ABSTRAK

Pencarian sumber daya yang efektif dalam lingkungan yang tidak pasti merupakan tantangan mendasar dalam pengembangan agen otonom. Strategi pencarian acak seperti *Lévy Flight* telah banyak diteliti karena efisiensinya yang mendekati optimal; namun, performanya sangat bergantung pada nilai parameter dan karakteristik lingkungan sehingga sulit beradaptasi secara mandiri. Penelitian ini mengusulkan *Monte Carlo Reinforcement Learning* (MCRL) sebagai strategi pencarian adaptif yang dapat mempelajari dan menyesuaikan pola pencariannya melalui interaksi langsung dengan lingkungan, tanpa memerlukan pengetahuan awal tentang lokasi target. Agen dimodelkan dengan dua aksi jalan lurus (*straight*) dan belok (*turn*) di mana *state* didefinisikan sebagai jumlah langkah lurus berturut-turut sebelum berpindah arah, sehingga karakteristik distribusi panjang langkah *Lévy Flight* dapat direpresentasikan dalam kerangka *reinforcement learning*. Eksperimen dilakukan menggunakan *ONE Simulator* pada empat skenario lingkungan yang merupakan kombinasi antara dua pola persebaran target (*Homogeneous Poisson Process* dan *Thomas Cluster Process*) dan dua kondisi penempatan target tetap (*Fixed*) dan teracak (*Randomized*). Tiga strategi eksplorasi-eksploitasi dievaluasi: *Epsilon-Greedy*, *Upper Confidence Bound* (UCB), dan *Thompson Sampling* dalam dua varian (*Beta-Binomial* dan *Gaussian*). Hasil menunjukkan bahwa MCRL dengan skema pembaruan *every-visit* dan *Beta-Binomial Thompson Sampling* mencapai *cumulative reward* dan cakupan target tertinggi di sebagian besar skenario, dengan cakupan hingga 100% pada hari ke-7 simulasi. Meskipun *Lévy Flight* mendominasi pada fase awal karena tidak memerlukan fase pembelajaran, agen berbasis MCRL secara bertahap melampaui performanya seiring bertambahnya episode. Temuan ini menunjukkan bahwa bahkan dengan representasi *state* yang sederhana, MCRL mampu mengembangkan strategi pencarian yang adaptif dan melampaui strategi pencarian acak konvensional.

Kata kunci: strategi pencarian sumber daya, *lévy flight*, *q-learning*, strategi eksplorasi-eksploitasi

ABSTRACT

Effective resource foraging in uncertain environments is a fundamental challenge in autonomous agent research. Random search strategies such as *Lévy Flight* have been widely studied for their near-optimal efficiency in target discovery; however, their performance is highly sensitive to parameter values and environmental characteristics, limiting their ability to adapt independently. This study proposes *Monte Carlo Reinforcement Learning* (MCRL) as an adaptive foraging strategy that learns and adjusts its search behavior through direct interaction with the environment, without requiring prior knowledge of target locations. The agent is modeled with two actions *straight* and *turn* where the state is defined as the number of consecutive straight steps taken before changing direction, effectively capturing the characteristic step-length distribution of *Lévy Flight* within a reinforcement learning framework. Experiments were conducted using the ONE Simulator across four environmental scenarios, combining two target distribution patterns (*Homogeneous Poisson Process* and *Thomas Cluster Process*) with two target placement conditions *Fixed* and *Randomized*. Three exploration-exploitation strategies were evaluated: *Epsilon-Greedy*, *Upper Confidence Bound* (UCB), and *Thompson Sampling* in two variants (*Beta-Binomial* and *Gaussian*). Results show that MCRL with the *every-visit* update scheme and *Beta-Binomial Thompson Sampling* achieves the highest cumulative reward and target coverage across most scenarios, reaching up to 100% coverage by the seventh simulation day. While *Lévy Flight* remains dominant in the early episodes due to its parameter-free nature requiring no learning phase, MCRL-based agents progressively surpass its performance as learning accumulates. These findings demonstrate that even a simple single-state representation is sufficient for MCRL to develop an adaptive foraging strategy that surpasses conventional random search.

Keywords: foraging strategy, lévy flight, q-learning, exploration-exploitation strategy