



توسعه مدل ریاضی الگوریتم یادگیری تقویتی برای خودکارسازی آب‌بند‌های کشویی در کانال‌ها

کاظم شاهوردی¹ - محمدجواد منعم^{2*}

تاریخ دریافت: 1392/05/28

تاریخ پذیرش: 1394/03/31

چکیده

امروزه با توجه به کمبود آب و ضعف مدیریت در بخش آب کشاورزی و به منظور استفاده بهینه از آب، بهبود عملکرد شبکه‌های آبیاری ضروری است. در سال‌های اخیر، بکارگیری سامانه‌های کنترل خودکار پیشرفته به منظور بهبود عملکرد شبکه‌های آبیاری مورد توجه محققان قرار گرفته است. برای استفاده از این سامانه‌ها، تهیه مدل ریاضی سیستم‌های کنترل خودکار و سازه‌های مربوط بصورت تلفیقی با مدل‌های هیدرودینامیکی ضروری است. هدف اصلی این تحقیق، توسعه مدل ریاضی سیستم کنترل خودکار بالادست دریچه کشویی با استفاده از الگوریتم یادگیری تقویتی بصورت یک زیربرنامه در مدل هیدرودینامیکی ICSS است. دو سناریوی افزایشی و کاهش برای جریان ورودی با دبی اولیه 25 لیتر در ثانیه در نظر گرفته شد و شبیه‌سازی انجام گرفت. شاخص‌های حداکثر خطای مطلق، میانگین متوسط خطاها و زمان پاسخگویی سیستم برای ارزیابی عملکرد سیستم کنترل توسعه داده شد. مقادیر این شاخص‌ها برای سناریوی کاهش جریان برابر صفر بدست آمد. نتایج حاصل نشان می‌دهد مدل توسعه یافته عملکرد مناسبی داشته و سطح آب را بخوبی کنترل می‌کند. از این رو توسعه آن برای کانال‌های آبیاری پیشنهاد می‌شود.

واژه‌های کلیدی: تلفیق مدل، بهبود عملکرد، یادگیری

مقدمه

آبیاری مورد استفاده قرار گرفته‌اند عبارتند از: (2) PI، (12) Bival، (3) CARRD، (3) Fuzzy، (16) MPC، (15) Adaptive control، (13). همانطور که مشاهده می‌شود کنترل بهتر پدیده‌های هیدرولیکی همچنان توسط محققین مختلف مورد بررسی قرار گرفته و تکنیک‌های کنترل بهتر در حال توسعه می‌باشد. در همین راستا الگوریتم یادگیری تقویتی در این تحقیق مورد توجه قرار گرفته است.

الگوریتم‌های هوش مصنوعی یکی از الگوریتم‌های کنترل پیشرفته می‌باشد که می‌توان از آنها به منظور کنترل خودکار سازه‌ها در کانال‌های آبیاری استفاده کرد. هوش مصنوعی، دانش ساختن ماشین‌ها یا برنامه‌های هوشمند است. یکی از شاخه‌های مهم و پرکاربرد هوش مصنوعی، یادگیری تقویتی³ (RL) می‌باشد. نحوه عملکرد این الگوریتم و در نتیجه افزایش دقت در طول شبیه‌سازی با استفاده از این روش بررسی گردید (5) و سپس روندنما این الگوریتم به منظور استفاده محققین ارائه شد (4). الگوریتم RL برای بهره‌برداری سازه دریچه کشویی در یک کانال روباز با دو بازه مورد استفاده قرار گرفت (9). هدف از این تحقیق توسعه مدل ریاضی سیستم کنترل خودکار بالادست دریچه کشویی با استفاده از الگوریتم

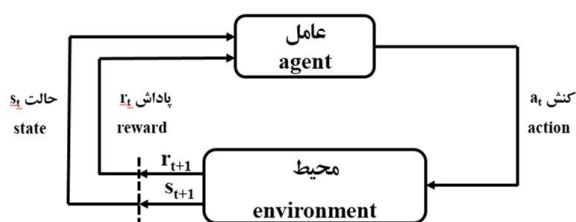
به دلیل پایین بودن عملکرد شبکه‌های آبیاری، ایده بهبود روش‌های تحویل و توزیع آب و افزایش انعطاف‌پذیری موضوع مطرح در بین محققین است. یکی از بهترین راه‌های ممکن برای بهبود عملکرد و افزایش انعطاف‌پذیری تحویل آب در کانال‌های روباز، اتوماسیون یا خودکارسازی سیستم‌های تحویل آب در کانال‌ها می‌باشد. خودکارسازی کانال‌های آبیاری اولین بار با سامانه‌های الکترومکانیکی Littleman و Colvin (6) و سپس با دریچه‌های خودکار هیدرولیکی آمیل، اویس و آویو توسعه یافت. در سال‌های اخیر با توسعه مدل‌های ریاضی، ترکیب مدل‌های شبیه‌ساز جریان کانال‌ها با الگوریتم‌های کنترل پیشرفته، مورد توجه محققین قرار گرفته است. برخی از محققین الگوریتم‌های موجود را مورد بررسی قرار داده و مشخصات، مزایا و معایب هر یک از این الگوریتم‌ها را مورد بحث قرار داده‌اند (7 و 10). نمونه‌هایی از این الگوریتم‌ها که در پروژه‌های

1 و 2- دانش آموخته دکتری و دانشیار گروه مهندسی سازه‌های آبی، دانشگاه تربیت مدرس

* - نویسنده مسئول: (Email: javadmonem@gmail.com)

شده است. این مراحل به ترتیب شامل مشاهده حالت S_t ، تصمیم‌گیری عامل برای انجام کنش a_t ، انجام کنش a_t ، شبیه‌سازی سیستم و مشاهده حالت جدید (S_{t+1})، اختصاص پاداش R_{t+1} برای کنش a_t و یادگیری از تجربیات می‌باشد که تا رسیدن به شرط پایداری این روند تکرار می‌شود (14).

در شبکه‌های آبیاری، در سازه‌های آب‌بند ابتدا با توجه به جهت کنترل (بالادست یا پایین‌دست) عمق جریان مشاهده می‌شود (S_t) و با توجه به مقدار این عمق و میزان انحراف آن از عمق هدف، عامل میزان تغییر در بازشدگی دریاچه را محاسبه و آنرا اعمال می‌کند (a_t). شبیه‌سازی انجام می‌شود و عمق جدید مشاهده می‌شود (S_{t+1}) و سپس به بازشدگی a_t پاداش داده می‌شود و این فرآیند تا زمانی که عمق جریان در عمق هدف تثبیت شود ادامه می‌یابد.



شکل 1- فرایند یادگیری تقویتی (14)

Figure 1- Reinforcement Learning process [14]

مدل هیدرودینامیکی ICSS

مدل ICSS توانایی شبیه‌سازی تمامی مشخصات مهم هیدرولیکی و مورد نیاز در کانالهای آبیاری را دارد و توسط مانز در سال 1990 تهیه گردید. با استفاده از این مدل می‌توان به بهره‌برداری از سازه‌ها در شبکه‌های آبیاری اقدام کرد. صحت عملکرد این مدل با استفاده از داده‌های دقیق آزمایشگاهی مورد بررسی و تایید قرار گرفته است و به تایید انجمن مهندسين عمران کانادا رسیده است (11). در این مدل امکان شبیه‌سازی طیف قابل توجهی از سازه‌های آبی وجود دارد که هر کدام به صورت یک شرط مرزی در نظر گرفته می‌شوند. دو زیربرنامه مهم از مدل ICSS که در این تحقیق از اهمیت برخوردارند زیربرنامه‌های WRRST (WRite ReStarT) و RDRST (ReAd) می‌باشد. زیربرنامه WRRST تمامی اطلاعات شبیه‌سازی شده در زمان t را در یک فایل متنی ذخیره و زیربرنامه RDRST این اطلاعات را برای ادامه شبیه‌سازی فراخوانی می‌کند.

مدل ریاضی سیستم کنترل RL

در این تحقیق مدل ریاضی الگوریتم RL، توسعه داده شد و با مدل هیدرودینامیکی ICSS تلفیق گردید. روندنمای این الگوریتم در شکل 2 نشان داده شده است. کادر route که در این شکل دیده

یادگیری و بصورت تلفیق با مدل هیدرودینامیکی ICSS می‌باشد.

معرفی مبانی الگوریتم RL

در یادگیری ماشین، هدف برنامه‌ریزی رایانه‌ها بنحوی است که بتواند از تجربیات گذشته برای حل یک مسئله داده شده استفاده کند. یادگیری تقویتی در یک بیان کلی یعنی با توجه به شناخت محیط و بر اساس نتایج تعاملاتی که با محیط وجود داشته و سودها و زیانهایی که در نتیجه انجام کنش‌های مختلف بدست آمده است، راهبردی اتخاذ شود که با عمل به آن در بلند مدت، مطلوبیت بیشینه شود (14).

بطور کلی اجزاء اصلی مسائل یادگیری تقویتی شامل پنج عنصر عامل¹، محیط²، سیاست³، تابع پاداش⁴ و مدل هستند. عامل همان یادگیرنده است که از طریق مجموعه‌ای از حسگرها و محرک‌هایش می‌تواند با محیط در تعامل باشد. محیط فضایی است که عامل کنشهای⁵ مربوطه را در آن انجام داده و روی آن تاثیر می‌گذارد. سیاست، هسته اصلی عامل یادگیرنده تقویتی است که مشخص می‌کند که عامل چه کنشی را در هر حالت باید انجام دهد. با توجه به میزان مطلوبیت کنش انجام شده، پاداشی به آن داده می‌شود که مقدار آن از طریق تابع پاداش مشخص می‌شود. مدل رفتار محیط را شبیه‌سازی کرده و حالت بعدی سیستم را پیش‌بینی می‌کند (1).

در این تحقیق در شبکه‌های آبیاری منظور از عامل، سازه‌های آب‌بند می‌باشد که میزان بازشدگی آنها برای شرایط مختلف یاد داده می‌شود. منظور از محیط، کانال یا شبکه‌ای از کانالها است که شرایط هیدرولیکی جریان درون آن (که معادل حالت⁶ در یادگیری تقویتی است)، به عنوان ورودی به الگوریتم RL داده می‌شود و یادگیری بر اساس تغییرات آن انجام می‌گیرد. در این تحقیق، عمق آب به عنوان حالت می‌باشد. منظور از مدل، مدل هیدرودینامیکی می‌باشد که در این تحقیق ICSS مورد استفاده قرار گرفته است. عمق، میزان بازشدگی و پاداش مرتبط با آن به عنوان یک عضو تلقی می‌شود که یک کلاسیفایر⁷ نامیده می‌شود. مجموعه‌ای از کلاسیفایرها پس از یادگیری، سیاست الگوریتم یادگیری تقویتی را تشکیل می‌دهد بطوریکه برای هر عمق مشخص، میزان بازشدگی دریاچه مرتبط با آن که موجب تثبیت عمق آب در مقدار هدف می‌شود در آن ذخیره می‌گردد.

مراحل انجام عملیات در یادگیری تقویتی در شکل 1 نشان داده

- 1- Agent
- 2- Environment
- 3- Policy
- 4- Reward function
- 5- Actions
- 6- State
- 7- Classifier

مقایسه می‌شود. اگر CR کمتر از PR باشد، عضو تولید شده مستقیماً به جمعیت اضافه می‌شود، در غیر اینصورت R برابر با CR قرار داده شده و PR نیز برابر با CR قرار داده می‌شود و سپس عضو تولید شده به جمعیت اولیه اضافه می‌شود. مقدار R با Rmax مقایسه می‌شود. اگر R کمتر از Rmax باشد این فرایند ادامه می‌یابد در غیر اینصورت WRRST فراخوانی شده و اطلاعات زمان مذکور ذخیره می‌شود سپس زیربرنامه report فراخوانی شده و خروجی برنامه گزارش می‌شود و برنامه به گام زمانی بعدی می‌رود. این فرایند تا اتمام زمان شبیه‌سازی برای هر گام زمانی در داخل مدل ICSS تکرار می‌شود.

فلوچارت زیربرنامه action در شکل 3 نشان داده شده است. در این فلوچارت Rclass پاداش کلاسیفایر انتخاب شده می‌باشد. GO بازشدگی دریچه می‌باشد که در شکل به اختصار نوشته شده است. وقتی که این زیربرنامه فراخوانی می‌شود، سعی می‌شود که بازشدگی دریچه مناسب که عمق آب را در عمق هدف تثبیت می‌کند، پیدا شود. برای این منظور، ابتدا در جمعیت موجود جستجو انجام می‌شود و MS⁶ تشکیل می‌شود. MS شامل تمام کلاسیفایرهایی است که عمق در آنها برابر با عمق موجود می‌باشد. اگر پاداش یکی از کلاسیفایرهای MS برابر با Rmax باشد در اینصورت میزان بازشدگی آن اعمال می‌شود در غیر اینصورت یک کلاسیفایر بصورت تصادفی تولید می‌شود. عمق این کلاسیفایر برابر با عمق موجود و پاداش آن نیز پس از شبیه‌سازی اختصاص می‌یابد و نهایتاً کلاسیفایر تکمیل شده و به جمعیت اضافه می‌شود.

در زیربرنامه reward مقدار پاداش با استفاده از معادلات 1 و 2 محاسبه می‌شود.

$$A = \frac{|Y_t - Y|}{Y_t} \quad (1)$$

که در آن $Y =$ مقدار عمق مشاهده شده و $Y_t =$ مقدار عمق هدف می‌باشد.

$$R = (1 - A)R_0 \quad (2)$$

که در آن $R_0 =$ حداکثر پاداش ممکن می‌باشد، بدین معنی که هر وقت عمق آب برابر با عمق هدف باشد $R_0 = R$. فرایندهای تشریح شده در بالا برای یک بازه از کانال قابل اجراست.

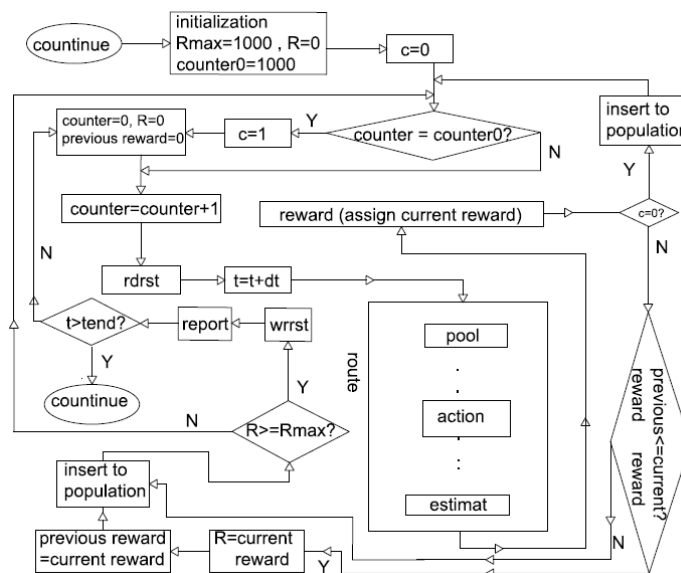
برای یک کانال با چندین بازه، فرایند الگوریتم شکل 2 ابتدا برای بازه اول انجام می‌شود و سپس برای بازه 2 و سپس برای بازه‌های بعدی به ترتیب تکرار می‌شود.

برنامه route¹ می‌باشد که جریان غیرماندگار را شبیه‌سازی می‌کند و خود شامل چندین زیربرنامه² می‌باشد. در داخل برنامه route زیربرنامه‌هایی مانند pool، estimat و action قرار دارند که زیربرنامه action یکی از زیربرنامه‌هایی است که در این تحقیق توسعه داده شده و سایر زیربرنامه‌ها توسط مانز توسعه داده شده است و جز بحث این تحقیق نیست. این زیربرنامه‌ها برای اجرای شبیه‌سازی مورد نیاز می‌باشند.

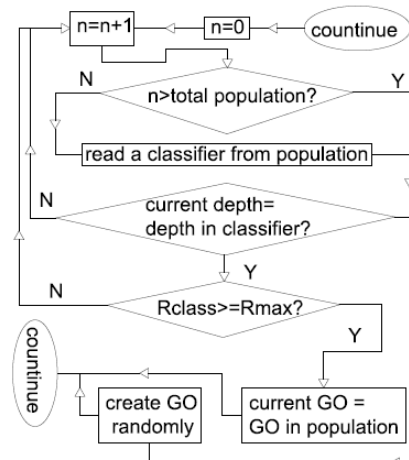
وقتی که نرم‌افزار ICSS اجرا می‌شود اطلاعات فیزیکی و هیدرولیکی توسط زیربرنامه WRRST ذخیره می‌شود. در مرحله بعد، بهره‌برداری انجام می‌شود و جریان ورودی تغییر می‌کند. تغییر جریان ورودی باعث تغییر عمق آب می‌شود. در ادامه برنامه (continue) در شکل 2، برنامه RL فراخوانی شده و تمامی اطلاعات خوانده شده و بهره‌برداری شده به RL منتقل می‌شود. در RL، ابتدا مقادیر اولیه مانند Rmax (حداکثر پاداش) و حداکثر شمارش³ به الگوریتم RL معرفی می‌شود. سپس به اندازه حداکثر شمارش (1000)، کلاسیفایرهای جمعیت اولیه بصورت تصادفی توسط الگوریتم تولید می‌شود. برای این منظور، تا زمانی که تعداد شمارشها کمتر از شمارش حداکثر باشد شمارشها افزایش یافته و RDRST فراخوانی می‌شود. زمان به اندازه dt افزایش می‌یابد و یک گام زمانی مورد توجه قرار می‌گیرد ($t=t+dt$). برنامه route فراخوانی می‌شود و زیربرنامه pool و سایر زیربرنامه‌ها تا رسیدن به زیربرنامه action اجرا می‌شوند. در ادامه بازشدگی دریچه توسط زیربرنامه action محاسبه می‌شود سپس زیربرنامه estimat و سایر زیربرنامه‌های route اجرا می‌شود که در نتیجه آن عمق جدید و سایر پارامترهای هیدرولیکی مشخص می‌گردد. در ادامه زیربرنامه reward فراخوانی شده و با توجه به عمق جدید، به بازشدگی اعمال شده پاداش داده می‌شود. شرط c کنترل می‌شود اگر $c=0$ کلاسیفایر تولید شده به جمعیت موجود اضافه می‌شود. این فرایند تا زمانی که حداکثر شمارش حاصل شود تکرار می‌شود. در اینصورت تولید جمعیت اولیه به پایان رسیده و $c=1$. با توجه به اینکه زیربرنامه WRRST در این پروسه فراخوانی نمی‌شود زمان تغییری نمی‌کند. برای انجام شبیه‌سازی مراحل زیر انجام می‌شود.

در گام اول از فرایند یادگیری در هر سیکل از شبیه‌سازی مقدار پاداش R، شمارنده و پاداش قبلی⁴ (PR) برابر با صفر قرار داده می‌شود. زیربرنامه‌های RDRST، route و reward به ترتیب برای یک گام زمانی فراخوانی شده و اجرا می‌شوند که در نتیجه آن یک عضو تولید می‌شود. سپس پاداش کنونی⁵ (CR) محاسبه شده و با PR

- 1- Program
- 2- subprogram
- 3- Counter0
- 4- Previous Reward
- 5- Current Reward



شکل 2- روندنمای الگوریتم توسعه یافته RL
Figure 2- The flowchart of developed RL



شکل 3- روندنمای محاسبه بازشدگی دریچه در الگوریتم RL
Figure 3- The flowchart of gate opening calculation in RL algorithm

شاخص‌های ارزیابی

برای ارزیابی مدل توسعه یافته از شاخص‌های حداکثر خطای مطلق (MAE¹)، میانگین مطلق خطاها (IAE²) و زمان عکس‌العمل سیستم (SRT³) (14) استفاده شد. شاخص‌های حداکثر خطای مطلق و میانگین مطلق خطاها به صورت روابط 3 و 4 معرفی می‌شوند:

$$MAE = \frac{\max(|Y_t - Y|)}{Y_t} \quad (3)$$

$$IAE = \frac{\frac{\Delta t}{T} \sum_0^T |Y_t - Y|}{Y_t} \quad (4)$$

که در آن، T= طول دوره شبیه‌سازی می‌باشد و بقیه پارامترها قبلا

- 1- Maximum Absolute Error
- 2- Integral of Absolute magnitude of Error
- 3- System Response Time

معرفی شده‌اند. شاخص حداکثر خطای مطلق نشان‌دهنده حداکثر انحراف عمق از عمق هدف است. شاخص میانگین مطلق خطا، نشان‌دهنده متوسط انحراف عمق آب از عمق هدف در طول دوره شبیه‌سازی است. فاصله زمانی بین شروع انحراف عمق از محدوده مجاز تا تثبیت دوباره عمق در درون این محدوده را زمان عکس‌العمل می‌نامند و آن را با SRT نشان می‌دهند. هر چه مقدار این شاخص کمتر باشد، عمق آب سریعتر در داخل محدوده مجاز تثبیت می‌شود. دامنه تغییرات شاخص‌های حداکثر خطای مطلق و میانگین مطلق خطا بین صفر و 100 درصد است. در صورتی که تغییرات عمق در درون محدوده مجاز صورت گیرد، زمان عکس‌العمل برابر صفر خواهد بود.

سناریوهای شبیه‌سازی

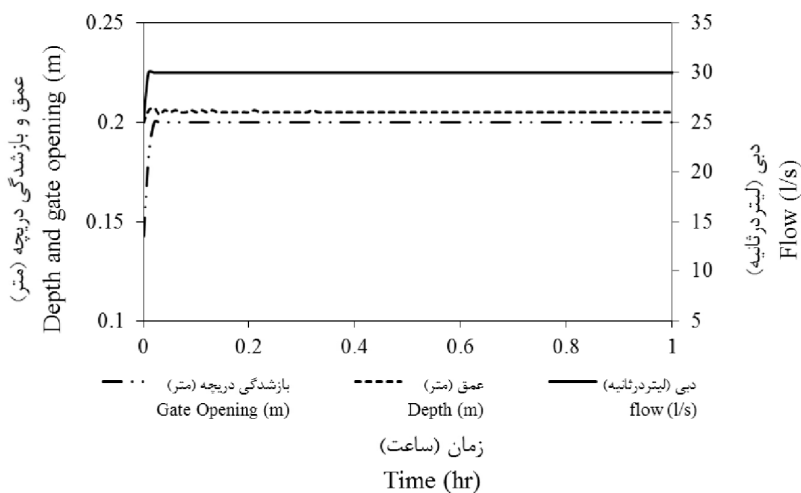
برای آزمون و ارزیابی سیستم کنترل RL توسعه یافته، فلومی به طول 10 متر و با ارتفاع و عرض به ترتیب 0/45 و 0/3 متر در نظر گرفته شد. این فلوم دارای دو بازه بوده که بوسیله یک دریچه خودکار با منطق RL از یکدیگر جدا شده‌اند. این دریچه در فاصله 4 متری از ابتدای فلوم قرار گرفته است. جریان ورودی به فلوم با استفاده از یک شیر فلکه که در ابتدای فلوم قرار گرفته است، کنترل می‌شود. عمق آب در بالادست دریچه خودکار کنترل می‌شود. به منظور بررسی عملکرد سیستم کنترل خودکار، ابتدا مشخصات فیزیکی و هیدرولیکی فلوم به مدل ICSS معرفی شد و سپس مقادیر دبی اولیه، بازشدگی اولیه دریچه و عمق هدف در مقادیر 25 l/s، 14.3 cm و 20 cm تنظیم شدند. دو سناریوی افزایش و کاهش جریان مورد توجه قرار گرفت. در سناریوی افزایش جریان، دبی اولیه 25 l/s به 30 l/s افزایش می‌یافت و در سناریوی کاهش جریان دبی اولیه 25 l/s به 20 l/s کاهش می‌یافت.

نتایج و بحث

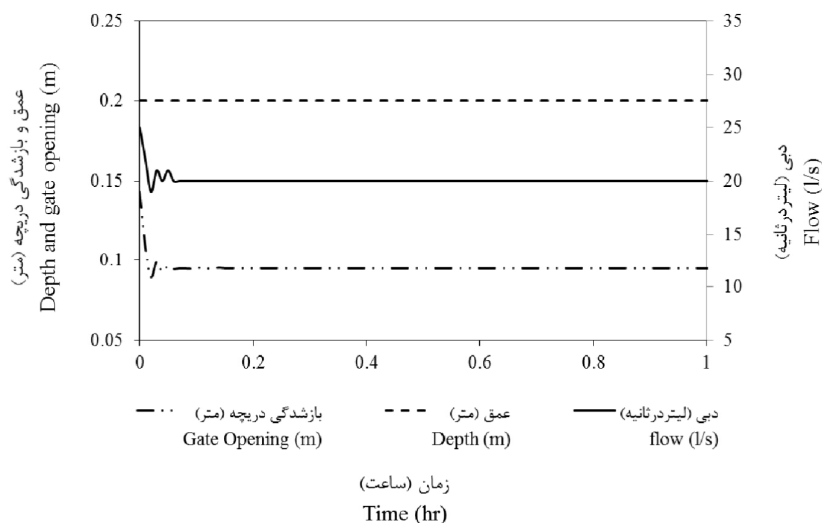
با توجه به سناریوهای اشاره شده، شبیه‌سازی انجام شد و نتایج آنالیز گردید. در پاسخ به تغییرات جریان، دریچه خودکار انحرافات سطح آب را کنترل کرده و آنرا در داخل محدوده مجاز تثبیت می‌کند. شکل شماتیک تغییرات دبی و سطح آب در بالادست دریچه و مقدار بازشدگی دریچه برای سناریوها در شکل‌های 4 و 5 نشان داده شده است. همچنین مقادیر شاخص‌های ارزیابی در جدول 1 ارائه شده است. برای سناریوی افزایش جریان، جریان بالادست از 25 l/s به 30 l/s افزایش یافت. پس از انجام یادگیری در پاسخ به این افزایش، سطح آب در بالادست دریچه به 20.7 cm افزایش یافت و سپس به مقدار

نتیجه‌گیری کلی

در این تحقیق، مدل ریاضی کنترل خودکار سازه دریچه کشویی با الگوریتم RL و بصورت تلفیقی با مدل هیدرودینامیکی ICSS تهیه شد و مورد آزمون و ارزیابی قرار گرفت. دو سناریوی افزایش و کاهش دبی ورودی با دبی اولیه 25 لیتر در ثانیه در نظر گرفته شد. تغییرات عمق آب در نقطه کنترل (بالادست دریچه خودکار) در پاسخ به تغییرات جریان بالادست مشاهده شد که نشان می‌دهد در مدت زمان کوتاهی عمق آب را در عمق هدف تثبیت کرده است.



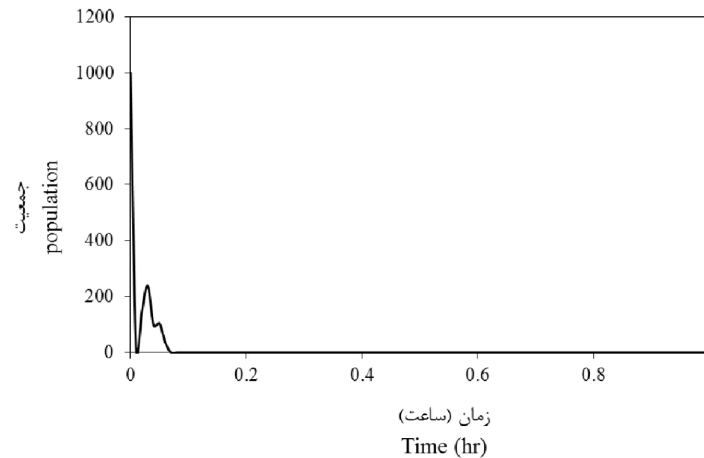
شکل 4- تغییرات جریان، عمق و بازشدگی دریچه در سناریوی افزایش جریان
 Figure 4- Flow, water depth and gate opening variations in flow increase scenario



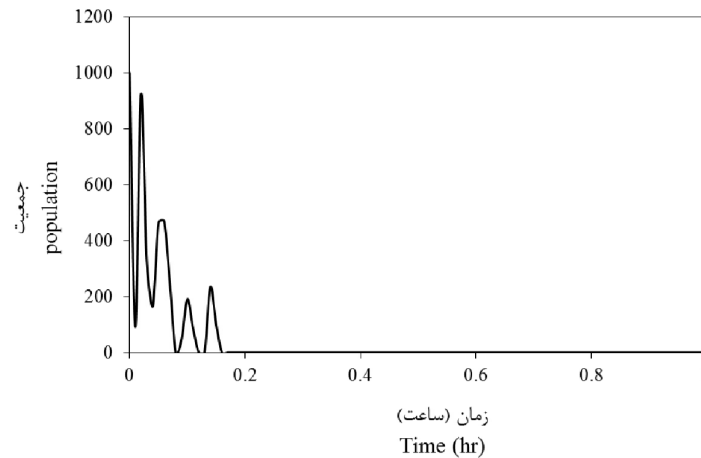
شکل 5- تغییرات جریان، عمق و بازشدگی دریچه در سناریوی کاهش جریان
 Figure 5- Flow, water depth and gate opening variations in flow decrease scenario

جدول 1- مقادیر شاخص های ارزیابی برای سناریوهای افزایش و کاهش جریان
 Table 1- Performance indicators for flow increase and decrease scenarios

	افزایش جریان (Flow increase)	کاهش جریان (Flow decrease)
Yt(cm)	20.00	20.00
Max Y(cm)	20.07	20.00
Min Y(cm)	20.00	20.00
MAE	3.50	0.00
IAE	2.57	0.00
SRT10(min)	0.00	0.00



شکل 6 - نمودار تعداد جمعیت جدید تولید شده در سناریوی افزایش جریان
Figure 6- Generation of population diagram in flow increase scenario



شکل 7 - نمودار تعداد جمعیت جدید تولید شده در سناریوی کاهش جریان
Figure 7- Generation of population diagram in flow decrease scenario

بیشتر برای کنترل سازه‌ها در کانالهای آبیاری مورد استفاده قرار گیرد.

نتایج نشان می‌دهد سیستم کنترل RL از نظر دقت و زمان پاسخگویی دارای عملکرد مطلوبی می‌باشد و می‌تواند برای مطالعات

منابع

- 1- Alpaydin E. 2004. Introduction to Machine Learning. The MIT Press.
- 2- Bennett S. 1993. A History of Control Engineering. Peter Peregrinus Ltd.
- 3- Burt C.M. 1983. Regulation of Sloping Canals by Automatic Downstream Control. Dept. of Agricultural and Irrigation Engineering, Utah State University.
- 4- Butz M., and Wilson S. 2001a. An algorithmic description of XCS. Advances in Learning Classifier Systems, 267-274.
- 5- Butz M.V., Kovacs T., Lanzi P.L., and Wilson S.W. 2001b. How XCS evolves accurate classifiers. Proceedings of the Third Genetic and Evolutionary Computation Conference (GECCO).
- 6- Buyalski C.P., Ehler D., Falvey, H.T., Rogers D.C., and Serfozo E.A. 1991. Canal Systems Automation Manual: Volume 1. US Department of Interior.

- 7- Clemmens A. and Replogle J. 1989. Control of irrigation canal networks. *Journal of irrigation and drainage engineering*, 115(1): 96-110.
- 8-Clemmens A.J., Kacerek T. F., Grawitz B., and Schuurmans W. 1998. Test cases for canal control algorithms. *Journal of irrigation and drainage engineering*, 124(1): 23-30.
- 9- Hernández J. and Merkley G. 2011. Canal structure automation rules using an accuracy-based learning classifier system, a genetic algorithm, and a hydraulic simulation model. I: result. *Journal of irrigation and drainage engineering*, 137(1), 12–16.
- 10- Malaterre P.O., Rogers D. C., and Schuurmans J. 1998. Classification of canal control algorithms. *Journal of irrigation and drainage engineering*, 124(1): 3-10.
- 11- Monem M.J., and Manz D. H. 1994. Application of simulation techniques for improving the performance of irrigation conveyance systems. *Journal of Water Resources Engineering*, 2: 1-22.
- 12-Rogers D.C., and Goussard J. 1998. Canal control algorithms currently in use. *Journal of irrigation and drainage engineering*, 124(1): 11-15.
- 13- Sawadogo S., Faye R., Benhammou A., and Akouz K. 2000. Decentralized adaptive predictive control of multireach irrigation canal. *Institute of Electrical and Electronics Engineers*: 5: 3438-3442.
- 14-Sutton R.S., and Barto A. G. 1998. *Reinforcement Learning: An Introduction*. Cambridge University Press.
- 15- Wagemaker R. 2005. Model predictive control on irrigation canals application of various internal models. Delft University of Technology, Netherland.
- 16- Zhang Q., Wu C.H., and Tilt K. 1996. Application of fuzzy logic in an irrigation control system. *Industrial Technology, Proceedings of Institute of Electrical and Electronics Engineers International Conference on*, Institute of Electrical and Electronics Engineers, New York.



Development of Reinforcement Learning Algorithm for Automation of Slide Gate Check Structure in Canals

K. Shahverdi¹ – M.J. Monem^{2*}

Received: 15-11-2013

Accepted: 21-06-2015

Introduction: Nowadays considering water shortage and weak management in agricultural water sector and for optimal uses of water, irrigation networks performance need to be improved. Recently, intelligent management of water conveyance and delivery, and better control technologies have been considered for improving the performance of irrigation networks and their operation. For this affair, providing of mathematical model of automatic control system and related structures, which connected with hydrodynamic models, is necessary. The main objective of this research, is development of mathematical model of RL upstream control algorithm inside ICSS hydrodynamic model as a subroutine.

Materials and Methods: In the learning systems, a set of state-action rules called classifiers compete to control the system based on the system's receipt from the environment. One could be identified five main elements of the RL: an agent, an environment, a policy, a reward function, and a simulator. The learner (decision-maker) is called the agent. The thing it interacts with, comprising everything outside the agent, is called the environment. The agent selects an action based on existing state in the environment. When the agent takes an action and performs on environment, the environment goes new state and reward is assigned based on it. The agent and the environment continually interact to maximize the reward. The policy is a set of state-action pair, which have higher rewards. It defines the agent's behavior and says which action must be taken in which state. The reward function defines the goal in a RL problem. The reward function defines what the good and bad events are for the agent. The higher the reward, the better the action. The simulator provides environment information. In irrigation canals, the agent is the check structures. The action and state are the check structures adjustment and the water depth, respectively. The environment comprises the hydraulic information existing in the canal. Policy is a map of water depth-check structure pairs. Reward function is defined based on the difference between water depth and target depth, and the simulator is a hydrodynamic model which, in the present study, was Irrigation Conveyance System Simulation (ICSS). In the developed RL, the RL begins with required initializations, and then the canal structures are operated. While the maximum reward is reached at the time step of t , the agent receives some representation of the state of the environment and, on that basis, selects an action. The simulator performs the action and provides information on the state of the new environment by simulating the canal system. Finally, the reward is assigned. Maximizing the reward, the RL goes on to the next time step. This process is continued until the final simulation time step is reached. The learning process is similar for all operations. The ICSS hydrodynamic model was used to simulate the canal system and provide the environmental information. Input to the ICSS was the new selected action. The ICSS performed the action and simulated the canal system. The output from the ICSS was information on the new environment for use in the next time step. Two scenarios of flow increase and decrease with initial flow of 25 l/s were simulated. MAE (Maximum Absolute Error), IAE (Integral of Absolute magnitude of Error) and SRT (System Response Time) indicators have been used to assess developed model. For flow decrease scenario, the indicators value are obtained zero.

Results and Discussion: Results were obtained from the performed scenarios. In the flow increase scenario, water depth variations were inside the dead band, therefore, SRT indicator was obtained zero. The MAE and IAE indicators were obtained 3.5% and 2.57%, respectively, which showed the water depth deviations from target depth was very low. In the flow decrease scenario, the all indicators values were obtained zero. At time zero in two scenarios, 1000 populations were generated and tested. As the RL controlled the water depth, it generated new populations, too. The reason for this is that the RL generates a new population if there is no classifier with maximum reward in the population. There are no new generation after 0.03 hr and 0.16 hr in flow increase and flow decrease scenarios, respectively. Considering the results, it could be concluded that the developed control system is a powerful technique in terms of accuracy and response time for water depth

1 , 2- Ph.D. Graduated and Professor of Water Structures Engineering Department, Tarbiat Modares University
(*- Corresponding Author Email: javadmonem@gmail.com)

control.

Conclusion: In this research, the RL upstream control system was developed and connected with ICSS hydrodynamic model and evaluated in two scenarios of flow increase and flow decrease. The results showed an ability to control of deviations, short response time and accurate performance of the developed RL control system, which could be used for further study in irrigation canals.

Keywords: Learning, Model development, Performance improvement