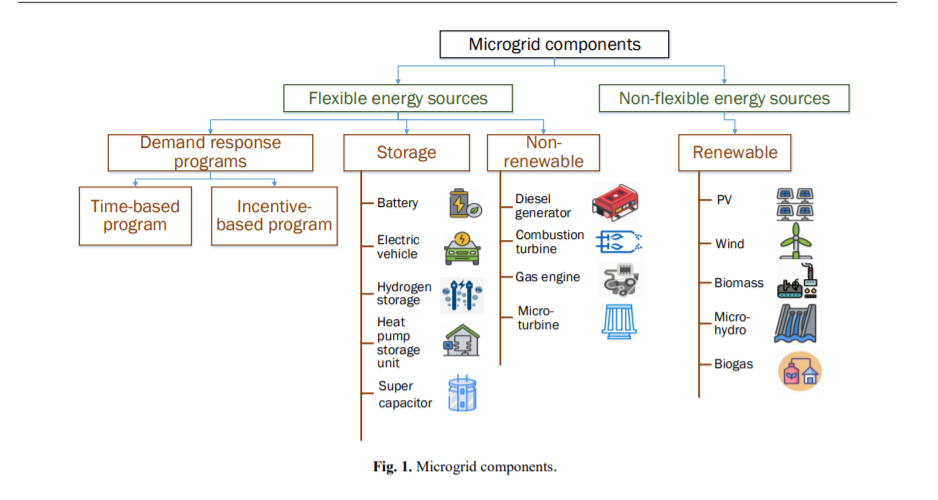
Отличие краткосрочного хранения от долгосрочного — это в основном вопрос стоимости: батареи в настоящее время слишком дороги, чтобы использовать их для решения сезонных колебаний. Энергетические микросети сталкиваются с двойной стохастико-детерминированной структурой: одной из основных задач при эксплуатации МГС является поиск стратегий хранения, способных справиться с неопределенностями, связанными с будущим производством и потреблением электроэнергии; кроме того, МГС также имеют характеристики, на которые их динамика детерминированно реагирует на действия по управлению хранением (Francois-Lavet et al., 2016).

Эта статья организована следующим образом: в разделе 2 подробно рассматривается проблема управления энергопотреблением с учетом гибких вариантов источников. Раздел 3 описывает глубокую структуру обучения с поедерплением. Раздел 4 представляет нашу глубокую структуру обучения с подкреплением, посвященную управлению микросети, а также эмпирические результаты, соответствующие случаю внесетевых микросетей, расположенных в Бельгии и России. Раздел 5 завершает.

# 2. Управление энергией микросетей

2.1 Концепции микросети

Распределенные поколения исключают потребность для системы передачи путем установления около клиентов (Aboli et al., 2019). Интеграция и управление распределенными генераторами вместе с запоминающими устройствами и гибкими нагрузками может составлять низковольтную распределительную сеть, называемую микросеть, которая может работать в изолированном или сетевом режиме (Sedig hizadeh et al., 2019). Родовое понятие микросети показано на рисунке 1. 

возобновляемые

Не возобновляемые

Программы, основанные на времени

Стимулирующие программы

Программы реагирования на спрос

Негибкие энергетические ресурсы

Гибкие энергетические ресурсы

Компоненты микросети

В режиме grid-connected вспомогательные услуги могут предоставляться посредством торговой деятельности между микросети и основной сеткой(Table 1, 2).

Таблица 1 Механизмы создания сетки-привязки микросети

|  |  |  |
| --- | --- | --- |
| **Функции** | **Описание** | **Аккумуляторная система** |
| Энергетический рынок | Решите, какую цену вы готовы заплатить / продать | ++ |
| Дополнительные услуги | Продажа услуг в сеть | ++ |
| Пиковое уменьшение | Через локальную и общинную оптимизацию | ++ |
| Функциональность ИБП | Работайте в островном режиме | ++ |
| Эффективность | Через оптимизированное управление нагрузки и поколения |  |

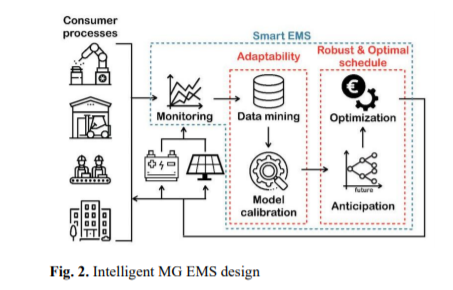
Таблица 2 Преимущества для публичной сети

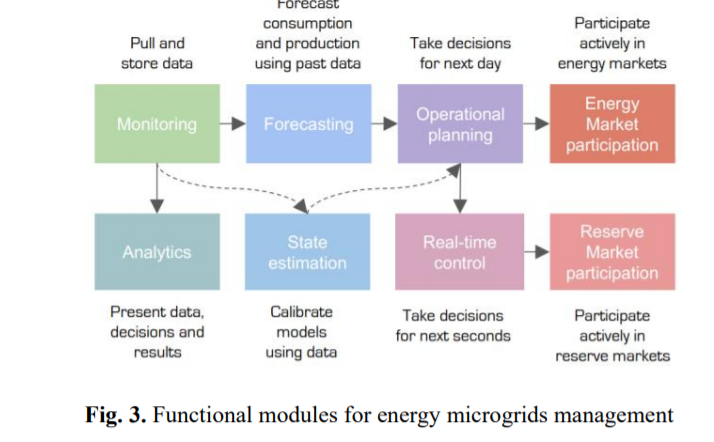
|  |  |  |
| --- | --- | --- |
| **Функции** | **Описание** | **Аккумуляторная система** |
| Пиковое управление уменьшения / подачи | Мгновенно установит ограничения на микросеть | ++ |
| Поддержка напряжения | Реактивная гибкость мощность аккумуляторной батареи и фотоэлектрической | ++ |
| Балансировка фаз | Использование буфера постоянного тока для хранения | ++ |
| Коррекция коэффициента мощности | Гибкость инверторов | ++ |
| Поддержка частоты | Первичный или вторичный резерв | ++ |

Существуют и другие возможные источники дохода (Karavas et al., 2015). В островном режиме реальная и реактивная мощность, генерируемая в микросеть, в том числе обеспечиваемая системой аккумулирования энергии, должна быть сбалансирована с потребностью локальных нагрузок. Микросети предлагают возможность сбалансировать необходимость сокращения выбросов углерода, продолжая обеспечивать надежную электроэнергию в периоды, когда возобновляемые источники энергии недоступны. Микросеть может переходить между этими двумя режимами из-за планового технического обслуживания, деградированной мощности, изменения потока энергии через её компоненты, микросети облегчают интеграцию производства возобновляемых источников энергии, таких как генерация PV, ветра и топливных элементов, не требуя перепроектирования национальной системы распределения (Sfikas et al., 2015; Logenthiran et al., 2011). Современные методы оптимизации также могут быть включены в MG EMS для повышения эффективности, экономичности и отказоустойчивости.

2.2 Интеллектуальный дизайн EMS

Целью EMS является определение оптимального использования распределенных генераторов для питания электрических нагрузок (Li et al., 2019). EMS может работать в двух режимах: централизованном и децентрализованном. В централизованном режиме центральный контроллер направлен на оптимизацию обмена мощностью микросетей на основе рыночных цен и ограничений безопасности. В децентрализованном режиме микросети и управляемые нагрузки имеют большую степень свободы (Dou et al., 2015). В результате компоненты микросети считаются интеллектуальными и пытаются максимизировать доход от взаимодействия микросетей друг с другом (Katiraei et al., 2008). Первоначальная обязанность EMS как в централизованном, так и в децентрализованном режиме заключается в обеспечении микросетей созданием баланса генерации нагрузки (Theo et al., 2017). Стандартная EMS обычно может обеспечить минимальную функцию управления, прежде всего, мониторинг энергии и фиксированные правила работы хранилища. Интеллектуальный MG EMS может использовать данные, чтобы сделать микросеть гибкой, надежной и извлечь максимальную ценность и имеет функцию управления сообществом (Fig 2). Функциональные модули, использующие данные из микросети, представлены на фиг. 3. Рассмотрим некоторые блоки более подробно.





В последнее время появились предложения по разработке онлайн-алгоритмов управления энергией в реальном времени в МГС для оптимизации долгосрочных затрат, которые учитывают неопределенность возобновляемых источников энергии, потребности и рынок. В следующем режиме реального времени действия могут быть обеспечены:

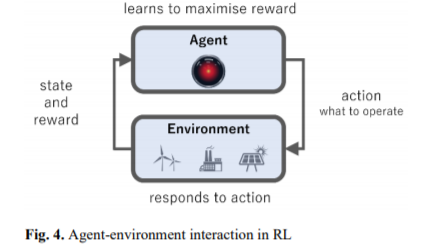
* Режим привязки к сетке: реализует решения по оперативному планированию, исправляет ошибки и осуществляет диспетчеризацию между источниками гибкости и управляет системами хранения для ограничения их деградации
* Островной режим: контролирует и отправляет источники гибкости для поддержания частоты системы и отправки гибридных систем хранения энергии

Участие на рынке передовых энергетических / вспомогательных услуг для MG может включать:

* Оптимальные торги на рынке на день вперед с использованием ожидаемой нагрузки, генерации и цен.
* Корректировка энергетических обменов на внутридневном рынке в соответствии с изменениями нагрузки, выработки и цен
* Использование возможностей балансировки путем реагирования на сигналы TSO (операторов систем передачи).
* Предоставление вознагражденных полей гибкости, которые TSO может активировать для целей балансировки.

# 3. Глубокое машинное обучение с подкреплением в управлении энергетической микросетью

Обучение с подкреплением (RL) дает машине возможность научиться принимать меры (Sutton et al., 2018). Машина выполняет действия в среде для оптимизации сигнала вознаграждения. В контексте микросетей этот сигнал вознаграждения может быть стоимостью энергии, пиком нагрузки или безопасностью - любым поведением, которое мы хотим стимулировать (рис. 4).



В контексте RL агент учится действовать, используя формализм процесса принятия решений Маркова (MDP). Однако пространство состояний очень велико в современных электрических сетях, и поэтому типичный алгоритм RL не имеет эффективного решения. Для решения такой задачи мы можем использовать глубокую нейронную сеть (NN) для моделирования нужных политик, функций значений, которая поэтому называется Deep RL (François-Lavet et al., 2018).

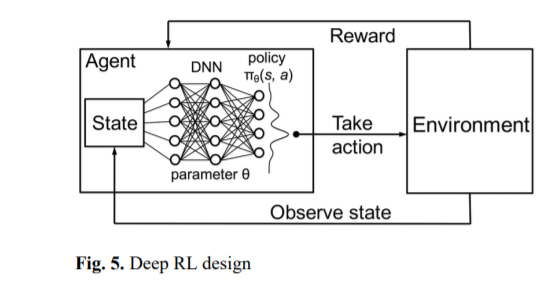
3.1 Глубокие решения RL для последовательного принятия решений

Оптимально работающий микросеть может быть формализован как частично наблюдаемый MDP, где MG рассматривается как агент, взаимодействующий со своей средой. Для того, чтобы приблизиться к Маркова собственность, состояние системы 𝑠𝑡 ∈ 𝑆 состоит из истории особенности наблюдений 𝑂𝑡 𝑖 , 𝑖 ∈ {1, ... , 𝑁𝑓}, где 𝑁 𝑓 ∈ 𝑁 - общее число функций. Каждый 𝑂𝑡 𝑖 представляет собой последовательность пунктуальный наблюдений за выбранной истории длины ℎ𝑖 : 𝑂𝑡 𝑖 = [𝑜𝑡-ℎ𝑖+1 𝑖 , ... , 𝑜𝑡 𝑖 ] (история продолжительность может зависеть от характеристика). На каждом временном шаге, агент отмечает переменной состояния st , принимает меры 𝑎𝑡 ∈ 𝐴 и переходит в состояние 𝑠𝑡+1~𝑃(∙ |𝑠𝑡 , 𝑎𝑡 ). Награда сигнала 𝑟𝑡 = 𝜌(𝑠𝑡 , 𝑎𝑡 , 𝑠𝑡+1) связан с переходом (𝑠𝑡 , 𝑎𝑡

, 𝑠𝑡+1), где 𝜌: 𝑆 × 𝐴 × 𝑆 → 𝑅 награда функцию. Затем мы определяем гамма-дисконтированную оптимальную функцию Q-значения:

* 𝑄 ∗ (𝑠, 𝑎) = max 𝜋 𝐸 [∑ 𝛾 ∞ 𝑘−𝑇 𝑘=𝑡 𝑟𝑘 |𝑠𝑡 = 𝑠, 𝑎𝑡 = 𝑎, 𝜋]

По аналогии (Франсуа-лаве и др., 2016), мы предлагаем аппроксимировать 𝑄 ∗ ( 𝑠 ,𝑎) с помощью глубокого NN. Обозначим через 𝑄 ( ∙ ,∙; Θ 𝑘) так называемую Q-сеть (рис.5).



Глубокие NNs предлагают свойства обобщения, которые адаптированы к высокомерным сенсорным входам, таким как временные ряды. В НН параметров Θ𝑘 могут быть обновлены с помощью стохастического градиентного спуска путем отбора проб партий переходы (𝑠, 𝑎, 𝑟, 𝑠 ' ) в режиме воспроизведения памяти, обновления текущее значение 𝑄(𝑠, 𝑎; Θ𝑘) к целевому значению 𝑌𝑘 𝑄 = 𝑟 + 𝛾 𝑎𝑟𝑔 Макс 𝑎'∈𝐴 𝑄(𝑠 ', 𝑎 ' ,Θ𝑘 -) где Θ𝑘 - относится к параметру от некоторых предыдущих Q-сети под названием цель вопрос-сети или глубокий вопрос-сети (DQN), представленных здесь. При использовании в квадрате-убыток, полученный в обновление обучения выглядит следующим образом:

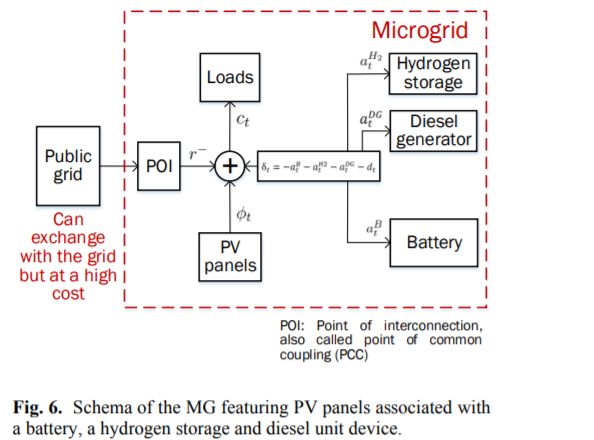
* Θ𝑘+1 = Θ𝑘 + 𝛼(𝑌𝑘 𝑄 − 𝑄(𝑠, 𝑎; Θ𝑘 ))∇Θ𝑘 𝑄(𝑠, 𝑎; Θ𝑘 ) (2)

где 𝛼 - скалярный размер шага, называемый скоростью обучения.

3.2 Микросеть: ориентир описание

Во-первых, обратите внимание, что модель микросети, описанная ниже, полностью описана в (Gemine et al., 2016). Мы обозначаем состояние операции хранения микрорешетки 𝑠𝑡 𝑀𝐺 ∈ 𝑆𝑀𝐺: оно описывает количество энергии в устройстве хранения. Количество энергии в батарее обозначается 𝑠𝑡^𝐵 [𝑊ℎ] ∈ 𝑆^𝐵 [𝑊ℎ], количество энергии в водородном баке обозначается 𝑠𝑡 𝐻2 [𝑊ℎ] ∈ 𝑆𝐻2[𝑊ℎ] и количество энергии, плотности дизельного генератора обозначается 𝑠𝑡 𝐷𝐺 [𝑊ℎ] ∈ 𝑆𝐷𝐺[𝑊ℎ/𝑘𝑔]. Введем x𝐵 [𝑊 ℎ] (ОТВ. x𝐻2 [ 𝑊 𝑝 ]) в качестве аккумулятора (ОТВ. водород) хранение калибровки и мощность дизельного генератора x𝐷 𝐺 [𝑊]. Переменная η𝐵 (соотв. 𝜁𝐵) указывает на разряд батареи (респ. зарядка) эффективность. Аналогичным образом, эффективность электролиза/топливных элементов соответственно обозначается η𝐻2 (при хранении энергии) и 𝜁 𝐻 2 (при доставке энергии). Переменная 𝜁𝐷𝐺 обозначает эффективность дизельный генератор.

На каждом временном шаге, действие a𝑡 = [𝑎𝑡 𝐻2, 𝑎𝑡 𝐷𝐺, 𝑎𝑡 𝐵] ∈ 𝐴𝑡 , применяется в системе, где 𝑎 𝑡 𝐻 2-количество энергии, переданной в (если положительная) или из (если отрицательная) накопителя водорода, аналогично𝑎 𝑡 𝐵 - количество энергии, переданной в или из батареи и 𝑎 𝑡 𝐷 𝐺 - количество энергии из (все отрицательные) дизельного генератора. Динамика батареи дается: 𝑠𝑡+1 𝐵 = 𝑠𝑡 𝐵 + 𝜂𝑡 𝐵𝑎𝑡 𝐵 Если 𝑎𝑡 𝐵 ≥ 0 и 𝑠𝑡+1 𝐵 = 𝑠𝑡 𝐵 - 𝑎𝑡 𝐵 𝜁𝑡 𝐵 ⁄ в противном случае. Аналогично, водород динамика дается: 𝑠𝑡+1 𝐻2 = 𝑠𝑡 𝐻2 + 𝜂𝑡 𝐻2𝑎𝑡 𝐻2 если 𝑎𝑡 𝐻2 ≥ 0 и 𝑠𝑡+1 𝐻2 = 𝑠𝑡 𝐻2 - 𝑎𝑡 𝐻2 𝜁𝑡 𝐻2 в противном случае. Динамика дизель-генератора задается формулой: 𝑠𝑡+1 𝐷𝐺 = 𝑠𝑡 𝐷𝐺 − 𝑎𝑡 𝐷𝐺 𝜁𝑡 𝐷𝐺 ⁄ для всех случаев. Функция вознаграждения системы соответствует мгновенным операционным доходам 𝑟 𝑡 в момент времени𝑡 ∈ 𝑇. Мы использовали три количества, которые являются предпосылками к определению вознаграждения функция: (1) ϕ𝑡 [𝑊ℎ] ∈ ℝ+ электроэнергия генерируется локально фотоэлектрические установки (2) 𝑑𝑡 [𝑊ℎ] ∈ ℝ обозначает чистый спрос на электроэнергию, который определяется как разница между местным потребление C𝑡 и местное производство электричества φ𝑡 , (3) δ𝑡 [𝑊ℎ] ∈ ℝ [представляет баланс сил в микросети, учитывая contributins чистого спроса на электроэнергию и заряд или разряд устройства хранение: δ𝑡 = -𝑎𝑡 𝐵 - 𝑎𝑡 𝐻2 - 𝑎𝑡 𝐷𝐺 - 𝑑𝑡 (Рис. 6).



Мгновенное вознаграждение сигнал 𝑟𝑡 получается путем сложения доходов, получаемых от производства водорода 𝑟 𝐻2 с пенальти 𝑟 - из-за значение потери нагрузки: 𝑟𝑡 = 𝑟(𝑎𝑡 , 𝑑𝑡 ) = 𝑟𝐻2 + 𝑟-(𝑎𝑡 , 𝑑𝑡 ). Штраф 𝑟 - пропорционален общему количеству энергии, которая не была поставлена для удовлетворения спроса: 𝑟 -(𝑎𝑡 , 𝑑𝑡 ) = 𝑘δ𝑡 , когда δ𝑡 < 0 и null в противном случае (𝑘- это стоимость, переносимая за Wh, не поставляемый в пределах микрорешетки), в то время как 𝑟 𝐻 2 задается: 𝑟 𝐻2(𝑎𝑡 , 𝑑𝑡 ) = 𝑘𝐻2𝑎𝑡 𝐻2 (𝑘 𝐻2 является ли выручка / затраты на единицу произведенного/использованного водорода). В соответствии с постановкой задачи, нет возможности поставлять энергию извне (для общественных сетей) и система не получает за это вознаграждение. Однако гибридная PV-водород-дизельная система может обеспечить гораздо более существенную максимизацию операционных доходов.

Из серии награды 𝑟𝑡 , мы получаем операционные доходы за год y определены следующим образом: 𝑀𝑦 = ∑ 𝑦 𝑟𝑡𝑡∈𝜏 где 𝜏 𝑦 - это набор шагов, принадлежащих к году. Оптимизация работы мг требует определения последовательной стратегии принятия решений, которая приводит к максимизации 𝑀 𝑦 (Francois-Lavet et al., 2016).

## 4 Исследование конкретной ситуации

Предложенный подход эмпирически проиллюстрирован на примере интеллектуального энергоменеджмента внесетевых микросетей, расположенных в Бельгии и России. Наши примеры имитируют работу реалистичного микросеть, которая не подключена к основной инженерной сети (off-grid) и который оснащен фотоэлектрическими панелями, батареями и накопителем водорода и / или дизельным генератором. Мы использовали и модифицировали исходный код Python, предложенный в (Francois-Lavet et al., 2016).

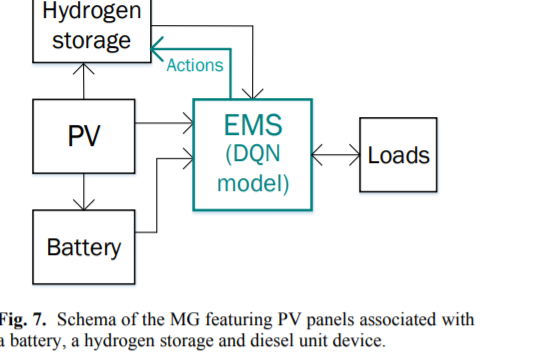
Мы предложили архитектуру DQN, где входы обеспечиваются вектором состояния, и где каждый отдельный выход представляет Q-значения для каждого дискретизированного действия. DQN обрабатывает временные ряды благодаря набору сверток с 16 фильтрами 2×1 с шагом 1, за которыми следует свертка с 16 фильтрами 2×2 с шагом 1. Выход свертки, а также другие входы затем следуют два полностью Соединенных слоя с 50 и 20 нейронами и выходной слой.

Начиная со случайного DQN, мы выполняем на каждом временном шаге обновление, указанное в Eq. 2 и, в то же время, мы пополняем воспроизведения памяти с все замечания, действия и награды, используя агентом, что следует 𝜖-жадные политики.Т. в 𝜋 политики(𝑠) = Макс 𝑎 ∈ 𝑄 𝐴 ( 𝑠 , 𝑎; Θ𝑘) выбирается с вероятностью 1 - 𝜖, и случайных действий (с однородной вероятностью в течение действия) выбирается с вероятностью. Мы используем уменьшающееся значение 𝜖 с течением времени. В ходе проверки и тестирования фазы, 𝜋 политики(𝑠) = Макс 𝑎 ∈ 𝑄 𝐴 ( 𝑠 , 𝑎; Θ𝑘) применяется (с 𝜖 = 0). Состояние агента DQN состоит из истории двух-четырех пунктуальных наблюдений:

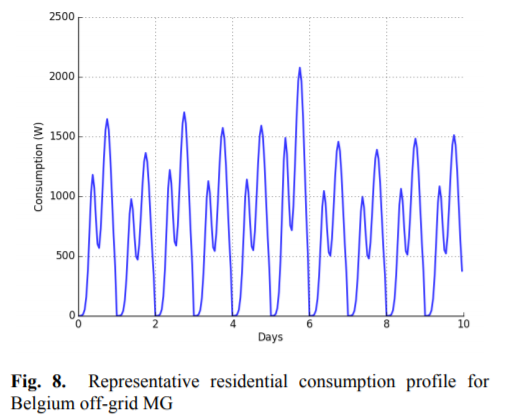
* Состояние зарядки кратковременного накопителя (0 пустое, 1 полное)
* Производство φ𝑡 и потребления с𝑡
* Расстояние до равноденствия
* Прогнозы будущего производства на ближайшие 24 часа и 48 часов

4.1 Случай 1: гибридная система водорода PV

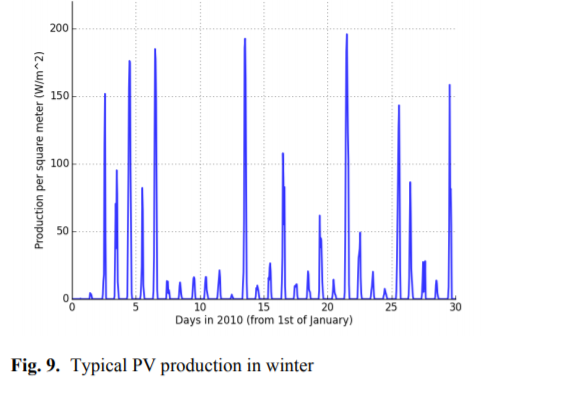
DQN-агент может либо хранить в долговременном хранилище (хранилище водорода), либо извлекать из него энергию, в то время как краткосрочное хранилище (батарея) в лучшем случае справляется с недостатком или избытком энергии, разряжая себя или заряжая себя соответственно (рис. 7). Мы рассматриваем три дискретизированных действия: (1) разрядка на полной скорости хранения водорода, (2) держать его вхолостую или (3) зарядить его на полной скорости. Всякий раз, когда краткосрочное хранилище пусто и не может обработать чистый спрос, штраф (отрицательное вознаграждение) получается равным значению убыточной нагрузки, установленной на 2 евро/кВт \* ч.

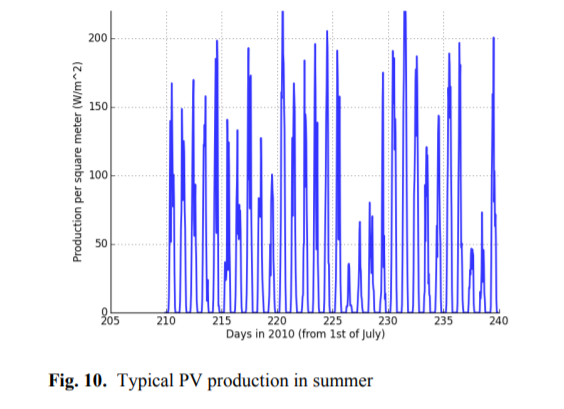


Рассмотрим случай бытового потребителя электроэнергии (в среднем 18 кВт \* ч / сут), расположенного в Бельгии и работающего вне сети MG (рис. 8). Размер батареи х𝐵 = 15 𝑘𝑊ℎ, мгновенная мощность хранения водорода х𝐻2 = 1.1 𝑘𝑊 и пик производства электроэнергии PV х𝑃𝑉 = 12 𝑘𝑊𝑝. Стоимость k за киловатт-час, не поставляемый в рамках МГ, устанавливается в размере 2 евро / киловатт-час. Другие параметры мг взяты из (Gemine et al., 2016).

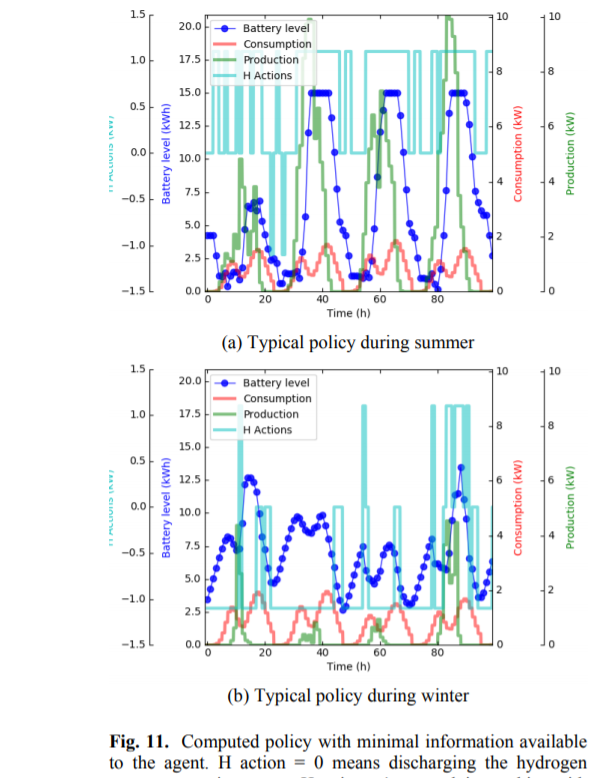


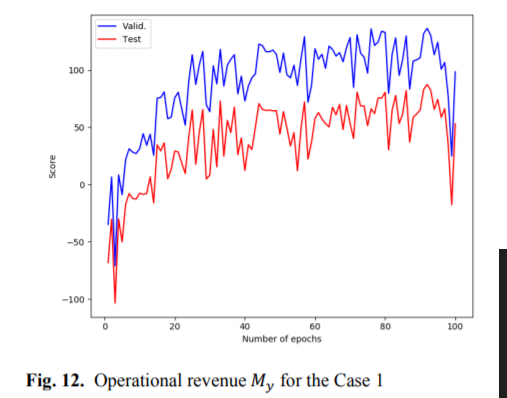
Солнечное излучение изменяется в течение года в зависимости от сезонов, а также изменяется в течение дня в зависимости от погоды и положения Солнца на небе относительно фотоэлектрических панелей. Главным различием между этими профилями является разница между летним и зимним производством фотоэлектрических материалов. В частности, производство варьируется с коэффициентом 1: 5 между зимой и летом, как видно из показателей производства фотоэлектрических панелей для жилого заказчика, расположенного в Бельгии, на рисунках 9 и 10.





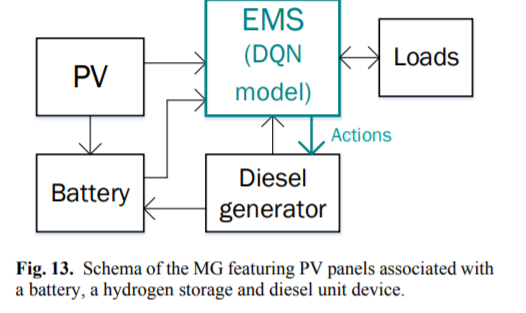
Типичное поведение политики 𝜋 ∗ показано на Рисунке 11 (тестовые данные). Поскольку микрогрид не имеет информации о будущем, он накапливает (в течение ночи) достаточный запас в устройстве кратковременного хранения, чтобы иметь возможность столкнуться с потреблением на следующий день, не испытывая слишком большой потери нагрузки. Он также позволяет избежать потери энергии (когда кратковременное хранилище заполнено) путем хранения в долговременном запоминающем устройстве, когда это возможно.



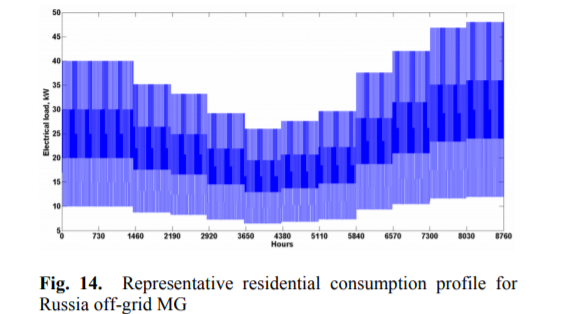
На рис. 12 представлен операционный доход по тестовым и валидационным данным 𝑀 𝑦 для случая 1 в зависимости от процесса обучения DQN. Лучший DQN получен после 92 эпох, с оценкой валидации 136,46 евро / год. Тестовая оценка этой нейронной сети составляет 87,34 евро / год. 

4.2 Случай 2: гибридная система дизеля PV

Преимуществом такой системы является то, что она требует низких инвестиционных затрат. Однако основные недостатки заключаются в том, что для работы генератора необходимо подавать топливо, а излишки энергии в хороший сезон не накапливаются (рис. 13).



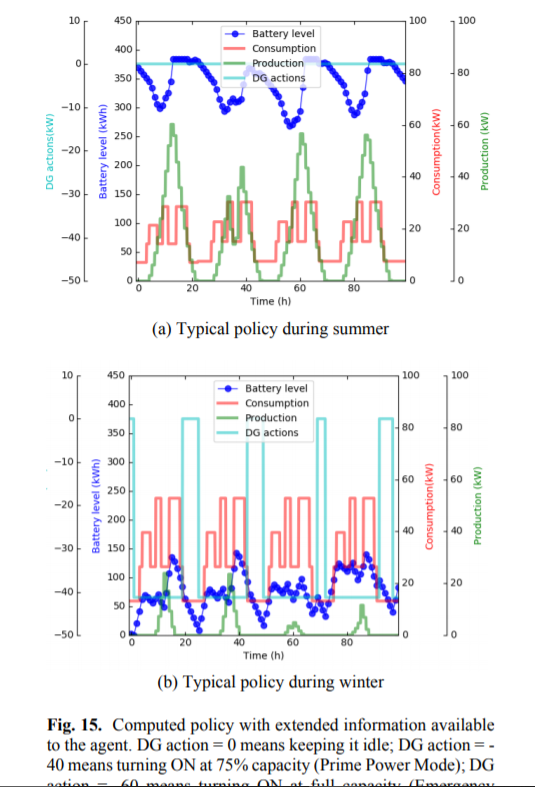
Рассмотрим случай 2 жилого электроснабжения потребитель (в среднем 48 кВт \* ч / сут), расположенный в Якутии, Россия, работающий вне сети микросети. В качестве базовой информации использовались ретроспективные данные (Временной интервал с 2005 по 2019 гг.) по суммарной солнечной радиации. Солнечное излучение было зафиксировано на метеорологической станции, расположенной в этом населенном пункте. Электрическая нагрузка показана на фиг. 14 был построен по реальным данным типичных дней относительно каждого месяца. Размер батареи х𝐵 = 384 𝑘𝑊ℎ, мощность дизельного генератора составляет х𝐷𝐺 = 100 𝑘𝑊 и пик производства электроэнергии PV х𝑃𝑉 = 75 𝑘𝑊𝑝. Стоимость k за киловатт-час, не поставляемый в рамках МГ, устанавливается в размере 2 евро / киловатт-час. Параметры MG взяты из (Sidorov et al., 2019).



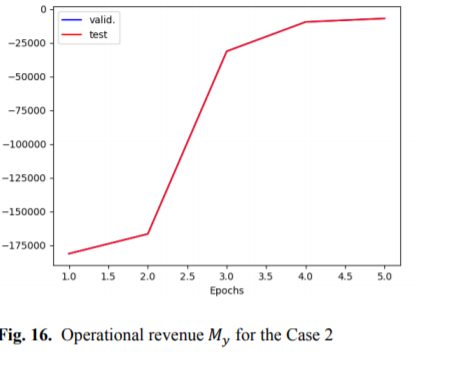
Состояние агента DQN складывается из истории двух-четырех пунктуальных наблюдений, которые такие же, как и в случае 1. Для DQN-агента возможны три действия:

* Действие DG -2 соответствует включению на полную мощность (режим аварийного резервного питания)
* Действие DG -1 соответствует включению на 75% мощности (режим Основной мощности)
* Действие DG 0 соответствует поддержанию его в режиме ожидания.

Типичное поведение политики показано на рисунке 15 (тестовые данные).



Из рисунка 15 видно, что летом дизель-генератор простаивает, а накопитель водорода заряжается; зимой-оба устройства работают для обеспечения баланса мощности для микросети. На рис. 16 показан операционный доход по тестовым и валидационным данным 𝑀 𝑦 для случая 2 в зависимости от процесса обучения DQN. Лучший DQN получен после 5 эпох, с оценкой валидации -100.056 евро / год. Тестовая оценка этой нейронной сети составляет -100.44 евро / год.



Отрицательный операционный доход 𝑀 𝑦 для оптимальной политики 𝜋 ∗ здесь означает, что анализируемый PV дизель MG не имеет возможности накапливать избыточное производство PV в долгосрочном хранилище. При этом основной целью является не максимизация операционных доходов, а фактически минимизация затрат.

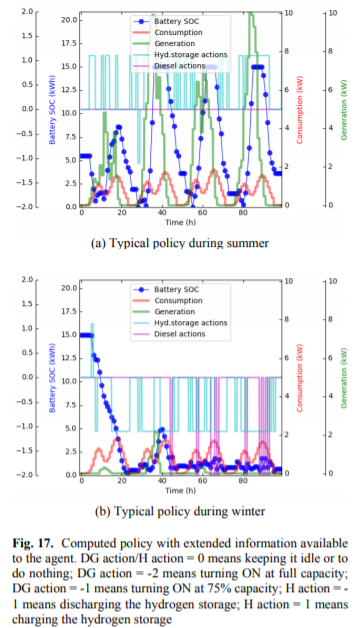
4.3 Случай 3: гибридная система водорода PV Тепловозная

Мы рассмотрели более сложную изолированную гибридную систему, когда DQN-агент может управлять двумя вариантами гибкости-накопителем водорода и дизельным генератором. Мы использовали данные для случая 1 (off-grid Belgium MG). Проектирование такой гибридной системы ставит более сложную задачу оптимизации (Dufo-Lopez et al., 2008). Состояние агента DQN складывается из истории двух-четырех пунктуальных наблюдений, которые такие же, как и в случаях 1 и 2.

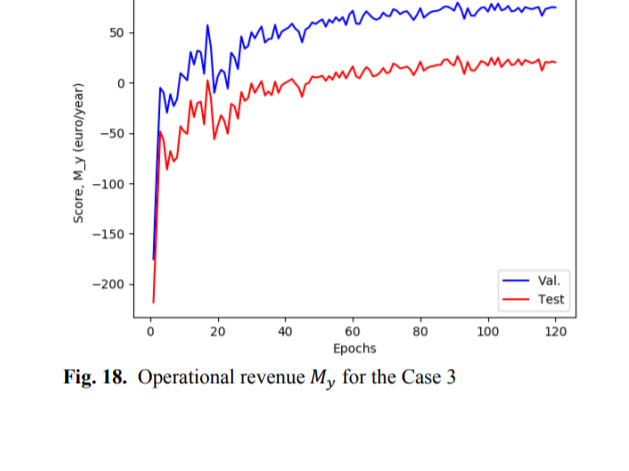
Для DQN-агента возможны шесть действий:

* ДГ действий/ч действие = 0 означает, что он бездействует и ничего не делать ;
* Действие DG = -2 означает включение на полную мощность;
* Действие DG = -2 означает включение на полную мощность;
* H действие = -1 означает сброс накопителя водорода;
* H действие = 1 означает зарядку накопителя водорода

Типичное поведение политики 𝜋 ∗ для случая 3 показано на рис. 17 (тестовые данные). Видно, что летом дизель-генератор простаивает, а накопитель водорода заряжается. Однако во время зимы, оба долгосрочных склада работают для того чтобы помочь панелям ПВ в обеспечивать требование. При этом накопитель водорода включается (разряжается) чаще, чем более дорогой дизель-генератор.



На рисунке 18 мы представляем операционный доход по тестовым и валидационным данным 𝑀 𝑦 для случая 3 в зависимости от процесса обучения DQN. Лучший DQN получен после 91 эпохи, с оценкой валидации 79.59 евро / год. Тестовая оценка этой нейронной сети составляет 26,44 евро / год.



Интересно отметить, что для случая 3, операционные доходы 𝑀𝑦 ниже, чем для случая 1. Это ожидаемый результат, так как добавление дизельного генератора приводит к дополнительным расходам топлива.

# 5 Выводы

Полагаться на традиционные математические модели для управления распределенной гибкостью в реальном времени для МГ не эффективно, так как вычислительные мощности для их обработки были бы чрезвычайно велики. Вероятным решением является использование комбинации двух подходов моделирования (задача оптимизации вариантов гибкости + реалистичные стохастические входы), которые могут быть объединены с обширным моделированием, основанным на глубоком обучении подкреплению.

В этой статье была введена архитектура глубокого обучения с усилением для решения проблемы работы электрического MG в стохастической среде. Экспериментальные результаты иллюстрируют тот факт, что NN-представление функции стоимости эффективно обобщает политику на ситуации, соответствующие невидимым конфигурациям спроса на электроэнергию и солнечной радиации.