

**МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ
ФЕДЕРАЦИИ
ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«АСТРАХАНСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
ИМЕНИ В.Н. ТАТИЩЕВА»**

Факультет цифровых технологий и кибербезопасности

Кафедра информационных технологий

КУРСОВАЯ РАБОТА

**Система климат-контроля помещений с поддержанием оптимальных
условий микроклимата по СанПиН 2.2.4 548-96 на основе марковских
процессов принятия решений**

выполнена в рамках изучения дисциплины
«Модели информационных процессов и систем»

Направление подготовки: 09.04.04 Программная инженерия Направленность
(профиль): «Проектирование и разработка систем искусственного интеллекта»

Исполнитель: студент группы ПИ-15

_____ Мартынов В.А.

Научный руководитель: доцент кафедры ИТ,

_____ Верига А.В.

Оценка: _____

«___» _____ 20__ г.

СОДЕРЖАНИЕ

Задание	3
Введение.....	4
Основная часть	5
1. Постановка задачи в терминах Марковского процесса принятия решения (МППР)	5
1.1 Среда.....	5
1.3 Состояния (S).....	6
1.4 Действия (A).....	6
1.5 Награды ($R(s,a)$).....	6
2. Пути решения задачи	7
2.1 Обучение с подкреплением для МППР	7
2.2 Алгоритмы для поиска оптимальной стратегии	8
2.3 Пример применения	9
Заключение	12
Список использованных источников	13

ЗАДАНИЕ

Необходимо разработать теоретическую модель, описывающую принцип работы системы климат-контроля помещений с поддержанием оптимальных условий микроклимата по СанПиН 2.2.4 548-96 в терминах Марковского процесса принятия решения (МППР). Для этого:

- Формализовать задачу, выделив основные компоненты МППР: состояния, действия, награды, вероятности переходов;
- Описать методы решения задачи на основе подходов обучения с подкреплением;
- Рассмотреть пример реализации рассматриваемой модели и её применение согласно прямому назначению.

Ключевая цель: показать, что процесс контроля климатических условий в помещениях можно эффективно моделировать как задачу принятия решений в рамках МППР.

ВВЕДЕНИЕ

Современные условия жизни и работы требуют создания комфортной и здоровой среды в помещениях, что особенно актуально в условиях изменяющегося климата и увеличения уровня загрязнения окружающей среды. Системы климат-контроля играют ключевую роль в обеспечении оптимальных параметров микроклимата, таких как температура, влажность и качество воздуха.

Целью данной курсовой работы является разработка системы климат-контроля помещений с использованием МППР, которая позволит автоматизировать процессы управления климатическими параметрами, обеспечивая при этом экономию энергоресурсов и улучшение качества воздуха. В рамках работы будет описана модель системы на основе МППР, алгоритмы управления. Ключевая особенность модели будет заключаться в ее способности адаптироваться к изменяющимся условиям и требованиям пользователей.

Актуальность темы обусловлена необходимостью создания интеллектуальных систем, способных не только поддерживать комфортные условия, но и минимизировать негативное воздействие на окружающую среду.

ОСНОВНАЯ ЧАСТЬ

1. Постановка задачи в терминах Марковского процесса принятия решения (МППР)

1.1 Среда

Средой для модели является замкнутое помещение с возможностью подачи воздуха определенного качества через систему взаимосвязанных электроустановок, целью которых будет придать воздуху определенное значение параметра.

Среда включает в себя следующие параметры:

- Относительная влажность воздуха в помещении;
- Температура в помещении.

Среда динамична, её состояние изменяется под влиянием внешних факторов окружающей среды, таких как температура и влажность атмосферы снаружи помещения.

1.2 Система

В рамках курсовой работы системой будет называться набор электроустановок, а именно:

- Обогреватель воздуха;
- Кондиционер;
- Увлажнитель воздуха.

Обладая данным набором инструментов, система может эффективно влиять на состояние атмосферы внутри помещения. Следует учитывать, что активация обогревателя или кондиционера будет осушать воздух, таким образом системе необходимо учитывать данное свойство для достижения оптимальных условий микроклимата. Основная особенность системы состоит в том, что она способна комбинировать работу электроустановок, для достижения максимальной эффективности своей работы.

1.3 Состояния (S)

Приведенные состояния описывают параметры среды, на которую влияет система. Примеры состояний:

- s_1 : Температура ниже установленной;
- s_2 : Температура оптимальная;
- s_3 : Температура выше установленной;
- s_4 : Относительная влажность ниже установленной;
- s_5 : Относительная влажность оптимальная;
- s_6 : Относительная влажность выше установленной;

Каждое состояние точно описывает текущую среду, исключая любые промежуточные состояния целевых параметров.

1.4 Действия (A)

Система на каждом этапе взаимодействия выбирает действие, направленное на достижение заданных параметров атмосферы внутри помещения. Следует учитывать, что кондиционер более оптимален для осушения воздуха, чем нагреватель. Данный факт обусловлен особенностями работы самого кондиционера, так как излишняя влага, содержащаяся в воздухе, конденсируется внутри кондиционера, а значит не попадает в среду. Примеры действий:

1. Активация обогревателя при низкой температуре;
2. Активация кондиционера при высокой температуре;
3. Активация кондиционера при высокой влажности;
4. Активация увлажнителя воздуха при низкой влажности.

1.5 Награды ($R(s,a)$)

Для корректной работы системы необходимо обозначить количество наград, которое будет лежать в диапазоне от 0 до 100 очков. За соответствие условий микроклимата установленной норме, за единицу времени система будет получать положительные очки. Например, за 1 секунду соответствия – начисление +1 балла. Активация любой электроустановки, которые доступны системе, её количество баллов будет уменьшаться на 1 за единицу времени. Так как системе будет

доступно комбинирование разных электроустановок для достижения необходимого результата, активируя несколько электроустановок общая награда системы будет уменьшаться на 1 за каждую активную электроустановку.

Таким образом, система получает численную оценку за каждое действие:

- Соответствие условий установленной норме: +1 в секунду;
- Активация электроустановки: -1 в секунду.

Для того, чтобы система имела возможность фиксировать результаты своей работы, необходимо использовать терминальное состояние по времени. Оптимальным значением времени будет – 1 час. Таким образом, каждый час система будет считывать и сбрасывать свой итоговый балл и относительно полученного значения будет приниматься решение о последующей стратегии поведения. После того, как результаты будут зафиксированы, данный цикл повторяется. Основная цель системы — максимизировать величину награды, которую она получает каждый час своей работы.

2. Пути решения задачи

В данном разделе рассматриваются основные методы и подходы к решению задачи моделирования системы климат-контроля помещений с поддержанием оптимальных условий микроклимата по СанПиН 2.2.4 548-96 с использованием Марковского процесса принятия решения (МППР). Основное внимание уделяется использованию методов обучения с подкреплением (Reinforcement Learning, RL), которые позволяют находить оптимальные стратегии поведения системы.

2.1 Обучение с подкреплением для МППР

Обучение с подкреплением (RL) представляет собой подход, при котором агент обучается взаимодействовать со средой, получая обратную связь в виде награды за свои действия. RL идеально подходит для задач, связанных с последовательным принятием решений, таких как задачи, описываемые МППР.

Основные этапы реализации RL для МППР:

1. Формализация задачи:
 - Определяются состояния (S), действия (A) и награды (R).

- Ставится задача максимизации суммарной ожидаемой награды.
- 2. Сбор данных:
 - Необходимо накопить данные о состоянии среды и о том, как воздействие системы изменяет среду.
- 3. Выбор алгоритма обучения:
 - Применяются алгоритмы, такие как Q-Learning, Deep Q-Learning, Policy Gradient или Actor-Critic, в зависимости от сложности задачи и размерности пространства состояний.
- 4. Оценка стратегии:
 - После обучения стратегии производится её тестирование и оценка эффективности в виде таких метрик, как энергопотребление системы и итоговое время, когда атмосфера не соответствовала установленным требованиям.

2.2 Алгоритмы для поиска оптимальной стратегии

2.2.1 Q-Learning

Q-Learning — один из самых популярных алгоритмов обучения с подкреплением, подходящий для задач с дискретными состояниями и действиями.

Основная идея заключается в обновлении таблицы $Q(s, a)$, которая отражает ценность выполнения действия a в состоянии s . Обновление происходит по формуле:

$$Q(s, a) \leftarrow Q(s, a) + \alpha [r + \gamma \max_{a'} Q(s', a') - Q(s, a)],$$

где α — скорость обучения, r — полученная награда, γ — коэффициент дисконтирования, $\max_{a'} Q(s', a')$ — максимальная ценность следующего состояния.

Особенности Q-Learning:

- Простота реализации.
- Подходит для небольших задач с ограниченным числом состояний и действий.
- Сложность масштабирования на задачи с высокой размерностью.

2.2.2 Deep Q-Learning (DQN)

Deep Q-Learning использует нейронные сети для аппроксимации функции $Q(s,a)$, что позволяет решать задачи с непрерывными или высокоразмерными состояниями.

Основные шаги алгоритма:

1. Инициализация нейронной сети: создаётся модель, принимающая на вход состояние s и возвращающая значения $Q(s,a)$ для каждого действия.
2. Replay Memory: собираются данные взаимодействий агента со средой для эффективного обучения.
3. Обновление весов сети: минимизируется ошибка между предсказанным значением $Q(s,a)$ и целевым значением:

$$y = r + \gamma \max_{a'} Q(s', a'; \theta),$$

где θ — веса сети.

Преимущества DQN:

- Подходит для сложных задач с большой размерностью состояний.
- Может использоваться в реальном сценарии обеспечения надлежащих условий микроклимата, так как необходимо контролировать несколько состояний среды одновременно.

2.2.3 Политики на основе градиента (Policy Gradient)

Алгоритмы на основе градиента политики, такие как REINFORCE, оптимизируют стратегию напрямую, обновляя параметры политики $\pi(a|s;\theta)$:

$$\theta \leftarrow \theta + \alpha \nabla_{\theta} \log \pi(a|s;\theta) G_t,$$

где G_t — суммарная награда, α — скорость обучения.

Преимущества:

- Подходит для задач с непрерывным пространством действий.
- Эффективен при сложной стратегии принятия решений.

2.3 Пример применения

Рассмотрим пример работы модели:

1. Пользователь установил температуру 25 градусов по Цельсию и относительную влажность 60%.

2. Система осуществляет контроль состояния среды внутри помещения и обнаруживает 20 градусов по Цельсию и относительную влажность в 80%

3. Система реагирует активацией обогревателя и кондиционера ожидая, когда состояние среды нормализуется к заданным параметрам.

4. При достижении оптимальных параметров электроустановки отключаются.

5. На начальном этапе, когда система ещё не обучена, она не будет учитывать гистерезис процессов. Таким образом она станет активировать электроустановки после нормализации, когда будет обнаружен перегрев и пересушенный воздух.

6. Система в таком случае получает значительный штраф, что вынуждает её производить поиск оптимальных параметров среды, при которых следует отключить электроустановки, чтобы учесть гистерезис.

7. Когда система станет обученной, она будет отключать электроустановки заранее, что приведет к экономии электричества и снижению итогового времени работы электроустановок, что повлияет на значение общей награды системы.

2.4 Проблемы и ограничения подхода

Несмотря на высокую эффективность, подходы на основе RL и МППР имеют следующие ограничения:

1. Сложность сбора данных. Среда крайне инерционна, что сильно растягивает время сбора необходимых данных и обучения системы.

Таким образом, подходы, описанные в данном разделе, позволяют эффективно решать задачу климат-контроля помещений с поддержанием оптимальных условий микроклимата по СанПиН 2.2.4 548-96. Использование RL-алгоритмов обеспечивает гибкость и адаптивность модели, позволяя учитывать динамические изменения среды.

ЗАКЛЮЧЕНИЕ

В ходе выполнения курсовой работы была описана климат-контроль помещений с поддержанием оптимальных условий микроклимата по СанПиН 2.2.4 548-96 на основе марковских процессов принятия решений, что позволило создать эффективное решение для автоматизации управления климатическими параметрами. Проведенный анализ существующих технологий и методов управления климатом подтвердил актуальность применения МППР, которые обеспечивают высокую степень гибкости и адаптивности системы.

Были рассмотрены ключевые компоненты задачи, включая состояния, действия, награды и вероятности переходов, а также предложили пути решения через методы обучения с подкреплением. Использование таких алгоритмов, как Q-Learning, Deep Q-Learning и Policy Gradient, позволяет эффективно моделировать состояние системы и оптимизировать стратегию взаимодействия с инструментами для достижения высокой энергоэффективности системы.

Внедрение предложенных алгоритмов управления позволяет не только улучшить качество воздуха и поддерживать заданные параметры температуры и влажности, но и значительно сократить потребление энергоресурсов, что является важным аспектом в условиях современного экологического кризиса.

Таким образом, результаты данной работы могут быть использованы как основа для дальнейших исследований и разработок в области автоматизации климат-контроля. В будущем целесообразно рассмотреть возможность интеграции системы с другими технологиями, что позволит создать более комплексные и эффективные решения для управления микроклиматом в помещениях. Внедрение таких систем будет способствовать не только улучшению качества жизни, но и устойчивому развитию городской инфраструктуры.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ