

# 基于多 Agent 系统的自调节及协同工作的组合投资模型

王雪松<sup>1</sup>, 申群太<sup>2</sup>

(1. 中山火炬职业技术学院 信息工程系, 广东 中山 528436;

2. 中南大学 信息科学与工程学院, 湖南 长沙 410083)

**摘 要:** Agent 技术特别是多 Agent 系统(MAS, Multi-Agent System)为解决人工智能等领域复杂问题提供了一个新途径, 多 Agent 系统重点研究如何协调系统中的各个 Agent 的行为使其协同工作。针对多阶段组合投资问题, 提出了一个基于多 Agent 系统的自调节及协同工作的组合投资策略模型。该模型系统中的各个 Agent 通过通讯共享知识, 在求解问题的搜索空间中进行协同搜索, 在更短的搜索步长内得到问题的解, 极大地提高了系统性能。该模型具有不基于任何股票模型、时间复杂度低以及逼近最优投资策略速度较快等优点, 实验证明具有一定的实际意义。

**关键词:** 在线多阶段组合投资; 多 Agent 系统; 在线学习; 协同搜索

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2010)05-0117-04

## Combination Investment Model of Auto-Adjustment and Cooperate with Work Based on Multi-Agent System

WANG Xue-song<sup>1</sup>, SHEN Qun-tai<sup>2</sup>

(1. Department of Information Engineering, Zhongshan Torch College, Zhongshan 528436, China;

2. School of Information Science and Engineering, Central South University, Changsha 410083, China)

**Abstract:** Agent technology, especially MAS (Multi-Agent System) provides a new method that can greatly resolve the complex problems in fields such as artificial intelligence, focus on multi-agent system, how to coordinate the various systems work together so that it acts Agent. In the paper, propose to the combination investment model of auto-adjustment and cooperation based on multi-agent system by multi-phases combination investment problem. In the model, every agent carry through cooperation search by the search space of the solution problems by communication sharing knowledge. The solution of the problems is gained by the shorter search step, and the system performance is maximum increased. The model has these excellence of the any stock model no relation, the low of time complexity and the speed fast of the approach optimal investment. It is signification from the experimentation show.

**Key words:** online multi-phase combination investment; multi-agent system; online learning; cooperation search

## 0 引言

智能 Agent 是人工智能领域发展起来的具有感知能力、问题求解能力和外界进行通讯能力的一个实体<sup>[1]</sup>。多 Agent 系统是分布式人工智能研究的一个前沿领域, 它重点研究如何协调系统中的各个 Agent 的行为使其协同工作<sup>[2]</sup>。构造结构复杂、知识丰富和功能强大的单 Agent 系统和由多个结构、功能较为简单

的单 Agent 组合的多 Agent 系统<sup>[3]</sup>是 Agent 技术的两个主要发展方向。多 Agent 系统是指一个为了完成某些任务或者达到某些目标, 由多个 Agent 构成进行协同工作的计算机系统。它可由众多同构的或者异构的 Agent 组成。一个进行协同搜索的多 Agent 系统中, Agent 的特征包括多 Agent 系统的总体性能描述、单独的 Agent 的搜索算法、Agent 的搜索结果是否需要合并及如何合并等。

Clearwater 等人提出了 Agent 之间能够共享信息的机制<sup>[4]</sup>, Agent 之间的信息共享是通过一个所有 Agent 都能够存取的“中央黑板”进行的。不同类型、结构的数据通过独立的黑板来模型化, 从而使这些信息能够被所有不同构的 Agent 进行存取。在每一个搜索

收稿日期: 2009-07-02; 修回日期: 2009-10-26

基金项目: 湖南省教育科学“十一五”规划项目(XJK06CXJ022)

作者简介: 王雪松(1974-), 男, 湖南祁阳人, 副教授, 硕士, 研究方向为软件技术、数据挖掘、人工智能; 申群太, 教授, 博士生导师, 研究方向为决策控制、数据挖掘、人工智能。

阶段,一个 Agent 随机选择一个黑板上的信息(信息经过黑板系统的处理),然后使用此信息上的设定来替换 Agent 原来的一些状态参数。当一个 Agent 完成一个阶段的搜索工作时,它把自己的状态以及得到的可能有用的信息发布到黑板上。

证券组合投资的研究在经济领域和统计学领域中已经取得了相当的成果,基于不同的理论基础和市场模型上有很多不同的组合投资优化算法<sup>[5]</sup>。然而,这些算法往往受限于它所依赖的市场模型对实际市场模拟的精确程度或者算法的假设条件与实际市场的符合程度<sup>[6,7]</sup>。目前,博弈理论、在线学习等理论也应用到组合投资优化问题的解决上,并且取得了比较好的效果。文中基于在线学习的自动修正组合投资选择算法(AU)<sup>[8]</sup>的基础上,构建了一个自调节及协同工作的多 Agent 系统组合投资模型。

## 1 基于在线学习的自动修正组合投资选择算法

在每一个投资阶段的开始,算法将根据之前的投资策略  $W^t$  和相应的价格关系  $P^t$  来计算新的投资策略  $W^{t+1}$ 。根据 Kivinen 的思想,当选择一个与  $W^t$  接近的  $W^{t+1}$  时候能够取得好的效果。在函数(1)中找出使其达到最大值的  $W^{t+1}$ :

$$F(W^{t+1}) = \beta \log(W^{t+1} \cdot P^t) - d(W^{t+1}, W^t) \quad (1)$$

其中  $\beta$  是学习速率,而  $d$  是一个距离函数,用来保持  $W^{t+1}$  接近于  $W^t$ 。不同的距离函数  $d$  可能导致了不同修正规则。以相对熵作为距离函数:

$$d(W^{t+1} \| W^t) = \sum_{i=1}^N w_i^{t+1} \log \frac{w_i^{t+1}}{w_i^t} \quad (2)$$

因此对于函数(1)做出如下修改:

$$F(W^{t+1}, \alpha) = \beta(\log(W^t \cdot P^t) + \frac{P^t \cdot (W^{t+1} - W^t)}{W^t \cdot P^t}) - d(W^{t+1}, W^t) + \alpha(\sum_{i=1}^N W_i^{t+1} - 1) \quad (3)$$

考虑  $\sum_{i=1}^N w_i^{t+1} = 1$ , 即得:

$$w_i^{t+1} = \frac{w_i^t \exp(\beta p_i^t / W^t \cdot P^t)}{\sum_{j=1}^N w_j^t \exp(\beta p_j^t / W^t \cdot P^t)} \quad (4)$$

这即为算法中使用的投资策略修正规则。

## 2 多 Agent 的自调节及协同工作组合投资模型系统

在此多 Agent 系统中,各个 Agent 获得的价格关系不基于任何股票价格的统计模型。所有的 Agent 都共同相等享有初始资本,在进行一段时间的投资之后,

它们的资本的总和减去初始资本即为最后获得的投资收益。模型系统中,每个 Agent 使用不同的初始组合投资策略,利用 AU 算法在每一个投资阶段对自身的投资策略进行评估和优化。在 Agent 之间加入一个协同搜索的机制,通过对各个 Agent 当前的投资效果的交换、评估,在 Agent 之中进行策略的重新选择。一个 Agent 可能以某个概率选择将自己的投资策略改变为投资效果最好的 Agent 的策略。

在具体实现中,所有的 Agent 采用 AU 算法进行策略的修正计算,各个 Agent 选择不同的初始组合投资策略,Agent 采用随机的学习速率  $\eta$ ,并根据 AU 算法计算下阶段的组合投资策略  $W'$ ,  $W'$  的各个元素为非零实数并且它们的和等于 1。

由 AU 算法知道,如果 Agent 之间不进行策略的转换,那么 Agent 的未来组合投资策略序列就由它的学习速率、当前的组合投资策略以及未来的价格序列  $\{P^t\}$  所决定。由于所有 Agent 都采用同样的策略修正算法,当 Agent A 把它的投资策略转换为 Agent B 时,它未来组合投资序列就会与 B 接近,除非它们之间的一个又改变了策略。可以发现,通过对 Agent 之中最成功的策略的信息的通讯,并由此改变某些 Agent 的组合投资策略,可以使多 Agent 系统的有效组合投资策略更快地逼近 BCRP(Best Constant Rebalance Portfolio, 取得最好收益的固定重平衡组合投资)组合投资策略。

多 Agent 系统的有效组合投资策略是指在同样的初始资本的基础上,该策略在投资结束的时候取得的投资收益跟系统中所有 Agent 的投资收益的总和相同。定义有效组合投资策略  $W_e^t$  为:

$$W_e^t = \sum_{i=1}^M \left[ \frac{\text{wealth}_i^t}{\sum_{j=1}^M \text{wealth}_j^t} W_i^t \right]$$

其中,  $\text{wealth}_i^t$  是 Agent  $i$  在投资阶段  $t$  的资产,  $M$  是系统中 Agent 的数量。

多 Agent 的自调节及协同工作组合投资模型: Agent 之间交换 Agent 的投资收益情况,并可选择将自己的组合投资策略转换到现阶段投资收益最好的 Agent 的组合投资策略。每个 Agent 维持其在最近  $\tau$  个投资阶段的平均投资收益,  $\tau$  称为收益窗口大小。一个在最近  $\tau$  个投资周期没有改变过投资策略的 Agent 在每一个阶段结束时将其现阶段组合投资策略以及平均投资收益粘贴到系统的中央黑板上,中央黑板维持当前最佳的组合投资策略以及其平均收益。在最近  $\tau$  个投资阶段没有改变过策略的 Agent 将自己的投资策略的平均收益与中央黑板进行比较,如果中央黑板上的最佳组合投资策略的平均收益大,则该 Agent 以一定

的概率  $p$  选择转换自己的组合投资策略为最佳组合投资策略,其中  $p$  称为转换概率。

### 3 模型的算法实现

算法是通过一个具有自调整能力的 Agent 组来实现的。参数  $p, \tau$  分别代表了系统中 Agent 的策略转换概率和收益窗口大小。函数  $m. average(m. W * P, \tau)$  用来计算 Agent  $m$  在最近  $\tau$  阶段内的平均每阶段投资收益。 $m. post()$  把 Agent  $m$  当前的投资策略以及投资收益发布到中央黑板上。 $m. switch(p)$  有两种返回情况:它或者返回 Agent  $m$  的投资策略以及  $count(m)$  的投资策略不改变),或者返回最佳的组合投资策略并把  $count$  设为 0( $m$  的投资策略改变)。

```
Global : T//投资阶段
        N//股票数量
        M//Agents 的数量
        V//初始投资
AdaptiveAndComMAS( $\eta$  Row,  $\eta$  High,  $p, \tau$ )
//  $\eta$  Row,  $\eta$  High 代表学习速率的范围,  $p$  是改变策略的概率;  $\tau$  是 Agent 之间对比收益的阶段数
{
    for each  $m \in M$ 
    {
         $m. wealth = V/M$ ; //按 agent 的数量平均初始资产
         $m. W = rand(N)$ ; //赋予 Agent 一个随机的 CRP 策略
         $m. \eta = Random(\eta Row, \eta High)$ ; //对学习速率赋一个随机值
         $m. count = 0$ ; //记录 Agent 从最近一次改变策略以来的阶段数
    }
    price = (1, 1, ..., 1); //初始化价格关系
    t = 1;
    while(t < T)
    {
        //开始投资阶段
        for each  $m$  in M
        {
             $m. trade(price)$ ;
        }
        (price, P) = price-relatives(t);
        //投资阶段结束
        for each  $m$  in M
        {
```

```
             $m. wealth = m. wealth * (m. W * P)$  //更新 agent 的资产
```

```
             $m. perf = m. average(m. W * P, \tau)$ ; //计算最近  $\tau$  阶段平均收益
```

```
             $m. count++$ ;
```

```
            if ( $m. count \geq \tau$ )
```

```
             $m. post()$ ; //与中央黑板进行比较
```

```
            }
```

```
        for each  $m$  in M
```

```
        {
```

```
            if ( $m. count \geq \tau$ )
```

```
            {
```

```
                [ $m. count, m. W$ ] =  $m. switch(p)$ ; //以概率  $p$  转化为最优策略
```

```
                 $m. W = m. update(P)$ ;
```

```
            }
```

```
        }
```

```
        t = t + 1
```

```
    }
```

```
    return  $\sum_{所有m} m. wealth$ ;
```

```
}
```

### 4 Agent 通讯的实现

利用 JKQML 实现 Agent 的通讯。在此方案中包括了三种 Agent: PortfolioAgent 是运用 AU 算法进行组合投资的 Agent, BlackboardAgent 是实现中央黑板的 Agent, 还有一个 Agent 就是 Facilitator。

在系统中,所有的 Agent 都是使用 AU 算法进行计算,它们的区别在于算法的学习速率不同,可以在 Agent 的初始化进行设定。下面给出系统的主要的通讯语句和流程:

1) BlackboardAgent 向所有的 Agent 发布接受投资策略和平均收益信息。

```
(advertise
```

```
:sender BlackboardAgent
```

```
:receiver facilitator
```

```
:reply - with id - BlackboardAgent
```

```
:language KQML
```

```
:ontology kqml - ontology
```

```
:content (insert
```

```
    :sender facilitator
```

```
    :receiver BlackboardAgent
```

```
    :in - reply - to id - BlackboardAgent
```

```
    :language Java
```

```
    :ontology kqml - ontology
```

```

:content (
    ":Stock1,
    :Stock2
    :Wealth"
)
)

```

在 BlackboardAgent 的内容解释程序中,对 Content 的内容做处理,保留最佳的策略和投资收益。

2) BlackboardAgent 向所有的 Agent 广播当前最优策略信息。

```

(broadcast
:sender BlackboardAgent
:receiver facilitator
:reply - with id - BlackboardAgent
:language Java
:ontology kqml - ontology
:content (":Stock1,
    :Stock2
    :Wealth"
)
)

```

在 PortfolioAgent 的内容解释程序中,对 Content 的内容做处理,做出自己是否选择转换策略。

3) PortfolioAgent 向 Facilitator 提交当前的策略和投资收益。

```

(insert
:sender PortfolioAgentX
:receiver facilitator
:in - reply - to id - facilitator
:reply - with id - PortfolioAgentX
:language Java
ontology kqml - ontology
content (
    ":Stock1 +
    :Stock2;
    :Stock3"
)
)

```

## 5 实验结果

在 AU 算法的实验数据上进行多 Agent 系统的实验。自调节及协同工作的多 Agent 系统组合投资模型与 AU 算法的比较。

通过考察在不同的投资周期长度下,自调节及协同工作的多 Agent 系统组合投资模型与 AU 算法对 BCRP 策略的逼近程度作分析和比较。实验的参数设

置为:AU 算法学习速率为 0.13;自调节及协同工作的多 Agent 系统组合投资模型:Agent 的数量为 20,Agent 的学习速率为[0.1,0.3]。具体的实验结果见表 1。

表 1 实验结果

算法	周期长度				
	100	200	300	400	500
自调节及协同工作多 Agent 模型	1.77	2.76	3.83	4.68	5.37
AU 算法	1.45	2.22	3.37	4.50	5.07
BCRP 策略	2.16	3.02	3.99	4.79	5.38

结论:通过多 Agent 系统的协同搜索,算法计算得到的有效组合投资策略能够比 AU 算法更快地接近 BCRP 策略。也就是说,在周期短的情况下,本模型能够获得比 AU 算法更好的投资收益。

## 6 结束语

在 AU 算法的基础上,构建了一个自调节及协同工作的多 Agent 组合投资模型。通过在模型中引入协同搜索机制,Agent 之间共享最优策略以及 Agent 转换自身的策略为最优策略等方法,加快了算法逼近 BCRP 策略的速度。由于 AU 算法无法在算法的初始阶段确定投资的最优学习速率(它的确定只能依靠先验知识),通过多个 Agent 之间不同的学习速率进行学习,从实践中可以证明它的有效组合投资策略更能够准确地逼近 BCRP 策略,取得更高的投资收益。

### 参考文献:

- [1] Davis S R. Negotiation as a metaphor for distributed problem solving[J]. Artificial Intelligence, 1983(20):63-109.
- [2] 伍尔德里奇. 多 Agent 系统引论[M]. 石纯一等译. 北京:电子工业出版社, 2003.
- [3] 赵龙文. 多 Agent 系统及其组织结构[J]. 计算机应用研究, 2000, 17(7):12-14.
- [4] Clearwater S H, Huberman B A, Hogg T. Cooperative problem solving[C]//Huberman B A. in Computation: The Micro and Macro View, World Scientific. Singapore: [s. n.], 1992: 33-70.
- [5] Jin Han-jun, Wang Hong-feng. Simulation analysis on the problem of portfolio investment based on genetic algorithm [J]. Journal of central China normal university: Nat. Sci., 2004(12):427-429.
- [6] Ross S A. The arbitrage theory of capital asset pricing[J]. Journal of Economic Theory, 1976(13):341-344.
- [7] 中国证券监督管理委员会. 证券组合管理[M]. 北京:科学出版社, 2001.
- [8] 彭佳红, 彭佳文, 贺智勇. 基于在线学习的自动修正组合投资选择算法[J]. 计算机工程与设计, 2005, 26(6):1460-1462.