

روش یادگیری تقویتی در بهینه سازی استوکاستیک سیستم مخزن

امیرمحمد مرادی^{1*}، علیرضا برهانی داریان²

1- دانشجوی دکتری، دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده عمران (a.m.moradi@mail.kntu.ac.ir)
2- دانشیار منابع آب، دانشگاه صنعتی خواجه نصیرالدین طوسی، دانشکده عمران (borhani@kntu.ac.ir)

چکیده

روش برنامه ریزی پویای استوکاستیک صریح (Stochastic Dynamic Programming) SDP برای حل مسایل پیچیده منابع آب با مشکلاتی نظیر بلای ابعادی و بلای مدلسازی مواجه می شود؛ این مساله موجب گردیده محققان همواره در پی آرایه و توسعه روشهای مختلف برای حل مسایل بهینه سازی استوکاستیک سیستمهای منابع آب باشند. در این تحقیق روش Q-learning که یکی از تکنیک های الگوریتم یادگیری تقویتی است برای بهره برداری بهینه از یک سیستم مخزن پیشنهاد شده و سپس نتایج آن با نتایج حاصل از روش SDP مقایسه گردیده است. مقایسه سیاست های بهینه مستخرج از این دو روش نشان می دهد که Q-learning در برخورد با پدیده های استوکاستیک جریان رودخانه و مسئله عدم قطعیت دخیل در آن عملکرد بهتری داشته است.

35 mm

واژه های کلیدی: بهینه سازی استوکاستیک مخزن، یادگیری تقویتی، برنامه ریزی پویای استوکاستیک، سیاست بهره برداری.

1- مقدمه

امروزه افزایش جمعیت و به دنبال آن افزایش نیازها و توسعه اقتصادی، باعث پیچیده تر و رقابتی تر شدن سیستمهای منابع آب شده است به گونه ای که لی و لبدی [1] بروز پیچیدگی های محاسباتی در این سیستمها را مورد تاکید قرار می دهند. بطور کلی، توسعه سیستم های منابع آب منجر به ایجاد یک فضای حالت - تصمیم با ابعاد بسیار بالا و اهداف چندگانه و غالباً متضاد تحت یک محیط استوکاستیک پیچیده و بزرگ مقیاس می شود که حل چنین سیستمهایی با روشهای سنتی ریاضیاتی بسیار مشکل بوده و یا امکان پذیر نمی باشد. در مواردی برای رفع این مشکل اقدام به ساده سازیهای بسیار می شود که این خود می تواند منجر به جوابهای غیربهینه توأم با هزینه بالا شود. در این خصوص استفاده از روشهای تطبیقی و هوش مصنوعی گامی موثر در حل مسائل چنین سیستمهای پیچیده ای قلمداد می شود.

روشهای بهینه سازی استوکاستیک طبق طبقه بندی لبدی [2] به دو گروه بهینه سازی ضمنی¹ (ISO) و بهینه سازی صریح² (ESO) تقسیم می شوند. تکنیکهای ISO شامل استفاده از سری زمانی تاریخی جریان روخانه یا سری زمانی مصنوعی می شود. در روشهای ESO برخلاف ISO، به جای اینکه از یک سری زمانی مشخص استفاده شود روابط احتمالاتی پارامترهای عدم قطعی به طور مستقیم وارد مدل می شود. عیب روشهای ESO در این است که برای سیستمهای چند مخزنه، پیچیدگی محاسباتی آن افزایش یافته و با مشکل ابعادی و یا مشکل مدلسازی مواجه می شوند. از جمله این روشها می توان برنامه ریزی پویای استوکاستیک صریح SDP را نام برد.

¹ Implicit Stochastic Optimization (ISO)

² Explicit Stochastic Optimization (ESO)