

控制 MMORPG 內通貨膨脹:

基於強化學習的NPC商店多商品動態定價

Controlling In-Game Inflation in MMORPGs:

Reinforcement Learning-Based Multi-Product Dynamic Pricing for NPC Stores

專題成員:張凱盛 指導教授:陳朝鈞

介紹

MMORPG中的通貨膨脹是指因貨幣供應過剩導致價值下降,常見於玩家透過重複性任務或資源採集迅速累積並販售給NPC 商店換取金幣。若商品價格未即時調整,將導致金幣大量流入市場,購買力下降、經濟失衡。人工每日調價既耗時又難以掌握市場即時變化,為實現自動化與即時的經濟調控,本研究引入強化學習模型,根據每日交易數據調整商品價格。透過與環境互動學習,模型可掌握市場變化並調整定價策略,實現供需平衡與通膨抑制。實驗結果顯示該方法能有效穩定遊戲內金幣流通與價值。

方法 Critic 2 Critic 1 Actor $H\pi(a|s)$ Loss function Loss function Loss function **NPC** Player Update 9 Update o Q1(s,a)Resource Q1'(s,a) Replay Memory <s,a,r,s'>

圖一、Soft Actor-Critic架構

1. 強化學習與 Soft Actor-Critic (SAC) 的應用動機

市場價格調控屬典型大數據問題,適合以AI進行分析與決策。傳統監督式學習仰賴靜態資料學習固定分布,難以應對市場即時變化;強化學習則透與過環境互動,調整策略以適應非靜態分布。而價格調控屬於連續動作,離散方法如DQN在商品數量增加時動作空間呈指數性擴張。圖一SAC模型原生則支援連續動作,並透過熵正則化提升探索效率,適用於多商品價格調控場景。

2. 強化學習架構設計

State (標準化至(-1, 1)):

- 1. 買賣價/買賣人次與交易量
- 2. 金幣 Sink/CPI/玩家持幣量

Action:

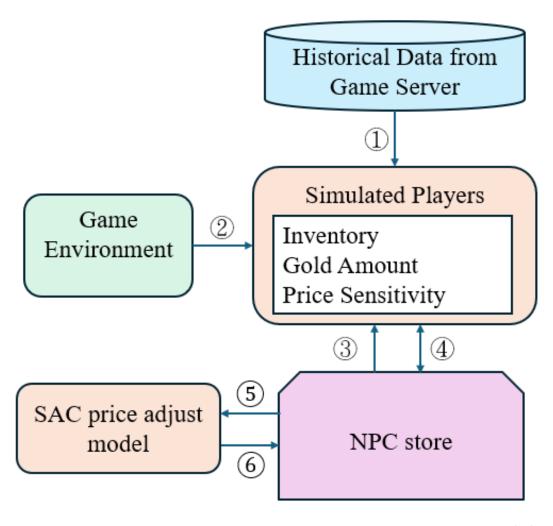
1. 每商品買/賣價連續微調

Reward (負偏差):

- 1. 人數偏差
- 2. 交易量偏差
- 3. CPI 變動率
- 4. 玩家金幣變動率

狀態設計讓模型全面感知經濟環境。動作的連續微調實 現細緻調控。**獎勵函數**兼顧市場活躍度與物價、貨幣穩定性。

3.強化學習模擬環境設計



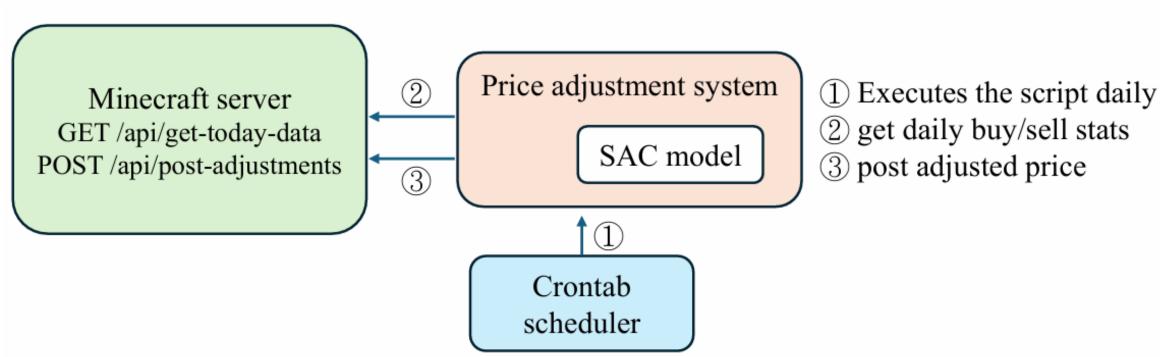
Simulated player behavior is inferred using logistic regression and item co-occurrence patterns from historical data.

- 1 Logistic regression + Apriori rules
- ② Resource Input (e.g., mining)
- ③ Product Prices from NPC Store
- 4 Player Trade Response
- 5 Daily buy/sell statistics sent to SAC model
- (6) Adjusted prices sent back to NPC store

圖二、模擬環境架構

圖二為本研究所設計並建構的模擬環境。實際遊戲每日交易次數有限,無法提供足夠資料支撐完整訓練流程。為此,本研究根據玩家歷史交易紀錄(Historical Data),並結合庫存(Inventory)、金幣數量(Gold Amount)與價格敏感度(Price Sensitivity)等心理特徵,設計了一套模擬環境,以確保在不干擾真實玩家體驗的前提下,驗證模型的學習成效。

4.真實遊戲伺服器自動調整價格系統



圖三、自動化價格調整系統架構

圖三說明了伺服器每日自動調整價格流程。系統由Crontab scheduler定時觸發,取得交易資料,價格調整系統(Price adjustment system)透過 SAC 模型計算調整價格,並更新至遊戲server,以實現動態調整與市場反應。

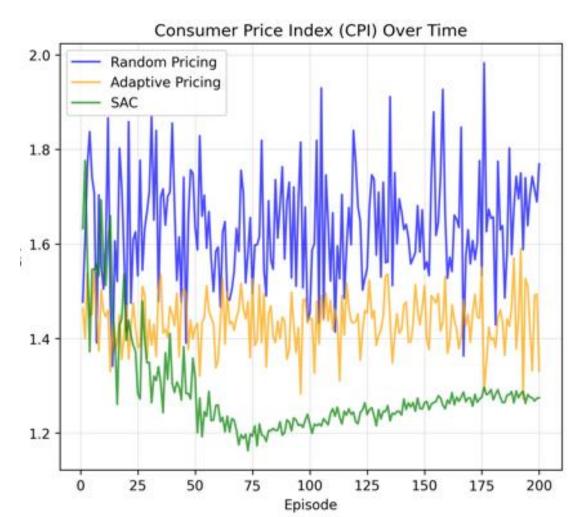
實驗結果

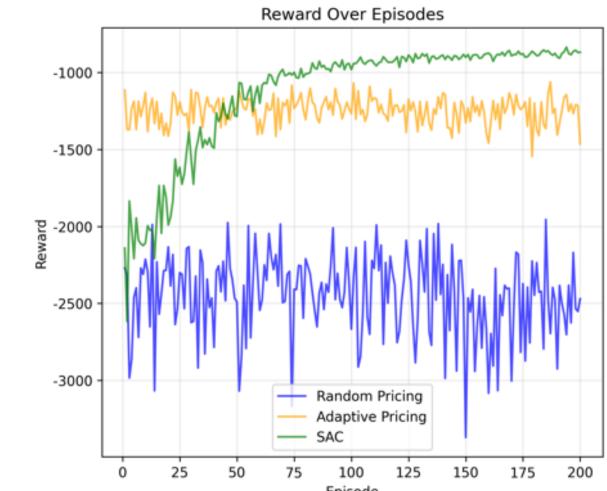
圖四為三種定價策略對物價指數(CPI)的控制效果比較: 消費者物價指數(CPI)是反映商品與價格變動情況的指標, 用來衡量通膨程度。

綠線(SAC強化學習):初期CPI較高但迅速下降並穩定於約1.2左右,顯示其能有效抑制物價膨脹,達成長期穩定。

藍線(Random Pricing):波動劇烈, CPI 長期維持在 1.4~2.0 區間,顯示隨機策略無法穩定物價,通膨風險高。

橘線 (Adaptive Pricing): CPI 稍低於隨機策略,但波動仍大,說明簡單規則調價有改善但效果有限。





圖四、三種定價策略物價指數變化

圖五、三種定價策略之Reward變化

圖五顯示 SAC 模型的 Reward 穩定上升,最終優於其他策略,代表其已學會有效且穩定的調控行為,達成本研究所追求的市場平衡目標。本模型亦已實際部署於遊戲伺服器,圖六為對應的 NPC 商店畫面。

此外,本研究亦設計了網頁模擬平台,用以展示環境運作機制並同步觀察市場變化。圖七為實際伺服器市場狀況之畫面,圖八則為所設計之模擬平台介面。





圖六、遊戲NPC商店畫面

圖七、遊戲市場交易狀況



圖八、模擬環境demo網頁

結論

本研究針對 MMORPG 中多商品價格難以即時調控與通膨失衡的問題,提出一套基於 Soft Actor-Critic 強化學習之價格調控方法。透過結合模擬環境進行訓練,並實際部署於遊戲伺服器,該方法能依據市場變化動態調整價格。實驗結果顯示,本模型在穩定物價指數與抑制通膨方面,表現優於隨機與傳統策略,展現出高度穩定性與自我調整能力,證實其應用於遊戲經濟調控的可行性與有效性。