانه هوشمند دریافت می‌کند.

سیستم HEMS برای حل مبتنی بر یادگیری Q برای MARL نیازمند ارتباط با اجزاء پنج‌گانه خانه هوشمند است که مدل هریک به‌طور مجزا در بخش‌های ۲-۲-۱ تا ۲-۲-۵ ارائه شده است. HEMS با دیدن حالت‌ها از محیط خانه و سایر اطلاعات در دسترس ورودی، محاسبات عامل‌های مربوط به هریک اجزاء را انجام می‌دهد که عمدتاً متمرکز بر پاداش ترکیبی شامل وجوه مالی (خرید و فروش در بازار انرژی و مشارکت در بازارهای انعطاف‌پذیری) و ابعاد پیچیده‌کننده (راحتی ساکنان و رعایت الزامات بالادستی) است. این محاسبات در نهایت منجر به حصول اقدامات بهینه اجزاء (همان برنامه‌ریزی روزانه خانه هوشمند) خواهد شد.

۲-۲ تجهیزات خانه هوشمند

پیش از پرداختن به توصیف مدل هریک از اجزاء در زیربخش‌های پیش رو، حالت مسئله RL (مجموعه اطلاعات قابل‌مشاهده توسط عوامل جهت تصمیم‌گیری مناسب) برای هر ساعت در (۱) ارائه شده که حاوی شش نوع اطلاعات از ساعت‌های پیش رو است. مقدار پیش‌بینی‌شده برای ۱) قیمت برق (بازار انرژی) و ۲) توان تولیدی پنل خورشیدی، قیمت تعیینی انعطاف‌پذیری ۳) رو به پایین و ۴) روبه بالا توسط نهاد بالادستی، ۵) سطح قطع تولید الزام‌شده توسط شبکه و ۶) توان دلخواه تنظیمی برای بارهای توان متغیر توسط کاربر.

$$s_t = \begin{Bmatrix} (\lambda_t^G, \lambda_{(t+1)}^G, ..., \lambda_T^G), (P_t^{PV}, P_{(t+1)}^{PV}, ..., P_T^{PV}), \\ (\lambda_t^{dn}, \lambda_{(t+1)}^{dn}, ..., \lambda_T^{dn}), (\lambda_t^{up}, \lambda_{(t+1)}^{up}, ..., \lambda_T^{up}), \\ (\alpha_t^f, \alpha_{(t+1)}^f, ..., \alpha_T^f), (P_t^{PS,us}, P_{(t+1)}^{PS,us}, ..., P_T^{PS,us}) \end{Bmatrix} \quad (۱)$$

۲-۲-۱ وسایل توان ثابت

وسایل ثابت به دلیل نقش حیاتی‌شان در ایمنی یا راحتی خانه هوشمند، باید بدون تغییر توانی یا زمانی تامین انرژی شوند (یعنی روشن شدن مطابق برنامه مورد درخواست). بنابراین، امکان برنامه‌ریزی خاصی برای مدیریت انرژی آن‌ها وجود ندارد و نقشی هم در مشارکت در بازار انعطاف‌پذیری ندارند. هدف از مدل‌سازی این وسایل در این مقاله، لحاظ هزینه مصرف آن‌ها بابت خرید مقدار توان مشخص از بازار انرژی (پاداش منفی) مطابق (۲) در HEMS مبتنی بر MARL است.

$$r_{i,t}^{FA} = -\lambda_t^G \times P_{i,t}^{d,FA}; \forall \Omega^{FA}, \forall t \quad (۲)$$

۲-۲-۲ وسایل با قابلیت جابجایی زمانی

بارهای با قابلیت جابجایی نقشی کلیدی در تعیین الگوی مصرفی مناسب خانه‌های هوشمند در مقابل قیمت‌های پویای انرژی دارند. این بارها قادرند از دوره‌های قیمت بالا (معمولاً اوج مصرف) به دوره‌های با قیمت پایین‌تر (معمولاً بار سبک‌تر شبکه) منتقل شوند تا به کاهش هزینه خرید انرژی ضمن توجه به راحتی ساکنان کمک کنند. این لوازم معمولاً فقط دارای دو وضعیت اصلی (روشن و خاموش) هستند و تنها می‌توان در این خصوص در بازه تعیینی برای آن‌ها، با حفظ تعادل بین هزینه و نارضایتی، تصمیم گرفت. با توجه به اینکه اغلب هم وقفه‌پذیر (مثلاً توجیه‌ناپذیر بودن خاموش کردن ماشین لباسشویی هنگام کار) نیستند، برای مشارکت در انعطاف‌پذیری لحظه‌ای مناسب نیستند.

رابطه (۳) پاداش این نوع وسایل در هر برش زمانی ارائه کرده که ترکیبی از هزینه خرید انرژی و نارضایتی ناشی از جابجائی زمانی است. همان‌طور که دیده می‌شود، هزینه انرژی علاوه به قیمت برق به ساعت انتخابی برای روشن کردن وسیله نیز وابسته است. توجه شود که پس از روشن کردن، وسیله باید به مدت مشخص (از پیش معلوم) با توان معین کار کند. نارضایتی ساکنان نیز متناسب با توان دوم فاصله از زمان اولیه (زمان انتظار) فرض شده و با یک ضریبی تنظیمی (بر اساس ترجیحات شخصی کاربر) به جنس هزینه درآمده است.

$$r_{m,t}^{TS} = -[\lambda_t^G \times u_{m,t} \times P_{m,t}^{d,TS}] - [DC_m^{TS} \times (t_m^s - t_m^{ini})^2]; \quad (۳)$$
$$\forall m \in \Omega^{TS}, \forall t = [t_m^{ini}, t_m^{end}]$$

۲-۲-۳ وسایل توان- متغیر

این وسایل (مثل سیستم تهویه برای سرمایش و گرمایش در فصول گرم و سرد) به شکلی مدل شده‌اند که بتوانند در پله‌های توانی از پیش تعیین‌شده (گسسته) عملکرد داشته باشند. این ویژگی علاوه به واکنش نسبت به قیمت پویای انرژی، امکان مشارکت در محصولات انعطاف‌پذیری رو به بالا و پایین را نیز فراهم می‌کند. در مطالعات پیشین، مشارکت این تجهیزات در بازارهای انعطاف‌پذیری از طریق بازیگران میانی (مثل تجمیع‌کننده‌ها) و به شکلی غیرمستقیم لحاظ شده است. حال اینکه مشارکت مستقیم می‌تواند منجر به بالا رفتن درآمد آن‌ها و محقق‌کننده مردم‌سالاری انرژی باشد. پاداش مربوط به این وسایل در (۴) ارائه شده که شامل هزینه‌های خرید انرژی، درآمد از انواع محصولات انعطاف‌پذیری و وجه نارضایتی ساکنان خانه است.

اولین بخش از این رابطه به قیمت پویای برق و توان متغیر تجهیز در هر لحظه وابسته است تا هزینه تامین انرژی را لحاظ کرده و کاهش دهد. توجه شود که مشارکت فعالانه در بازار انرژی به نوعی همان پاسخگویی بار سنتی مبتنی بر قیمت است. با این حال، بهره‌بردار شبکه برای مدیریت چالش‌های باقیمانده بهره‌برداری، ناچار است که محصولات انعطاف‌پذیری (مشابه رزرو انتقال) را تعریف و تامین کند. دو نوع محصول انعطاف‌پذیری رو به بالا (قابلیت کاهش لحظه‌ای مصرف) و رو به پایین (قابلیت افزایش لحظه‌ای مصرف) با قیمت مشخص نیز لحاظ شده است. طبیعتاً ظرفیت کاهش و افزایش میزان مصرف به فاصله توان لحظه‌ای تجهیز تا مقادیر کمینه و بیشینه توان ممکن وسیله وابسته است.

به جز هزینه‌ها و درآمدهای ناشی از مشارکت در بازارها، جزء دیگر این مدل، هزینه نارضایتی ساکنان است که با توان دوم فاصله مصرف تجهیزات با انتظارات ساکنان (مقادیر تنظیمی و موردانتظار کاربر) متناسب است. این نارضایتی به کمک ضریب هزینه تنظیمی در رابط کاربری تبدیل به مقدار دلاری شده تا با سایر اجزاء رابطه پاداش جمع‌پذیر باشد.

$$r_{j,t}^{PS} = [-\lambda_t^G \times P_{j,t}^{d,PS}] - [DC_j^{PS} \times (P_{j,t}^{d,PS} - P_{j,t}^{us,PS})^2] + [\lambda_t^{dn} \times (P_{j,t}^{d,PS} - P_j^{\min,PS})] + [\lambda_t^{up} \times (P_j^{\max,PS} - P_{j,t}^{d,PS})]; \quad (۴)$$
$$\forall j \in \Omega^{PS}, \forall t$$

۲-۲-۴ خودروی برقی

در این مقاله، بدون از دست دادن عمومیت مسئله، از تخلیه باتری خودروهای الکتریکی به دلایلی همچون اثرات منفی آن بر عمر و عملکرد تجهیز صرفنظر شده است. بنابراین این خودروها از لحاظ مفهومی مشابه بارهای توان متغیر مدل می‌شوند. بر این اساس خودروی برقی می‌تواند مصرف انرژی خود را برای انطباق با قیمت‌های کمتر بازار انرژی مدیریت کند و با توجه به امکان وقفه‌پذیر بودن شارژش و قابلیت کاهش و افزایش گسسته توانی، در بازارهای انعطاف‌پذیری نیز مشارکت داشته باشد. البته که باید به مسئله نارضایتی با توجه به فاصله از توان موردانتظار شارژ نیز توجه شود. بر اساس این توضیحات، پاداش این تجهیز در (۵) نمایش داده شده که ترکیبی کاملاً مشابه با بارهای توان متغیر، یعنی (۴)، دارد.

$$r_{n,t}^{EV} = [-\lambda_t^G \times P_{n,t}^{d,EV}] - [DC_j^{EV} \times (P_{n,t}^{d,EV} - P_{n,t}^{us,EV})^2] + [\lambda_t^{dn} \times (P_{n,t}^{d,EV} - P_n^{\min,EV})] + [\lambda_t^{up} \times (P_n^{\max,EV} - P_{n,t}^{d,EV})]; \quad (۵)$$
$$\forall j \in \Omega^{EV}, \forall t = [t^{arr}, t^{dep}]$$

۲-۲-۵ پنل خورشیدی

پنل‌های خورشیدی قادرند تا بخشی از انرژی موردنیاز خانه‌های هوشمند را تأمین کرده و حتی انرژی مازاد خود را به شبکه بفروشند. بر این اساس، علاوه بر نقش کلیدی در کربن‌زدائی و اقتصاد نوین انرژی، در

راستای کاهش وابستگی به شبکه (تحقق مردم‌سالاری و تمرکززدایی) نیز بسیار با اهمیت هستند. علاوه بر استفاده اقتصادی از قیمت پویای انرژی در کنار ذخیره‌سازها، پنل‌های خورشیدی می‌توانند با دنبال نکردن نقطه بیشینه توان خود (قطع تولید خودخواسته)، امکان مشارکت در بازارهای انعطاف‌پذیری رو به بالا و رو به پایین را نیز فراهم کنند. البته باید توجه داشت که بهره‌برداران شبکه، الزام مشخصی در این باره را برای شرایط پرچالش (مثلاً نفوذ لحظه‌ای بالای تجدیدپذیرها)، اجبار می‌کنند.

بر خلاف اغلب مقالات که پنل‌های خورشیدی را به شکل توان ثابت (مبتنی بر پیش‌بینی) مدل کرده‌اند، در این مقاله پله‌های توان تولیدی متغیر برای آن‌ها لحاظ شده است. رابطه (۶) پاداش پنل خورشیدی را در هر برش زمانی بر اساس توضیحات فوق ارائه کرده است. با فاصله گرفتن از بیشینه توان قابل استحصال (مقدار پیش‌بینی‌شده برای تولید با توجه به شرایط جوی) و از دست دادن بخشی از درآمد ناشی از فروش انرژی، امکان مشارکت در انعطاف‌پذیری رو به بالا (قابلیت افزایش توان تولیدی تا نقطه بیشینه) و رو به پایین (قابلیت کاهش توان تا مقدار کمینه) و کسب درآمد از آن‌ها فراهم می‌شود. البته که در کنار درآمدهای ناشی از بازارهای انرژی و انعطاف‌پذیری، باید به عامل پیچیده‌کننده‌ای الزام بالادستی قطع تولید و جریمه تخطی از آن نیز توجه داشت. مشابه نارضایتی کاربران در وسایل توان متغیر، برای تخطی از الزام بالادستی نیز یک رابطه درجه دوم متناسب با میزان تخطی فرض شده است.

$$r^{PV}_{k,t} = [\lambda^{G}_{t} \times \alpha_{k,t} \times P^{PV}_{k,t}] - [DC^{PV}_{k} \times (\alpha_{k,t} - \alpha^{f}_{k,t})^{2}] + [\lambda^{dn}_{t} \times (1-\alpha_{t}) \times P^{PV}_{k,t}] + [\lambda^{up}_{t} \times \alpha_{t} \times P^{PV}_{k,t}]; \quad \forall j \in \Omega^{EV}, \forall t = [t^{arr}, t^{dep}] \qquad (۶)$$

۲-۳ رویکرد حل

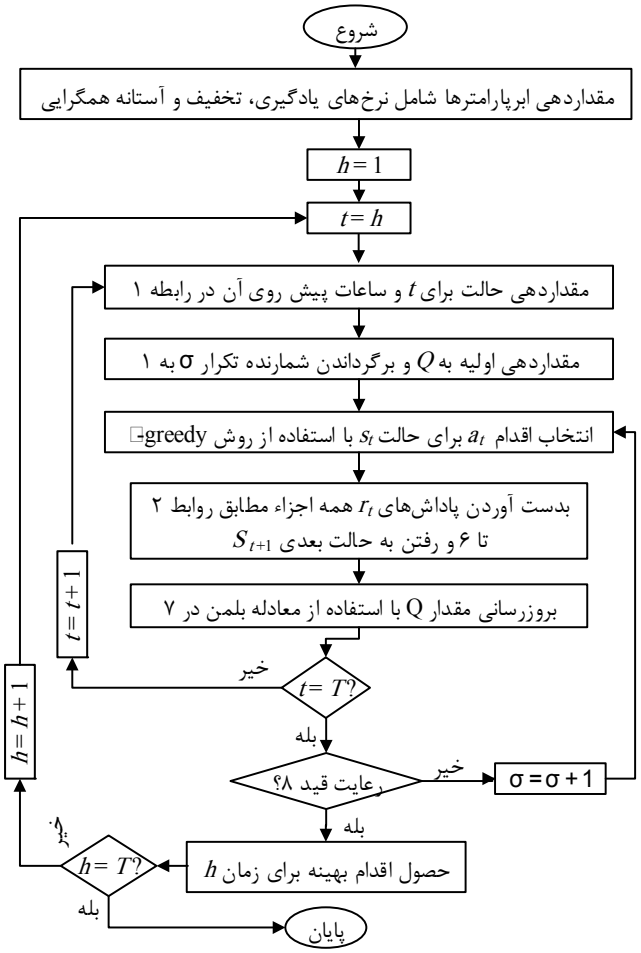
در این بخش دو ابزار کلیدی داده-محور حل مسئله، یعنی شبکه عصبی برای پیش‌بینی و یادگیری Q برای HEMS، توضیح داده شده است.

۲-۳-۱ شبکه عصبی برای پیش‌بینی

شبکه‌های عصبی یکی از مهم‌ترین روش‌های یادگیری ماشین بوده که با توجه به ساختار ساده و توانایی‌شان در یادگیری الگوهای پیچیده، در مسائل متنوعی کاربرد دارند. این شبکه‌ها از لایه‌های ورودی و خروجی بعلاوه چندین لایه پنهان (نمایانگر عمق شبکه) تشکیل شده که هر یک نقش ویژه‌ای در پردازش اطلاعات و استخراج ویژگی‌های مهم ایفا می‌کنند. فرایند یادگیری در این شبکه‌ها شامل اجرای رفت و برگشتی دو مرحله پیش‌خور (حرکت از نودهای ورودی به سمت نودهای خروجی با استفاده آخرین وزن‌های محاسبه‌شده و بهره‌گیری از توابع فعال‌سازی مثل سیگموید) و انتشار معکوس خطا (محاسبه تابع تلفات و به‌روزرسانی وزن‌ها با استفاده از مشتقات جزئی مبتنی بر الگوریتم گرادیان کاهشی) است. در مقاله حاضر برای پیش‌بینی قیمت انرژی (با استفاده از داده‌های تاریخی) و تولید پنل خورشیدی (با بهره‌گیری از داده‌های مربوط به وضعیت آب‌وهوا و شرایط محیطی) از شبکه عصبی [۲۶] استفاده شده و با توجه به عدم نوآوری در این‌باره، از توضیحات بیشتر فنی اجتناب شده است.

۲-۳-۲ یادگیری Q برای MARL

در فلسفه کلی MARL، هر عامل با انجام اقدامات مختلف در یک محیط پویا و دارای عدم‌قطعیت، تغییر حالت‌ها را ارزیابی کرده و با مشاهده پاداش منتجه، سیاست بهینه خود را به تدریج می‌آموزد. در این مقاله از



شکل ۳: روندنمای حل MARL به کمک یادگیری Q.

الگوریتم یادگیری Q برای به‌دست آوردن بیشینه پاداش تجمعی اجزا خانه هوشمند در کل دوره زمانی روزانه (۲۴ ساعته) استفاده شده است. فرایند اصلی در این الگوریتم، ساختن جدول Q و به‌روزرسانی آن تا برآورده شدن شرط همگرایی است که روندنمای آن در شکل ۳ نشان داده شده است.

رابطه (۷) که به معادله بلمن مشهور است، نحوه به‌روزرسانی مقدار Q را در هر تکرار نشان می‌دهد. نرخ یادگیری θ میزان تاثیر Q جدید را بر روی مقدار قدیمی آن نشان می‌دهد. در شرایط صفر بودن این نرخ، عامل فقط از اطلاعات قبلی استفاده کرده و زمانی که این مقدار یک است، عامل اطلاعات قبلی را نادیده گرفته و فقط برآورد فعلی را در به‌روزرسانی لحاظ می‌کند. نرخ تخفیف γ نیز، اهمیت نسبی نگاه آینده‌نگرانه (تمرکز بر پاداش آتی) نسبت به نگاه فرصت‌طلبانه (تاکید بر پاداش حال) است.

برای حصول اقدام بهینه در هر ساعت (با اندیس h مشخص شده) به‌روزرسانی Q به صورت تکراری (با شمارنده σ) مبتنی بر الگوریتم حریصانه ε و با ارزیابی تمامی ساعت‌های فعلی و بعدی (از $t = h$ تا $t = T$) تا رسیدن به شرط همگرایی (۸) (عدم بهبود در آخرین تکرار با مقایسه به مقدار آستانه معین τ) انجام می‌شود. مقدار مناسب ابرپارامترهای کلیدی θ (۰/۹)، γ (۰/۹)، ε (۰/۳) و τ (۰/۰۰۱) با آزمون و خطا تعیین شده است.

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \theta \times [r(s_t, a_t) + \gamma \times \max(Q(s_{t+1}, a_{t+1})) - Q(s_t, a_t)] \qquad (۷)$$

$$|Q^{\sigma} - Q^{\sigma-1}| \leq \tau \qquad (۸)$$

جدول ۲: مشخصات تجهیزات خانه هوشمند.

نوع	نام	مخفف انگلیسی	مجموعه اقدام		دوره استفاده	ضریب نارضایتی
توان ثابت	یخچال هشدار	REF AS	۱ (همیشه روشن)		۲۴ ساعته	–
قابل جابجایی	لباسشویی ظرفشویی	WM DW	۰ یا ۱ (خاموش یا روشن)		۷ تا ۱۰ شب ۸ تا ۱۰ شب	۰.۱۰ ۰.۰۶
توان متغیر	روشنایی گرمایش تهویه	L H AC	۰.۲، ۰.۳، ... و ۰.۶ ۰.۵، ۰.۶، ... و ۱.۵ ۰.۷، ۰.۸، ... و ۱.۴	توان‌های بار	۲۴ ساعته	۰.۰۲ ۰.۱۲ ۰.۱
خودرو برقی		EV	۰، ۳ و ۶ (توان‌های شارژ)		۰ بامداد تا ۶ صبح	۰.۰۴
پنل خورشیدی		PV	۰.۶، ۰.۷، ... و ۱ (ضرایب تولید)		۶ صبح تا ۸ شب	۰.۵۰

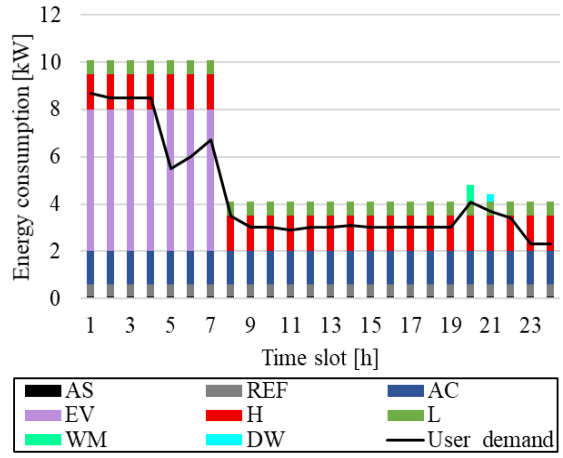
۳- نتایج عددی

۳-۱ ورودی‌ها، موارد مطالعاتی و بستر حل

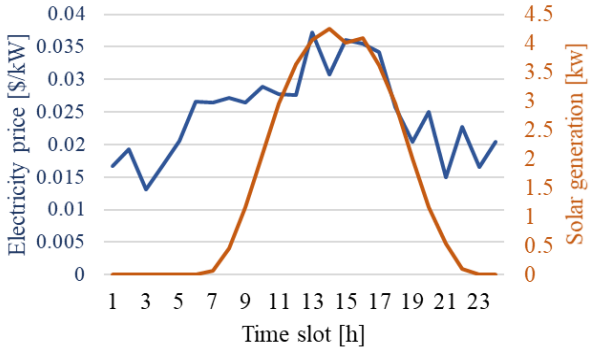
در این بخش شبیه‌سازی‌های متنوعی بر روی یک خانه هوشمند دارای یک پنل خورشیدی، دو وسیله توان ثابت (یخچال و سیستم هشدار)، دو وسیله با قابلیت جابجایی زمانی (ماشین‌های لباسشویی و ظرفشویی)، سه وسیله توان متغیر (سیستم‌های روشنایی، گرمایش و تهویه) و یک خودرو برقی انجام شده تا کارایی مدل پیشنهادی نشان داده شود. مشخصات کلیدی تجهیزات مختلف در جدول ۲ خلاصه شده و الگوی مصرف پایه خانه نیز در شکل ۴ نمایش داده شده است. خروجی شبکه عصبی برای پیش‌بینی روز- پیش قیمت برق و تولید خورشیدی هم در شکل ۵ به تصویر کشیده شده است. با توجه به اهمیت قیمت انعطاف‌پذیری‌های رو به بالا و پایین، سناریوهای مطالعاتی بر اساس ضرایب مختلف این قیمت‌ها نسبت به قیمت انرژی توسعه داده شده است.

در ادامه نتایج حاصله از موارد مطالعاتی مختلف ارائه شده است. مورد بدون انعطاف‌پذیری، مصرف انرژی خانه را متناسب با هزینه (صرفاً مشارکت در بازار انرژی)، نارضایتی ساکنین و جریمه الزام بالادستی بهینه می‌کند. موارد مطالعاتی لحاظ‌کننده انعطاف‌پذیری، در کنار هزینه انرژی و نارضایتی، اثر متقابل بازارهای انرژی و انعطاف‌پذیری بر یکدیگر را برجسته می‌سازد که برای دو نوع روبه بالا و روبه پایین نیز متفاوت است. در این موارد، تحلیل حساسیت نتایج به پارامترهای کلیدی نیز انجام شده و اثر همزمان این دو بازار انعطاف‌پذیری در کنار بازار انرژی نیز به‌عنوان مورد مطالعاتی آخر بررسی شده است.

پیاده‌سازی روش حل MARL به کمک زبان برنامه‌نویسی پایتون و بر اساس کتابخانه‌های numpy و pandas بر روی یک سیستم سخت‌افزاری با پردازنده Intel core i3 به همراه ۱۲ GB حافظه RAM انجام شده است. گفتنی است که استفاده از روش MARL نسبت به بهینه‌سازی کلاسیک (روش مدل-محور)، بهبودی از منظر بهینگی ایجاد نمی‌کند. در حقیقت مهم است که روش‌های داده-محور همچون MARL بتوانند به پاسخی مشابه با نقطه نهایی یافته‌شده توسط بهینه‌سازی همگرا شوند که در تمامی موارد مطالعاتی اجراشده چنین بوده است. از سویی دیگر، مزیت اصلی MARL در عدم‌نیاز به اطلاعات دقیق از جزئیات جهت مدل‌سازی در کاربردهای زمان‌واقعی، برای محیط‌های پویا و پرچالشی مثل خانه‌های هوشمند است. در واقع، مدل‌های کلاسیک



شکل ۴: الگوی بار پایه خانه هوشمند شامل همه تجهیزات.



شکل ۵: پروفیل پیشبینی‌شده برای قیمت برق و تولید خورشیدی.

دنبال‌کننده نقطه بهینه سراسری، تنها به کمک به‌روزرسانی مدل در شرایط متغیر و پیچیده (منجر به بار پردازشی بالا)، کارامد باقی خواهند ماند. حال اینکه MARL پس از یادگیری اولیه، بدون نیاز به اطلاع از جزئیات و بدون تحمیل بار محاسباتی قابل‌توجه، استراتژی بهینه HEMS را با تطبیق مداوم خود با محیط، به شکلی تدریجی به‌روزرسانی می‌کند.

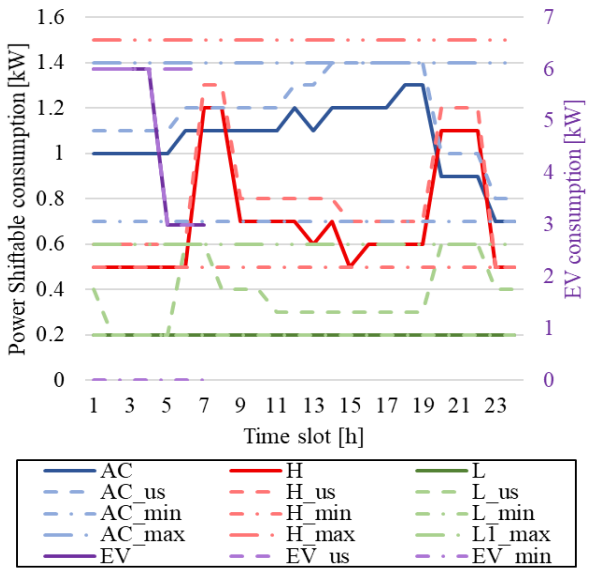
۳-۲ HEMS بدون لحاظ بازار انعطاف‌پذیری

۳-۲-۱ بدون لحاظ نارضایتی و الزام شبکه (بدون انعطاف‌پذیری)

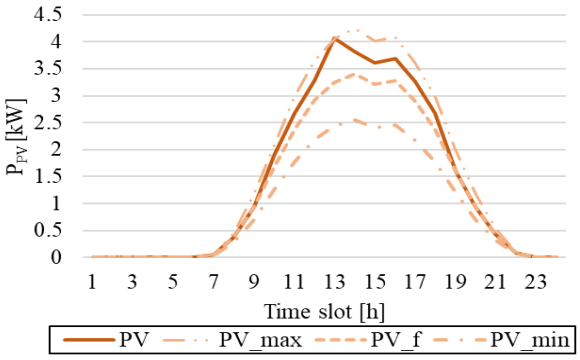
پاسخ HEMS در شرایطی که از بازار انعطاف‌پذیری صرفنظر شده و عوامل پیچیده‌کننده نارضایتی و الزام شبکه نیز لحاظ نشده‌اند، بدیهی است: بهینه‌سازی اقتصادی با توجه به قیمت‌های پویای بازار انرژی. به این منظور، نتیجه به سمت افزایش درآمد (بیشینه تولید پنل خورشیدی بدون توجه به الزام قطع تولید) و کاهش هزینه مصرف انرژی خواهد رفت. برای کاهش هزینه، نگاه در انواع مدل بار متفاوت هستند. بارهای توان ثابت که اساساً تغییری نمی‌کنند و بارهای قابل جابجایی به برش زمانی مجاز (در بازه مشخص‌شده) با کمترین قیمت انرژی (ساعت ۲۱) منتقل می‌شوند. بارهای توان متغیر و خودروی برقی نیز در حداقل توان ممکن‌شان (بدون توجه به مسئله نارضایتی کاربران) برنامه‌ریزی می‌شوند. با توجه به اینکه نتایج بارهای توان ثابت و قابل جابجایی زمانی (نظر به عدم امکان شرکت در انعطاف‌پذیری) در سایر موارد مطالعاتی تغییر نمی‌کند، در ادامه تنها به بارهای توان متغیر (و البته خودروی برقی که از لحاظ مفهومی مشابه آن‌ها است) و پنل خورشیدی تمرکز شده است.

۳-۲-۲ با لحاظ نارضایتی و الزام شبکه (بدون انعطاف‌پذیری)

کاهش هزینه (یا مصرف) با راحتی ساکنان و افزایش درآمد (یا تولید) با



شکل ۶: توان بهینه بارهای توان متغیر و خودروی برقی در شرایط بدون انعطاف‌پذیری به همراه مقادیر کمینه، دلخواه و بیشینه توان این تجهیزات.



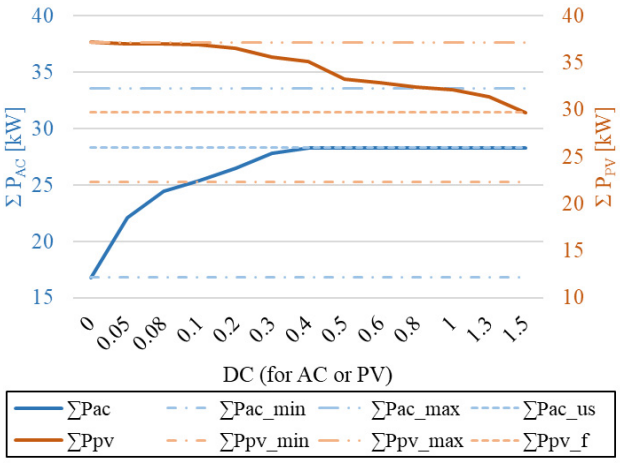
شکل ۷: توان بهینه پنل خورشیدی در شرایط بدون انعطاف‌پذیری به همراه مقادیر کمینه، الزام بالادستی و بیشینه توان این تجهیز.

جریمه الزام تخطی از الزام بالادستی در تعارض هستند. لحاظ ضرایب نارضایتی (یا جریمه) در پاداش تجهیزات، به برقراری تعادل بین بحث اقتصادی و این وجوه پیچیده‌کننده کمک می‌کند. شکل‌های ۶ و ۷ نتیجه HEMS در این شرایط را به تصویر کشیده‌اند که در آن‌ها مصرف وسایل توان متغیر از حداقل ممکن به سمت تنظیمات کاربری (حد رضایت) تمایل پیدا کرده و تولید خورشیدی نیز از مقدار حداکثر قابل‌تولید به سمت الزام شبکه تمایل پیدا کرده است.

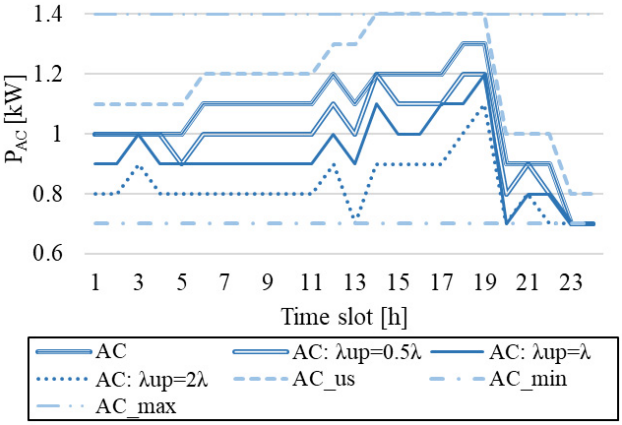
وسایل مختلف با توجه به ضریب نارضایتی متفاوت، رفتار متنوعی در مواجهه با قیمت انرژی نشان داده‌اند. روشنایی (ضریب نارضایتی کوچک) اصلاً متأثر نشده (کمینه توان) و خودروی برقی (ضریب نارضایتی بزرگ) با تنظیمات کاربری هم راستا شده (کمینه نارضایتی) است. برای سایر تجهیزات، با توجه به متعادل بودن ضرایب نارضایتی، مقدار مصرف جایی بین کمینه (کاهش هزینه) و تنظیمات کاربری (کاهش نارضایتی) قرار گرفته است. در ادامه، جهت خلوت‌تر شدن نتایج، تحلیل‌ها فقط بر روی سیستم تهویه به عنوان یک نمونه از تجهیزات انجام شده است. ادغام جریمه تخطی از الزام بالادست (سطح ۸۰٪ مطلوب)، سبب کاهش تولید پنل خورشیدی از مقدار بیشینه به میانگینی حدود ۹۰٪ شده است.

۳-۲-۳ تحلیل حساسیت به ضرایب نارضایتی و جریمه

نتیجه تحلیل حساسیت مجموع ۲۴ساعته توان مصرفی (AC به‌عنوان نمونه) و تولید (PV) خانه نسبت به ضرایب نارضایتی در شکل ۸ نشان



شکل ۸: تحلیل حساسیت مجموع بار (تهویه) و تولید روزانه به ضرایب نارضایتی (یا جریمه) در مورد مطالعاتی بدون انعطاف‌پذیری.



شکل ۹: توان AC در شرایط قیمتی مختلف انعطاف‌پذیری رو به بالا.

داده شده است. تهویه در مقدار ضریب نارضایتی صفر حداقل مصرف را دارد (جهت کاهش هزینه) و با افزایش آن تا ۰.۴، مصرفش دقیقاً هم‌راستا تنظیم کاربر (جهت راحتی) شده و بعد از آن طبیعتاً دیگر با افزایش ضریب نارضایتی تغییری در مصرف اتفاق نمی‌افتد. برای پنل خورشیدی نیز در ضریب جریمه صفر تولید حداکثر (بیشینه درآمد) شده و با افزایش این میزان تا ۱.۵، مقدار تولید با الزام شبکه هم‌راستا شده است.

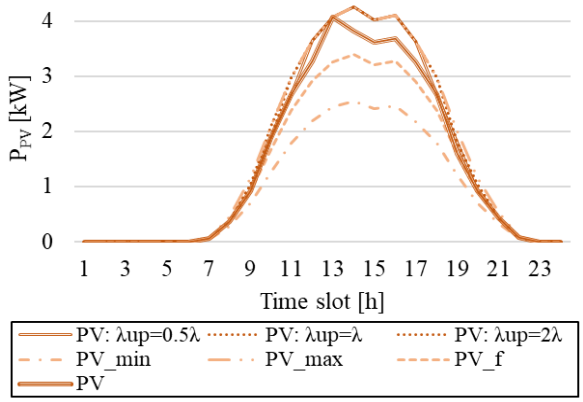
۳-۳ HEMS با لحاظ بازار انعطاف‌پذیری رو به بالا

۳-۳-۱ بدون عوامل پیچیده‌کننده (با انعطاف‌پذیری رو به بالا)

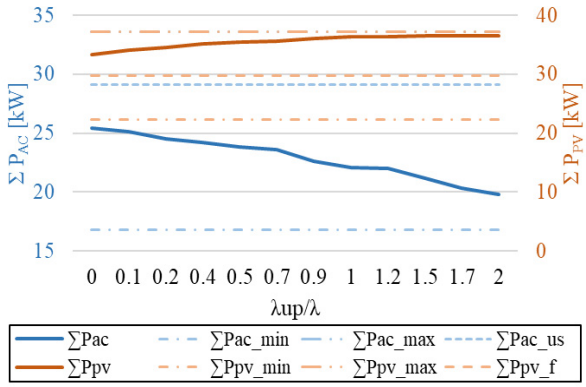
بازارهای انرژی و انعطاف‌پذیری رو به بالا از نگاه فنی تجهیزات همسو با هم هستند؛ یعنی هرچه بار کمتر (کاهش هزینه انرژی) و تولید بیشتر (افزایش درآمد انرژی) باشد، ظرفیت بیشتری برای مشارکت در انعطاف‌پذیری روبه بالا (افزایش بار و کاهش تولید) موجود است. در صورت صرفنظر از عوامل پیچیده‌کننده، قیمت این محصول هرچقدر هم که باشد، پنل خورشیدی به بیشینه توان تولیدی خود و تهویه به کمینه مقدار مصرفش می‌چسبد و تنها مجموع پرداختی و دریافتی خانه نسبت به شرایط بخش ۳-۲-۱ (فقط بازار انرژی) کمتر و بیشتر می‌شود.

۳-۳-۲ با عوامل پیچیده‌کننده (با انعطاف‌پذیری رو به بالا)

شکل‌های ۹ و ۱۰ به ترتیب توان بهینه مصرفی تهویه و تولید پنل خورشیدی را در شرایط قیمتی مختلف انعطاف‌پذیری رو به بالا در حضور عوامل پیچیده‌کننده نشان می‌دهند. با در نظر گرفتن ضرایب نارضایتی،



شکل ۱۰: توان PV در شرایط قیمتی مختلف انعطاف‌پذیری رو به بالا.



شکل ۱۱: تحلیل حساسیت نسبت به قیمت انعطاف‌پذیری رو به بالا.

تعادل بین وجه اقتصادی (کاهش بار و افزایش تولید برای برای نفع در هر بازار همسوی انرژی و انعطاف‌پذیری رو به بالا) و فنی (مصرف مطلوب با توجه به آسایش ساکنین و قطع تولید الزام‌شده)، به سمت دومی سبک‌تر شده است. در مقایسه با بخش ۳-۲-۲ (فقط بازار انرژی)، به افزایش درآمد ناشی از مشارکت در انعطاف‌پذیری رو به بالا، تولید به مقدار بیشینه و مصرف به مقدار کمینه نزدیک‌تر شده‌اند.

با افزایش قیمت انعطاف‌پذیری رو به بالا، تمایل به سمت وجه اقتصادی بیشتر شده و سهم نارضایتی و الزام شبکه کم‌رنگ‌تر شده است. میزان ۲۴ساعته مصرف تهویه و تولید پنل خورشیدی بدون لحاظ انعطاف‌پذیری رو به بالا، به ترتیب ۲۵٫۴ و ۳۳٫۳ کیلووات-ساعت بود. در شرایط شرکت در بازار انعطاف‌پذیری دارای قیمت دوبرابری نسبت به انرژی، این مقادیر به ۱۹٫۸ و ۳۶٫۵ کیلووات- ساعت رسیده‌اند.

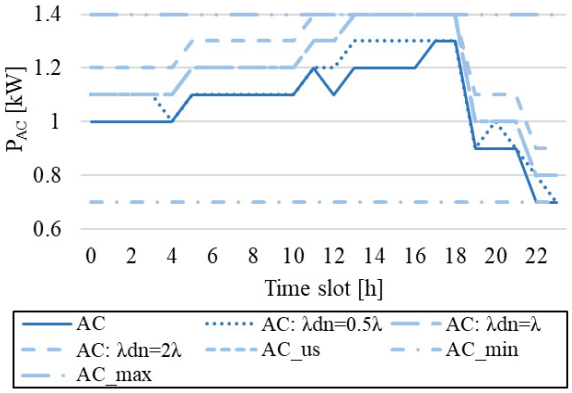
۳-۳-۳ تحلیل حساسیت به قیمت انعطاف‌پذیری رو به بالا

در بخش ۳-۲-۳ نشان داده شده که با افزایش ضرایب نارضایتی، مصرف و تولید به‌ترتیب به‌سمت توان‌های دلخواه کاربر و الزام‌شده بالادست حرکت می‌کنند. شکل ۱۱ روندی معکوس را با بیشتر شدن قیمت محصول انعطاف‌پذیری رو به بالا نشان می‌دهد. این مسئله کاملاً مشابه تأثیر افزایش قیمت انرژی بر رفتار تجهیزات بوده که با افزایش قیمت این محصول انعطاف‌پذیری، مصرف به سمت کمینه و تولید به سمت بیشینه سوق پیدا کرده که متعاقب آن نقش آسایش کاربر و جریمه تخطی از الزام بالادست در تصمیم‌گیری کم‌تر شده است.

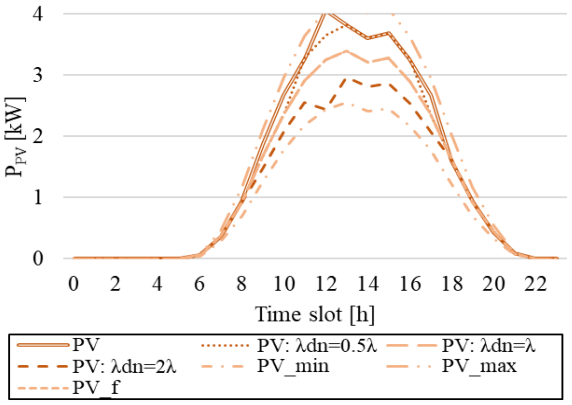
۳-۴ HEMS با لحاظ بازار انعطاف‌پذیری رو به پایین

۳-۴-۱ بدون عوامل پیچیده‌کننده (با انعطاف‌پذیری رو به پایین)

بر خلاف بخش ۳-۳-۱، بازارهای انرژی و انعطاف‌پذیری رو به پایین از



شکل ۱۲: توان AC در شرایط قیمتی مختلف انعطاف‌پذیری رو به پایین.

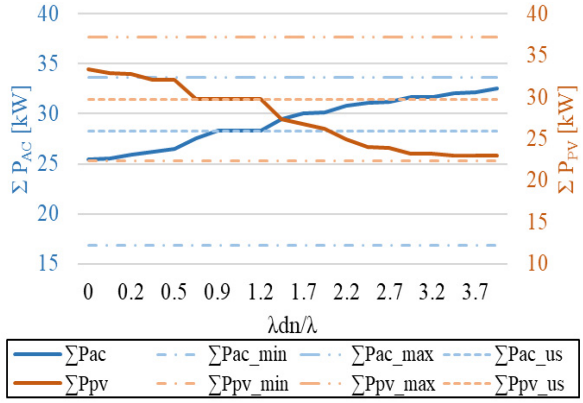


شکل ۱۳: توان PV در شرایط قیمتی مختلف انعطاف‌پذیری رو به پایین.

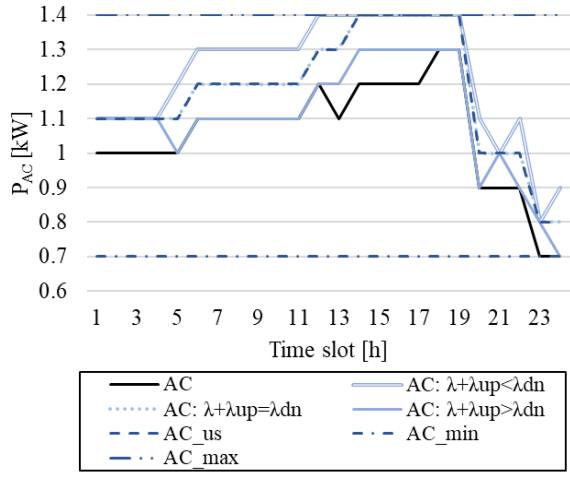
نظر فنی خلاف‌سو هستند؛ یعنی هرچه بار بیشتر (افزایش هزینه انرژی) و تولید کمتر (کاهش درآمد انرژی) باشند، ظرفیت مشارکت در انعطاف‌پذیری رو به پایین (کاهش بار و افزایش تولید) بیشتر است. در واقع، فرصت‌های تجاری این دو بازار با یکدیگر در تعارض هستند و تصمیم بهینه به تناسب میان قیمت این دو محصول بستگی دارد (البته بدون لحاظ عوامل پیچیده‌کننده). وقتی قیمت انعطاف‌پذیری رو به پایین به ترتیب کمتر و بیشتر از قیمت انرژی است، سیستم تهویه به ترتیب به بیشینه و کمینه مصرف خود چسبیده و پنل خورشیدی به ترتیب در کمینه و بیشینه تولید قرار می‌گیرد. در شرایط قیمت‌های مساوی این دو بازار، تعیین بار و تولید بهینه تصادفی است؛ چون هر تغییر توانی اثر اقتصادی کاملاً خلاف جهتی در دو بازار داشته که یکدیگر را خنثی نموه و سبب می‌شود همه مقادیر برای HEMS بهینه محلی (پاداش یکسان) باشند.

۳-۴-۲ با عوامل پیچیده‌کننده (با انعطاف‌پذیری رو به پایین)

با در نظر گرفتن ضرایب نارضایتی مربوط به عوامل پیچیده‌کننده، مقادیر بهینه مصرف تهویه و تولید پنل خورشیدی با لحاظ قیمت‌های مختلف انعطاف‌پذیری رو به پایین، مطابق شکل‌های ۱۲ و ۱۳ است. مشهود است که تعادل بین وجه اقتصادی (تضاد بین بازارهای انرژی و انعطاف‌پذیری) و فنی (آسایش کاربر یا تعهد به الزام شبکه)، به سمت دومی سنگین شده است. به عنوان نمونه، با قیمت انعطاف‌پذیری رو به پایین معادل نصف قیمت انرژی، تولید پنل بیشتر از مقدار بیشینه دور شده و مصرف سیستم تهویه از مقدار کمینه فاصله بیشتری گرفته است تا ضمن افزایش درآمد به واسطه مشارکت در انعطاف‌پذیری رو به پایین، مقادیر مصرف و تولید به ترتیب به میزان مصرف دلخواه کاربر و تولید الزام‌شده بالادست نزدیک شوند. مجموع ارزش افزوده ناشی از افزایش درآمد انعطاف‌پذیری و کاهش نارضایتی باید بتواند ضرر در بازار انرژی (افزایش هزینه و کاهش درآمد) را



شکل ۱۴: تحلیل حساسیت نسبت به قیمت انعطاف‌پذیری رو به پایین.



شکل ۱۵: توان AC در در حضور هر دو محصول انعطاف‌پذیری.

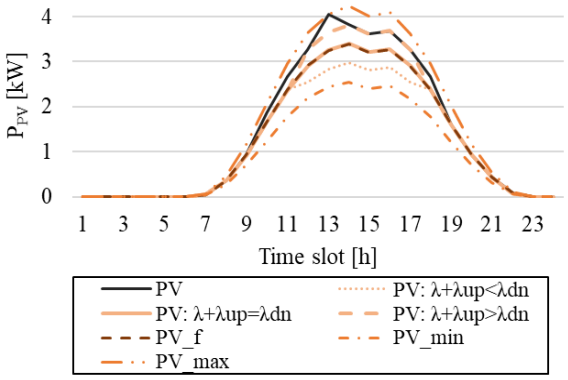
پوشش دهد.

با بالا رفتن قیمت محصول انعطاف‌پذیری رو به پایین، این نقش برجسته‌تر شده و در شرایط مرزی (قیمت یکسان انعطاف‌پذیری رو به پایین و انرژی)، اثر اقتصادی این بازارها همدیگر را خنثی کرده و توان بهینه سیستم تهویه و پنل خورشیدی دقیقاً معادل مصرف دلخواه کاربر و الزام تولیدی بالادستی خواهد بود. این حرکت با بالاتر رفتن قیمت انعطاف‌پذیری رو به پایین همچنان ادامه دارد؛ به‌طوری‌که میزان تولید و مصرف از حدود تعیینی الزام شبکه و تنظیم کاربر عبور کرده و به سمت کمینه تولید و بیشینه مصرف می‌رود. در این شرایط، مشارکت در بازار انعطاف‌پذیری رو به پایین، نه‌تنها بر ضرر ناشی از عدم مشارکت در بازار انرژی چیره شده بلکه جریمه ناشی از نارضایتی را هم پوشش داده است.

۳-٤-۳ تحلیل حساسیت به قیمت انعطاف‌پذیری رو به پایین

شکل ۱۴، میزان مصرف و تولید بهینه ۲۴ساعته را به‌ازای تغییر قیمت محصول انعطاف‌پذیری رو به پایین را نشان می‌دهد. با افزایش قیمت انعطاف‌پذیری، مصرف سیستم تهویه از حالت بهینه پایه به سمت مقدار دلخواه کاربر حرکت کرده و در قیمتی معادل ۰٫۸ انرژی، با تنظیمات کاربر یکسان شده است. این تعادل تا قیمت ۱٫۳ برابری نسبت به انرژی ادامه داشته و پس از آن مصرف به‌سمت مقدار بیشینه حرکت کرده است.

روند حساسیت تولید پنل خورشیدی کاملاً برعکس سیستم تهویه بوده و با افزایش قیمت انعطاف‌پذیری رو به پایین، از مقدار پایه کاهش پیدا کرده و در قیمت ۷۰ درصدی نسبت به قیمت انرژی، با الزام شبکه یکسان شده است. تعادل در این توان تا قیمت ۱٫۳ برابری ادامه داشته و بعد از این قیمت، تولید به سمت مقدار کمینه رفته است.



شکل ۱۶: توان PV در حضور هر دو محصول انعطاف‌پذیری.

۳-۵ HEMS با لحاظ بازار انعطاف‌پذیری در هر دو سمت

در این مورد مطالعاتی، چهار سناریو مختلف در نظر گرفته شده است. یک سناریو مربوط به بازار انرژی بوده که به نوعی قیمت هر دو محصول انعطاف‌پذیری در آن صفر فرض شده است. سه سناریو دیگر، با توجه به نتایج حاصله در دو بخش ۳-۴ و ۳-۵ (همسویی انعطاف‌پذیری رو به بالا و انرژی و خلاف سو بودن انعطاف‌پذیری رو به پایین نسبت به آن) توسعه داده شده است؛ به‌طوری‌که جمع قیمت‌های انرژی و انعطاف‌پذیری رو به بالا نسبت به محصول رو به پایین بیشتر، کمتر و مساوی باشد. اگر از عوامل پیچیده‌کننده صرفنظر شود، تحلیل ساده‌تر است. زمانی که مجموع قیمت انرژی و انعطاف‌پذیری رو به بالا بر قیمت انعطاف‌پذیری رو به پایین غالب باشد، مصرف و تولید بهینه به ترتیب در مقادیر کمینه و بیشینه خواهد بود. بالعکس، زمانی که قیمت انعطاف‌پذیری رو به پایین از مجموع دو مورد دیگر بیشتر باشد، مصرف و تولید بهینه به ترتیب مقادیر بیشینه و کمینه است. در شرایط مرزی برابری، تمامی مقادیر بار و تولید، بهینه محلی هستند. نتایج این سناریوها با لحاظ عوامل پیچیده‌کننده، در شکل‌های ۱۵ و ۱۶ به تصویر کشیده شده است.

تحلیل‌های این مورد مطالعاتی، مشابه بخش ۳-۴-۲ است. وقتی مجموع قیمت انرژی و انعطاف‌پذیری رو به بالا نسبت به انعطاف‌پذیری رو به پایین کمتر است، میزان بهینه مصرف تهویه (تولید پنل خورشیدی) بیشتر (کمتر) از مقدار دلخواه کاربر (الزام‌شده از بالادست) تعیین می‌شود. وقتی که این مجموع قیمت بر قیمت انعطاف‌پذیری رو به پایین غالب شده، HEMS مقدار بهینه مصرف تهویه (تولید پنل خورشیدی) را کمتر (بیشتر) از مقدار دلخواه کاربر (الزام‌شده از بالادست) به دست می‌آورد. در شرایط مرزی برابری قیمت‌ها هم، میزان بهینه مصرف و تولید به ترتیب منطبق با مقادیر دلخواه کاربر و الزام شده از بالادست خواهد بود.

٤- جمع‌بندی و پیشنهادها

با توجه به جایگاه کلیدی خانه‌های هوشمند در شبکه‌های پایین‌دست و روند روبه‌رشد رویکردهای محاسباتی مبتنی بر هوش مصنوعی، مقاله حاضر بر مدیریت انرژی بهینه خانگی مبتنی بر یادگیری تقویتی چندعامله متمرکز شده است. با مرور پیشینه مطالعات، دیده شد که انعطاف‌پذیری این خانه‌ها فقط از منظر فنی مورد توجه قرار گرفته و شرکت همزمان آن‌ها در بازارهای انرژی و انعطاف‌پذیری یک خلأهای تحقیقاتی است. بنابراین، مدل پیشنهادی ضمن رفع این خلاء، عوامل پیچیده‌کننده مرتبط با وسایل خانگی (نارضایتی کاربران) و تولیدات پراکنده تجدیدپذیر (الزامات بالادستی) را نیز در مسئله مدیریت نموده است.

نتایج اساسی حاصله از شبیه‌سازی‌های عددی پیاده‌سازی‌شده در محیط پایتون در قالب سه مورد کلیدی ارائه شده است. اول اینکه اگر فقط بازار انرژی لحاظ شود، تولید و بار خانه برای کسب سود اقتصادی بیشتر به ترتیب در مقادیر حداکثر و حداقل خود قرار می‌گیرند. لحاظ نارضایتی کاربر و الزام بالادستی در این شرایط، سبب نزدیک‌تر شدن پاسخ بهینه به مقادیر هدف تعیینی بار و تولید می‌شود. دوم اینکه انعطاف‌پذیری روبه‌بالا همسو با بازار انرژی بوده و آثار اقتصادی آن را تقویت می‌کند. همچنین این نوع از انعطاف‌پذیری تعادل بین وجه اقتصادی و نارضایتی بار (یا جریمه تولید) را به سمت وجه اقتصادی سنگین‌تر می‌کند. سوم اینکه انعطاف‌پذیری روبه‌پایین با بازار انرژی در تعارض است و سبب تضعیف آثار اقتصادی آن است. این نوع از انعطاف‌پذیری تعادل بین وجه اقتصادی و نارضایتی بار (یا جریمه تولید) را از بهینه اقتصادی دورتر می‌کند؛ به‌طوری‌که در شرایط مرزی قیمت یکسان انرژی و انعطاف‌پذیری روبه‌پایین، نارضایتی (یا جریمه) تنها عنصر اثرگذار است.

با وجود این نتایج، مدل پیشنهادی محدودیت‌هایی نیز دارد که می‌تواند در پژوهش‌های آتی رفع شود. به‌طور خاص، نارضایتی مشترکین مشابه بسیاری از مطالعات ادبیات تحقیق، نسبتاً ساده و وابسته به میزان انحراف از خواسته کاربران (متناسب با مجذور اختلاف زمانی و توانی به ترتیب در وسایل با قابلیت جابجایی زمانی و توان–متغیر) مدل شده است. اعمال سازوکارهای واقع‌گرایانه‌تر نارضایتی (مثل آسایش حرارتی بر اساس تغییرات دینامیک دما) در چارچوب پیشنهادی قابل پیاده‌سازی است. نگاه به خودروهای الکتریکی نیز نسبتاً ساده و مشابه بارهای توان–متغیر بوده است. مدل دوسویه (با هر دو بعد شارژ و دشارژ) برای این تجهیزات نیز می‌تواند در مدل ادغام شود. به عنوان نکته‌ای دیگر، قیمت‌های انرژی و انواع انعطاف‌پذیری به عنوان ورودی‌های معلوم و قطعی از طریق پیش‌بینی، در مسئله لحاظ شده‌اند. رویکردهای متنوع احتمالی و امکانی، برای مدیریت عدم‌قطعیت این پیش‌بینی‌ها قابل استفاده هستند.

مراجع

[1] M. Parsa Moghaddam, S. Nasiri, and M. Yousefian, "5D giga trends in future power systems," In M. Parsa Moghaddam, R. Zamani, H. Haes Alhelou, and P. Siani (eds.), *Decentralized Frameworks for Future Power Systems*, vol. 1, ch. 2, pp. 19-50, 2022.

[2] M. Jäntti, A. Jäntti, and M. Shafie-khah, "Toward customer-centric power grid: Residential EV charging simulator for smart homes," In M. Parsa Moghaddam, R. Zamani, H. Haes Alhelou, and P. Siani (eds.), *Decentralized Frameworks for Future Power Systems*, vol. 1, ch. 9, pp. 207-226, 2022.

[3] H. Delkhosh and M. Jorjani, "Green approaches in future power systems," In M. Parsa Moghaddam, R. Zamani, H. Haes Alhelou, and P. Siani (eds.), *Decentralized Frameworks for Future Power Systems*,vol. 1, ch. 5, pp. 99-127, 2022.

[4] Y. M. Rind, M. H. Raza, M. Zubair, M. Q. Mehmood, and Y. Massoud, "Smart energy meters for smart grids, an internet of things perspective," *Energies*, vol. 16, no. 4, Article ID: 1974, Feb. 2023.

[5] F. F. Alruwaili, *et al.,* "A decentralized approach to smart home security: blockchain with red-tailed hawk-enabled deep learning," *IEEE Access*, vol. 12, pp.14146-14156, 2024.

[6] A. Saif, S. K. Khadem, M. F. Conlon, and B. Norton, "Impact of distributed energy resources in smart homes and community-based electricity market," *IEEE Trans. on Industry Applications*, vol. 59, no. 1, pp. 59-69, Aug. 2022.

[7] H. Delkhosh, P. Emami, and M. P. Moghaddam, "Developments toward sustainable energy system operation," In *Hosting Capacity Aspects in Distribution Networks Towards Sustainable Energy Systems*, vol. 1, ch. 11, pp. 227-250, 2025.

[8] M. A. Khan, *et al.,* "Investigation and analysis of demand response approaches, bottlenecks, and future potential capabilities for IoT-enabled smart grid," *IET Renewable Power Generation*, vol. 18, no. 15, pp. 3509-3535, May 2024.

[9] M. Maldet, *et al.,* "Trends in local electricity market design: Regulatory barriers and the role of grid tariffs," *Journal of Cleaner Production*, vol. 358, Article ID: 131805, Feb. 2022.

[10] O. Rebenaque, C, Schmitt, K. Schumann, T. Dronne, and F. Roques, "Success of local flexibility market implementation: A review of current projects," *Utilities Policy*, vol. 80, Article ID: 101491, Feb. 2023.

[11] H. Tang, S. Wang, and H. Li, "Flexibility categorization, sources, capabilities and technologies for energy-flexible and grid-responsive buildings: State-of-the-art and future perspective," *Energy*, vol. 219, Article ID: 119598, Mar. 2021.

[12] H. Khajeh H. Firoozi, and H. Laaksonen, "Flexibility potential of a smart home to provide TSO-DSO-level services," *Electric Power Systems Research*, vol. 205, Article ID: 107767, Apr. 2022.

[13] F. Heider, A. Jahic, M. Plenz, and D. Schulz, "Extended residential power management interface for flexibility communication and uncertainty reduction for flexibility system operators," *Energies*, vol. 15, no. 4, Article ID: 1257, Feb. 2022.

[14] H. Karimi, G. B. Gharehpetian, R. Ahmadiahangar, and A. Rosin, "Optimal energy management of grid-connected multi-microgrid systems considering demand-side flexibility: A two-stage multi-objective approach," *Electric Power Systems Research*, vol. 214, pt. A, Article ID: 108902. Jan. 2023.

[15] C. Srithapon and D. Månsson, "Predictive control and coordination for energy community flexibility with electric vehicles, heat pumps and thermal energy storage," *Applied Energy*, vol. 347, Article ID: 121500, Oct. 2023.

[16] M. Nutakki and S. Mandava, "Review on optimization techniques and role of Artificial Intelligence in home energy management systems," *Engineering Applications of Artificial Intelligence*, vol. 119, Article ID: 105721, Mar. 2023.

[17] Z. Huang, F, Wang, Y, Lu, X.Chen, and Q. Wu, "Optimization model for home energy management system of rural dwellings," *Energy*, vol. 283, Article ID: 129039, Nov. 2023.

[18] I. L. R, Gomes, M. G. Ruano, and A. E, Ruano, "MILP-based model predictive control for home energy management systems: A real case study in Algarve, Portugal" *Energy and Buildings*, vol. 281, Article ID: 112774, Feb. 2023.

[19] S, Bahramara, "Robust optimization of the flexibility-constrained energy management problem for a smart home with rooftop photovoltaic and an energy storage," *Journal of Energy Storage*, vol. 36, Article ID: 102358, Apr. 2021.

[20] S, Dorahaki, M. Mollahassani-Pour, and M. Rashidinejad, "Optimizing energy payment, user satisfaction, and self-sufficiency in flexibility-constrained smart home energy management: a multi-objective optimization approach," *e-Prime-Advances in Electrical Engineering, Electronics and Energy*, vol. 6, Article ID: 100385, Dec. 2023.

[21] D, Vamvakas, P, Michailidis, C. Korkas, and E, Kosmatopoulos, "Review and evaluation of reinforcement learning frameworks on smart grid applications," *Energies*, vol. 16, no. 14, Article ID: 5326, Jul. 2023.

[22] A, Dolatabadi, H. Abdeltawab, and Y. A. R. I, Mohamed, "A novel model-free deep reinforcement learning framework for energy management of a PV integrated energy hub," *IEEE Trans. on Power Systems*, vol. 38, no. 5, pp.4840-4852, Sept. 2022.

[23] Y, Zhang, *et al.,* "Two-step diffusion policy deep reinforcement learning method for low-carbon multi-energy microgrid energy management," *IEEE Transactions on Smart Grid*, vol. 15, no. 5, pp.4576-4588, Sept. 2024.

[24] M, Hashemnezhad, H. Delkhosh, and M. P, Moghaddam, " Aggregator pricing strategy for community energy management based on multi-agent reinforcement learning considering customer loss or gain," *Sustainable Energy, Grids and Networks*, vol. 41, Article ID: 101607, Mar. 2025.

[25] W, Pinthurat, T. Surinkaew, and B, Hredzak, "An overview of reinforcement learning-based approaches for smart home energy management systems with energy storages," *Renewable and Sustainable Energy Reviews*, vol. 202, Article ID: 114648, Sept. 2024.

[26] X. Xu, *et al.,* "A multi-agent reinforcement learning-based data-driven method for home energy management," *IEEE Trans. on Smart Grid*, vol. 11, no. 4, pp.3201-3211, Jul. 2020.

[27] K, Ren, *et al.,* "A data-driven DRL-based home energy management system optimization framework considering uncertain household parameters," *Applied Energy*, vol. 355, Article ID: 122258, Mar. 2024.

[28] L. Langer, and T, Volling, "A reinforcement learning approach to home energy management for modulating heat pumps and photovoltaic systems," *Applied energy*, vol. 327, Article ID: 120020, Dec. 2022.

[29] A, Shahabi, H. Delkhosh, and M. P, Moghaddam, "Home energy management system based on multi-agent deep reinforcement learning handling the user's thermal preferences," in *Proc. 2025 10th Int. Conf. on Technology and Energy Management,* 5 pp., Tariz, Iran, Apr. 2025,

[30] R, Lu, *et al.,* "Reward shaping-based actor–critic deep reinforcement learning for residential energy management," *IEEE Trans. on Industrial Informatics*, vol. 19, no. 3, pp.2662-2673, Nov. 2022.

[31] G, Wei, *et al.,* "Deep reinforcement learning for real-time energy management in smart home," *IEEE Systems Journal*, vol. 17, no. 2, pp.2489-2499, Jun. 2023.

[32] X, Zhou, S, Xue, H. Du, and Z. Ma, , "Optimization of building demand flexibility using reinforcement learning and rule-based expert systems," *Applied Energy*, vol. 350, Article ID: 121792, Nov. 2023.

[33] Y. Zheng, *et al.,* "Interior-point policy optimization based multi-agent deep reinforcement learning method for secure home energy management under various uncertainties," *Applied Energy*, vol. 376, Article ID: 124155, Dec. 2024.

[34] A. A, Amer, K. Shaban, and A. M, Massoud, "DRL-HEMS: Deep reinforcement learning agent for demand response in home energy management systems considering customers and operators perspectives. *IEEE Trans. on Smart Grid*, vol. 14, no. 1, pp.239-250, Jan. 2022.

[35] S. S. Shuvo, and Y, Yilmaz, "Home energy recommendation system (HERSs): A deep reinforcement learning method based on residents' feedback and activity," *IEEE Trans. on Smart Grid*, vol. 13, no. 4, pp. 2812-2821, Jul. 2022.

پیمان مادح خاکسار تحصیلات خود را در مقاطع کارشناسی و کارشناسی ارشد مهندسی برق قدرت به‌ترتیب در سال‌های ۱۳۹۱ و ۱۴۰۳ در دانشگاه آزاد اسلامی و دانشگاه تربیت مدرس به پایان رسانده است و از سال ۱۳۹۳ به عنوان مدیر پروژه‌های EPC در شرکت صنام فعالیت دارد. ایشان همچنین به عنوان مهندس ناظر در سازمان نظام مهندسی ساختمان استان البرز مشغول به فعالیت می‌باشد. زمینه‌های تحقیقاتی مورد علاقه ایشان عبارتند از: بهینه‌سازی مصرف انرژی برق، انرژی‌های تجدیدپذیر و هوش مصنوعی.

سید حامد دلخوش تحصیلات خود را در مقاطع کارشناسی، کارشناسی ارشد و دکتری مهندسی برق گرایش سیستم‌های قدرت به ترتیب در سال‌های ۱۳۹۲، ۱۳۹۴ و ۱۳۹۸ و به ترتیب در دانشگاه های صنعتی امیرکبیر، صنعتی شریف و تربیت مدرس به پایان رسانده است. وی از سال ۱۳۹۹ عضو هیأت علمی دانشکده مهندسی برق و کامپیوتر دانشگاه تربیت مدرس می‌باشد. همچنین ایشان با مرکز ملی مطالعات و برنامه‌ریزی شبکه‌های قدرت دانشگاه تربیت مدرس به عنوان مدیر پروژه‌های تحقیقاتی صنعتی همکاری دارد. زمینه های تحقیقاتی مورد علاقه ایشان عبارتند از: امنیت سایبری- فیزیکی شبکه‌ی قدرت، بهره‌برداری و کنترل شبکه‌های قدرت و انرژی‌های تجدیدپذیر.