



НАЦИОНАЛЬНЫЙ ИССЛЕДОВАТЕЛЬСКИЙ
УНИВЕРСИТЕТ

Алгоритм субоптимального управления процессом проветривания подземных горнодобывающих предприятий на основе аппроксимированного динамического программирования

Проект НУГ № 21-04-039

«Динамическая оптимизация параметров контура управления киберфизической системы проветривания подземного горнодобывающего предприятия»

А.В. Кычкин, к.т.н., зав. НУЛ МЭИ, НИУ ВШЭ – Пермь



История проекта



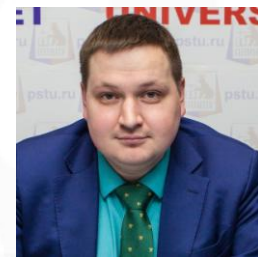
Кафедра технологий электропривода
и транспортных систем
Высшей инженерной школы
(Высшей школы технических наук)
им. Георга Агриколы (г. Бохум, Германия)



Штефан Фёт



Кафедра горной электромеханики
Пермского национального
исследовательского
политехнического университета
(г. Пермь)



Николаев А.В.

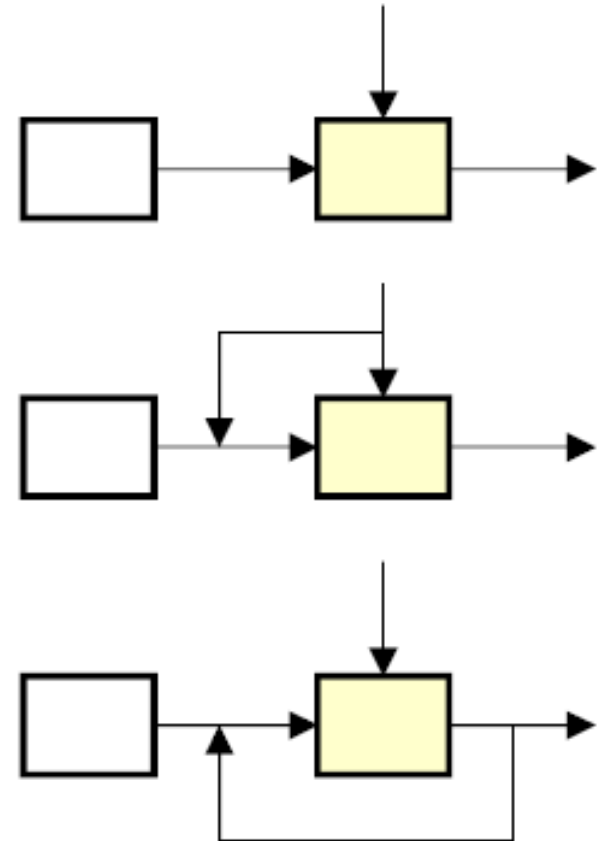


ООО «Фабрика цифровых решений»
(г. Пермь)



Русаков Ю.А.

- **Принцип программного управления** (*разомкнутое управление*).
- **Принцип компенсации** (*управление по возмущению*).
- **Принцип обратной связи** (*управление по отклонению*).



Управление «как есть»

Управление проветриванием «как есть» подразумевает ручное формирование воздействий на процесс подачи воздуха на всем рассматриваемом интервале времени (начальное и конечное состояния детерминированы)

Промежуточные сигналы управления выдаются согласно ранее определенного синтезированного плана (программы)

Сигналы не зависят от фактического поведения процесса проветривания

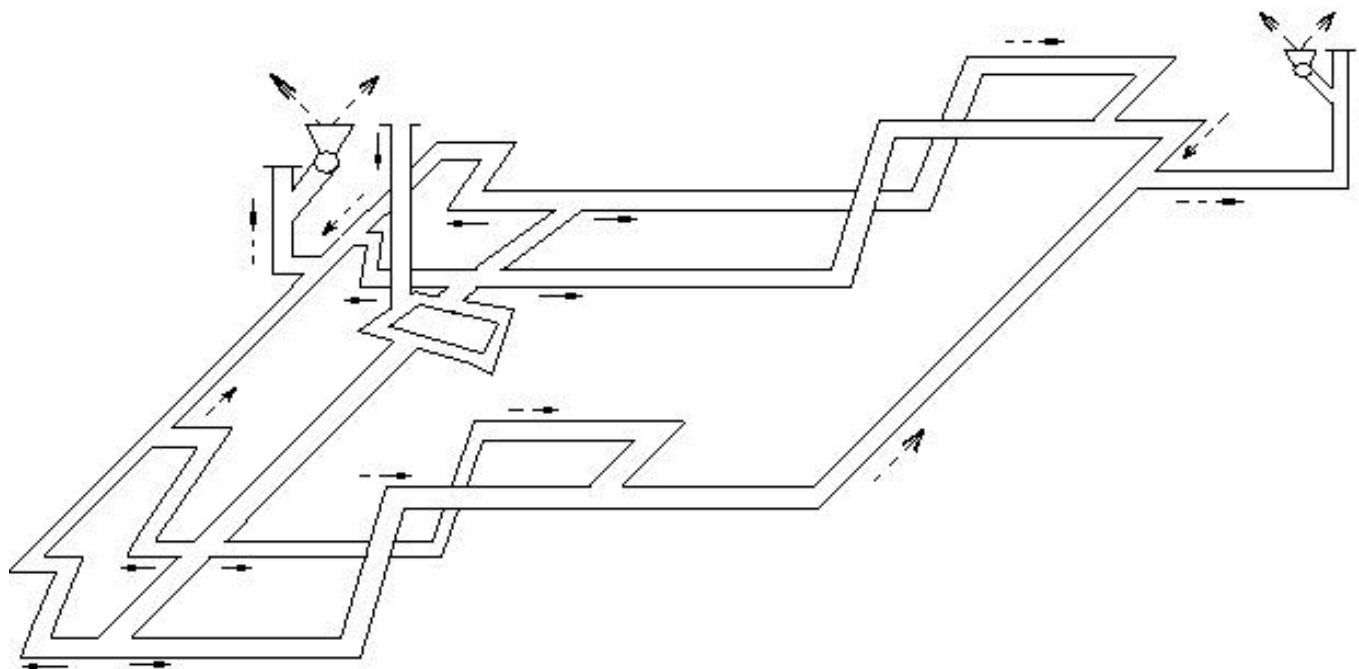
Такое автоматизированное (или автоматическое) программное управление не является оптимальным, оно не связано с состоянием процесса проветривания.



Бытовой пример системы программного управления

Системная динамика процесса

Динамика описывает зависимость состояния распределения воздуха и общих энергозатрат на реализацию этого процесса в каждый новый момент времени от всех предыдущих состояний и, как следствие, всех предыдущих управляющих воздействий



Сложность системной динамики

Причина динамики – сложность (инерционность) объекта, физические законы подготовки, подачи и распределения воздуха в подземных горных выработках

Поиск оптимальных управленческих решений требует анализа большого количества шагов вперед, чтобы моделировать возможные сценарии изменения состояний процесса проветривания и связанные с ним суммарные величины затрат на энергоресурсы, а затем выбрать наилучший сценарий управления

Это приводит к резкому нелинейному росту числа возможных последовательностей (траекторий) управляющих воздействий при линейном росте числа шагов

Имеются ограничения по исполнению алгоритмов на ПК

Ричард Беллман

«Оптимальная траектория обладает тем свойством, что каково бы ни было данное состояние системы, выбор частного решения на этом этапе должен быть таким, чтобы в совокупности с частными решениями на всех последующих этапах они составляли оптимальный план»

Данный метод эффективен при решении комплексных расчетных задач на дискретных множествах

Постановка задачи динамического управления на конечном интервале времени

Базовые характеристики задачи динамического управления проветриванием:

1. *finite horizon*
2. Это задача управления дискретным по времени динамическим процессом
3. Число шагов последовательности управляющих воздействий = 96
4. Управляющее воздействие - изменение объемного расхода воздуха, подаваемого ГВУ, относительно предустановленного значения, которое приводит к изменению энергопотребления и затрат на энергоресурсы
5. Состояние в момент времени t_i , неразрывно связано с состоянием в момент t_{i-1}
6. Стоимость затрат на энергоресурсы при принятии управленческих решений на каждом шаге накапливается с течением времени, в результате чего формируется общая стоимость затрат на заданном интервале

Постановка задачи динамического управления на конечном интервале времени

Общая запись процесса динамического управления процессом проветривания:

$$x_{k+1} = f_k(x_k, u_k, w_k)$$

k – дискретные моменты времени, $k=0, 1 \dots N-1$;

x_k – состояние процесса проветривания, которое сформировалось за счет всех предыдущих действий, оно может быть использовано для будущей оптимизации;

u_k – управляющее воздействие на ГВУ для изменения объемного расхода воздуха относительно расчетного (*baseline*) уровня, которое будет выбрано во время k . Управление имеет ограничения (*control constraint*), вызванное требованиями безопасности, например $u_{min} > u_k > u_{max}$, то есть нельзя задавать значение управления подачей воздуха ниже уровня u_{min} и выше u_{max} ;

w_k – помеха в виде случайного изменения требуемого объема воздуха в зависимости от действия внешних факторов, независимая случайная величина (*independent random parameter*);

N – горизонт оптимизации в виде числа шагов для управления процессом проветривания;

f_k – функция, которая описывает динамику процесса проветривания, механизм для обновления состояния.

Постановка задачи динамического управления на конечном интервале времени

Функция стоимости затрат на энергоресурсы аккумулирует во времени (*additive cost*) все стоимости по шагам.

Если для момента времени k обозначим ее как $g_k(x_k, u_k, w_k)$, то суммарная стоимость будет равна:

$$J(x_0) = g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k)$$

где $g_N(x_N)$ – стоимость на последнем шаге оптимизации, то есть в момент времени, когда процесс управления переходит в свое окончное заранее определенное (детерминированное) состояние.

$$J(x_0) = E \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\} \quad \text{expected cost}$$

Постановка задачи динамического управления на конечном интервале времени

Оптимальное решение задачи управления проветриванием в конкретный день менее важно, чем поиск «оптимального» правила выбора для произвольного момента времени k нужного управляющего воздействия u_k для любых возможных допустимых состояний процесса x_k .

В этом заключается противопоставление «действия» сейчас и «стратегии действий» на всем интервале оптимизации

Постановка задачи динамического управления на конечном интервале времени

Для решения поставленной задачи более рационально искать последовательность функций μ_k $k = 0, \dots, N-1$, ставивших в соответствие состоянию процесса проветривания x_k управляющее воздействие u_k с целью минимизации накапливаемой ожидаемой стоимости расходов на энергоресурсы в условиях ограничений по безопасности.

Функция $\mu_k(x_k)$ в момент времени k возвращает значение управляющего воздействия u_k для состояния процесса проветривания x_k .

Последовательность $\pi = \{\mu_0, \dots, \mu_{N-1}\}$

назовем как «стратегия» или закон управления (*policies*), как набор правил выбора управляющих воздействий u_k для каждого момента времени k и каждого допустимого состояния процесса проветривания x_k

Постановка задачи динамического управления на конечном интервале времени

при выбранной стратегии задача динамической оптимизации процесса проветривания будет заключаться в поиске минимальной ожидаемой стоимости затрат на энергоресурсы

$$J_{\pi}(x_0) = E \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\} \rightarrow \min$$

При оптимизации процесса проветривания на интервале длиной N шагов, только на последнем шаге перед переходом процесса проветривания в окончательное состояние можно принять решение о выборе наилучшего управляющего воздействия, не принимая во внимание последующие моменты времени

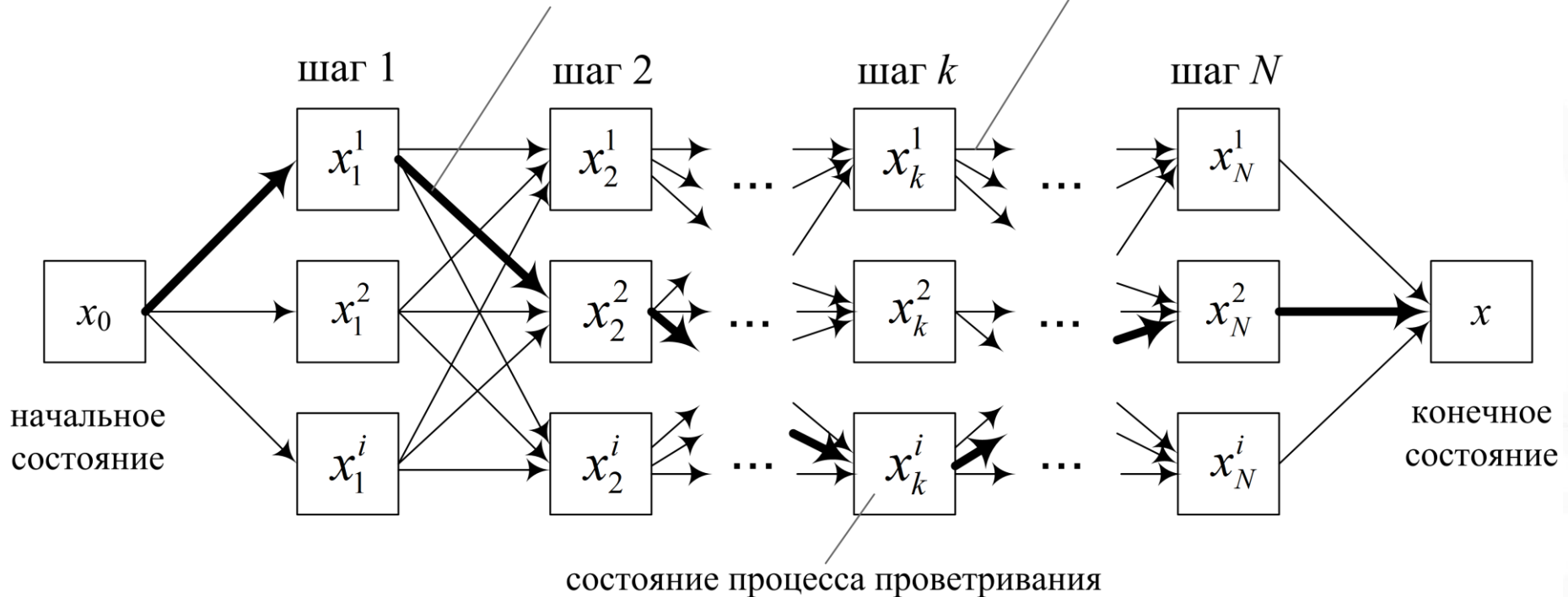
Разработка алгоритма динамического управления процессом проветривания

1. Вводится размер диапазона построения оптимального плана, например сутки вперед.
2. Диапазон разбивается на N этапов (шагов), например $N=96$ для 15-минутных шагов за сутки. Определяется исходное состояние процесса проветривания x_0 и стратегия управления $\pi = \{\mu_0, \dots, \mu_{N-1}\}$, которая будет задавать функцию f , $x_{k+1} = f_k(x_k, u_k, w_k)$, определяющую новое состояние процесса проветривания x_{k+1} , в которое оно может перейти под воздействием управляющего частного решения по изменению подачи воздуха u_k и возмущения w_k , вызванного погодными условиями, поведением людей или машин, другими факторами.
3. Для каждого состояния процесса проветривания на всех N этапах должна быть рассчитана стоимость затрат на энергоресурсы $g_k(x_k, u_k, w_k)$, которые будут иметь место, если для изменения подачи воздуха принять управление u_k , когда процесс проветривания был в состоянии x_k при случайном возмущении w_k .
4. Для каждого из альтернативных состояний процесса проветривания на последнем этапе N решается уравнение $J_\pi(x_N) = E\{g_N(x_N)\} \rightarrow \min$.
5. Следуя из конца в начало рекурсивно для каждого шага $\{N-1, N-2, \dots, 3, 2, 1\}$ для каждого из возможных на данном этапе состояний процесса проветривания x_k решается уравнение $J_\pi(x_0) = E\left\{g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k)\right\} \rightarrow \min$, из которого определяются частные оптимальные управляющие воздействия u_k при случайном возмущении w_k .
6. Для исходного состояния процесса проветривания x_0 и выбранной стратегии управления $\pi_1 = \{\mu_0^1, \dots, \mu_{N-1}^1\}$ определяется последовательность (траектория) промежуточных частных оптимальных состояний и формирующих их частных управляющих воздействий: $x_0 \xrightarrow{u_0} x_1 \xrightarrow{u_1} x_2 \xrightarrow{u_2} \dots x_{N-1} \xrightarrow{u_{N-1}} x_N$.
7. Далее выбирается следующая стратегия $\pi_2 = \{\mu_0^2, \dots, \mu_{N-1}^2\}$ и снова определяется последовательность (траектория) промежуточных частных оптимальных состояний и формирующих их частных управляющих воздействий согласно п.6.
8. Решением задачи оптимального управления процессом проветривания подземных горнодобывающих предприятий на основе метода динамического программирования будет являться та стратегия, которая обеспечит минимальное значение затрат на энергоресурсы $J^*(x) = \min_{\pi \in \Pi} J_\pi(x)$.
9. Зная оптимальную стратегию управления, формируется план в виде вектора оптимальных управляющих воздействий $U = \{u_0, u_1, u_2 \dots u_{N-1}\}$, переводящий процесс проветривания из исходного состояния x_0 в конечное x_N с минимальными затратами на энергоресурсы $J_\pi(x_0)$.

Постановка задачи динамического управления процессом проветривания как поиск оптимальной траектории

оптимальное управление на шаге

выбранное управление u_k и стоимость энергозатрат $g_k(x_k, u_k)$



процесс управления - детерминированный
возмущение w_k принимает одно конкретное значение



Постановка задачи динамического управления процессом проветривания как поиск оптимальной траектории

Задача оптимального управления заключается в поиске оптимальной траектории (последовательности) управляющих воздействий от начального до конечного состояния процесса проветривания на рассматриваемом интервале времени, имеющей наименьшую суммарную стоимость энергозатрат при всех заданных ограничениях по безопасности и технологии работы предприятия

$$J_{\pi}(x_0) = E \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k) \right\} \rightarrow \min$$

$$g_{22}^{*1,j}(x_{22}^1, u_{22}^{*1,j}) = \min \begin{bmatrix} g_{22}^{1,1}(x_{22}^1, u_{22}^{1,1}) + g_{23}^{*1}(x_{23}^{*1}, u_{23}^{*1}) \\ g_{22}^{1,2}(x_{22}^1, u_{22}^{1,2}) + g_{23}^{*2}(x_{23}^{*2}, u_{23}^{*2}) \\ \dots \\ g_{22}^{1,i}(x_{22}^1, u_{22}^{1,i}) + g_{23}^{*i}(x_{23}^{*i}, u_{23}^{*i}) \end{bmatrix}$$

$$g_{21}^{*1,j}(x_{21}^1, u_{21}^{*1,j}) = \min \begin{bmatrix} g_{21}^{1,1}(x_{21}^1, u_{21}^{1,1}) + g_{22}^{*1}(x_{22}^{*1}, u_{22}^{*1}) \\ g_{21}^{1,2}(x_{21}^1, u_{21}^{1,2}) + g_{22}^{*2}(x_{22}^{*2}, u_{22}^{*2}) \\ \dots \\ g_{21}^{1,i}(x_{21}^1, u_{21}^{1,i}) + g_{22}^{*i}(x_{22}^{*i}, u_{22}^{*i}) \end{bmatrix}$$

$$g_1^{*1,j}(x_1^1, u_1^{*1,j}) = \min \begin{bmatrix} g_1^{1,1}(x_1^1, u_1^{1,1}) + g_2^{*1}(x_2^{*1}, u_2^{*1}) \\ g_1^{1,2}(x_1^1, u_1^{1,2}) + g_2^{*2}(x_2^{*2}, u_2^{*2}) \\ \dots \\ g_1^{1,i}(x_1^1, u_1^{1,i}) + g_2^{*i}(x_2^{*i}, u_2^{*i}) \end{bmatrix}$$

$$g_0^{*j}(x_0, u_0^{*j}) = \min \begin{bmatrix} g_0^1(x_0, u_0^1) + g_1^{*1}(x_1^{*1}, u_1^{*1}) \\ g_0^2(x_0, u_0^2) + g_1^{*2}(x_1^{*2}, u_1^{*2}) \\ \dots \\ g_0^i(x_0, u_0^i) + g_1^{*i}(x_1^{*i}, u_1^{*i}) \end{bmatrix}$$

Стратегия ограниченных перспектив

«проклятие размерности» Р. Беллмана

Для снижения вычислительной нагрузки алгоритма предлагается сократить горизонт оптимизации и использовать в каждый момент времени (на каждом шаге) решения, основанные на краткосрочном прогнозировании изменений на некоторое число шагов вперед. В самом простом случае можно использовать стратегию с прогнозом на один шаг вперед.

$$\min_{u_k \in U_k(x_k)} E \left\{ g_k(x_k, u_k, w_k) + \bar{J}_{k+1}(f_k(x_k, u_k, w_k)) \right\}$$

где \bar{J}_{k+1} - аппроксимированное значение истинной функции затрат на энергоресурсы J_{k+1} ,
при $\bar{J}_N = g_N$

Переход к алгоритму субоптимального управления за счет применения инструмента аппроксимированного динамического программирования (ADP)

Далее рассмотрим стратегию с прогнозом на два шага вперед

$$\bar{J}_{k+1}(x_{k+1}) = \min_{u_{k+1} \in U_{k+1}(x_{k+1})} E \left\{ g_{k+1}(x_{k+1}, u_{k+1}, w_{k+1}) + \bar{J}_{k+2}(f_{k+1}(x_{k+1}, u_{k+1}, w_{k+1})) \right\}$$

где \bar{J}_{k+2} - аппроксимированное значение функции затрат на энергоресурсы J_{k+2} . Аналогичным образом определяются стратегии с прогнозом на большее число шагов, что позволяет применять данный подход к решению оптимизационной задачи на бесконечном интервале времени

В случае использования стратегии с прогнозом на два шага вперед число задач поиска минимума затрат равно одному плюс число всех возможных состояний процесса проветривания x_{k+1} , получаемых из состояния x_k . Рост вычислительной сложности можно описать линейной функцией от значения числа шагов периода оптимизации.

Переход к алгоритму субоптимального управления за счет применения инструмента аппроксимированного динамического программирования (ADP)

В рамках стратегии ограниченных перспектив существует несколько подходов к аппроксимации функции затрат:

1. Аппроксимация на основе решения аналогичной задачи предметной области. В этом случае оптимальная стоимость затрат на процесс проветривания аппроксимируется связанной с этой, но более простой и меньшей по масштабу функцией.
2. Эвристическая аппроксимация функции стоимости. Оптимальная стоимость аппроксимируется некоторой функцией, параметры которой выбираются эвристически
3. Оценка процесса последовательного внедрения (*Rollout*) на основе имитационного моделирования. Оптимальная стоимость аппроксимируется на основе некоторой субоптимальной стратегии, которая показывает наилучшие результаты «раскатывающегося» использования управляющих воздействий при имитационном моделировании
4. Комбинированный подход

Аспекты реализации вычислительных алгоритмов последовательного внедрения на основе имитационного моделирования

Для расчета функции выбора управляющего воздействия $\mu_k(x_k)$ при работе *Rollout* алгоритма потребуется для каждого $u_k \in U(x_k)$ в момент времени k рассчитать значение Q -фактора:

$$Q_k(x_k, u_k) = E\{g_k(x_k, u_k, w_k) + H_{k+1}(f_k(x_k, u_k, w_k))\}$$

В этом случае будем понимать под $H_{k+1}(f_k(x_k, u_k, w_k))$ такую функцию затрат на энергоресурсы, которая достигается за счет использования некоторой базовой стратегии управления процессом проветривания $f_k(x_k, u_k, w_k)$, относительно которой будет производиться описание качественного улучшения (результатов оптимизации) с помощью оценки Q -фактора

Аспекты реализации вычислительных алгоритмов последовательного внедрения на основе имитационного моделирования

$$\begin{aligned} \bar{J}_k(x_k) &= E\left\{g_k(x_k, \bar{\mu}(x_k), w_k) + \bar{J}_{k+1}(f_k(x_k, \bar{\mu}(x_k), w_k))\right\} \leq \\ &\leq E\left\{g_k(x_k, \bar{\mu}(x_k), w_k) + H_{k+1}(f_k(x_k, \bar{\mu}(x_k), w_k))\right\} \leq \\ &\leq E\left\{g_k(x_k, \mu(x_k), w_k) + \bar{H}_{k+1}(f_k(x_k, \mu(x_k), w_k))\right\} = H_k(x_k), \end{aligned}$$

Набор приведенных неравенств показывает, что наличие потенциала энергоэффективности при аппроксимации функции затрат с помощью *Rollout* алгоритма, означает эффективность новой стратегии $\bar{\pi}$ по отношению к стратегии π , формируемой на основе базового алгоритма выбора управляющих воздействий.

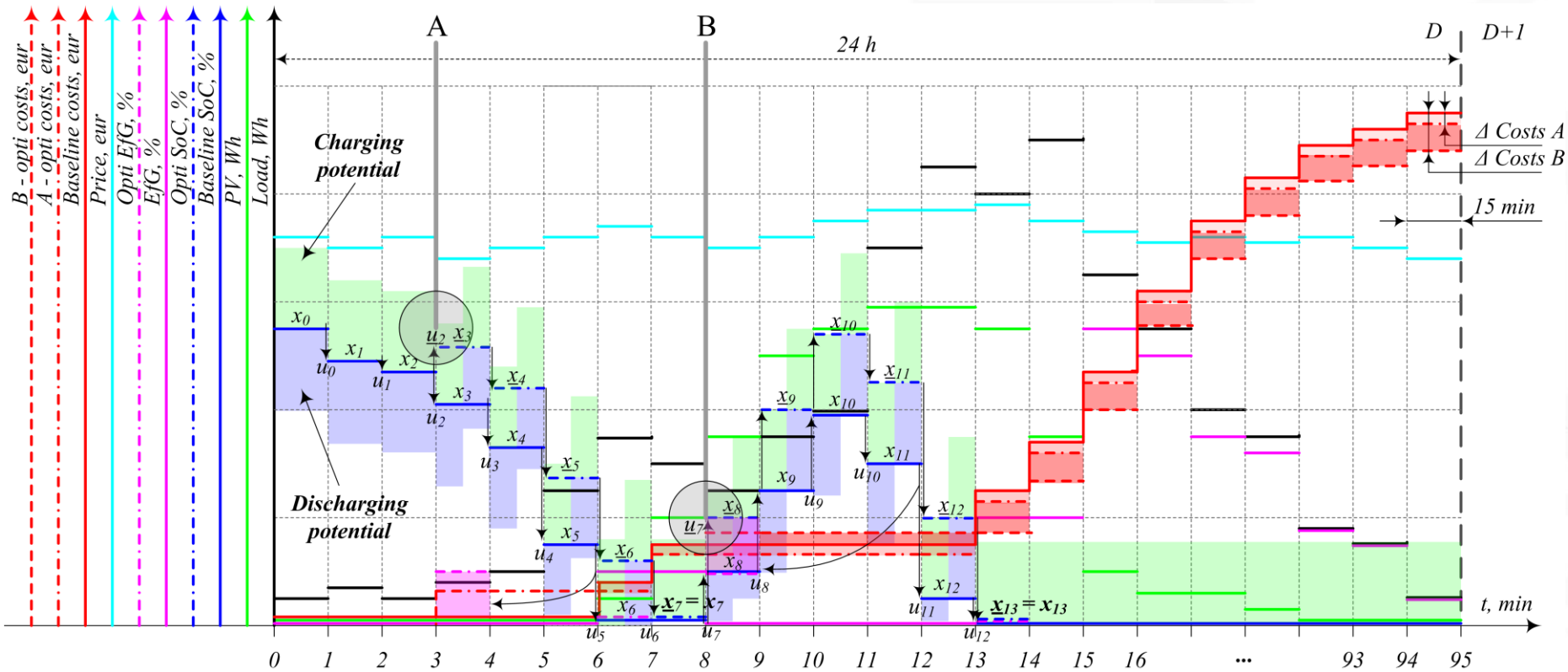
Аспекты реализации вычислительных алгоритмов последовательного внедрения на основе имитационного моделирования

Определить Q -фактор можно за счет решения двух типов задач:

1. Стохастическая задача с оценкой Q -фактора на основе имитационного моделирования с использованием метода Монте-Карло. Данный стохастический подход был представлен в работе [Abramson B. 1990, «Expected outcome: a general model of static evaluation». IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol. 12, pp. 182-193]. В этом случае для состояния процесса проветривания x_k , в результате имитационного моделирования должны быть получены все возможные значения Q -фактора $Q_k(x_k, u_k)$ для всех $u_k \in U(x_k)$.

2. Стохастическая задача с аппроксимацией Q -фактора в соответствии с алгоритмом оптимизации. В этом случае значение Q -фактора может быть получено путем решения задачи оптимизации, например методом наименьших квадратов, для некоторой аппроксимированной функции с заранее определенными параметрами.

Заключение





Спасибо за внимание!