Machine Learning-Based Energy Management in a Hybrid Electric Vehicle for better optimization of power

1 Atul Dwivedi 1 dwivedi.20@gmail.com 2 Balendu Bhushan Pandey 2 blissbalendu@gmail.com

Abstract:

This study looks into the issue of energy management in hybrid electric vehicles (HEVs), with a particular emphasis on how to keep a HEV's running costs, which include both gasoline and battery replacement costs, to a minimum. More specifically, the study proposes a nested learning framework in which both the best courses of action (such as choosing between an internal combustion engine and an electric motor to power the car) and the range restrictions imposed by the battery's state-of-charge are dynamically learnt. The inner-loop learning process is essential for reducing fuel consumption, whereas the outer-loop learning process is crucial for reducing the amortised cost of battery replacement. Experimental findings show that the suggested HEV energy can reduce operational costs by up to 48%.

I. Introduction:

Due to worries about excessive fuel consumption and pollution from traditional internal combustion engine (ICE) vehicles, electric vehicles (EVs) and hybrid electric vehicles (HEVs) are currently gaining market share in the automotive industry. Compared to conventional vehicles, EVs and HEVs have improved energy efficiency and fewer emissions thanks to the incorporation of electric motors (EMs) into the propulsion system. HEVs, which serve as a bridge between fully electric vehicles and regular ICE vehicles, are more fuel-efficient than ICE vehicles and have less battery-related issues than EVs. To fully explore the benefits of HEVs, however, complex HEV energy management techniques are required due to the hybrid structure of the propulsion system. An ICE and one or more EMs are the components of a HEV's hybrid propulsion system. The ICE transforms the fuel's chemical energy into mechanical energy to move the vehicle. Regenerative braking is a method that increases the energy efficiency of EVs and HEVs. The EM utilises electrical energy stored in the battery pack to move the vehicle. It can also function as a generator collecting kinetic energy during braking to charge the battery pack. In order to increase the energy efficiency of HEVs, HEV energy management strategies coordinate the functioning of ICE and EM.

One significant element of the HEV's operational costs is the cost of fuel. As a result, the majority of earlier research on HEV energy management focused on increasing fuel efficiency. The driver-controlled pedal action is translated into the necessary propulsion power by rule-based techniques for HEV energy management, which then decide how much power should be distributed between the ICE and the EM using intuition, human judgement, or fuzzy logic. The optimization-based control strategies either reduce the amount of gasoline used during a trip with a specified, anticipated, or stochastic driving profile, or carry out control by converting battery charge into equal fuel consumption (ECMS and adaptive-ECMS approaches).

Due to the frequent charging and discharging of the battery pack by the EM, the state-of-health (SoH) of the HEV battery pack is declining with HEV operation. The paper examined the state-of-charge (SoC) swing, number of charging/discharging cycles, and other factors as they relate to the SoH degradation model for the EV/HEV battery pack. When its SoH deteriorates to 80% or 70%, the battery pack will approach the end of its useful life. Replacing the battery pack will increase the HEV's running costs. To enhance the HEV's fuel efficiency, it is preferable to increase the battery pack's energy capacity within size, weight, and cost limits. In particular, the plug-in HEV (PHEV) uses a higher-capacity battery pack that can store more energy. The amortised battery replacement cost must not be overlooked in the HEV because the cost of replacing a battery rises dramatically when battery capacity is increased. When maximising fuel efficiency, some work has been done to take battery SoH degradation into account. These works, however, suffer from one or more of the following drawbacks:

I The ECMS or adaptive ECMS approaches, which depend on knowledge of the future driving profile, provide the foundation for the HEV energy management rules they employ. The efficacy of these ECMS and adapative-ECMS based techniques may suffer if the future driving profile prediction is inaccurate.

(ii) They employ Ah-throughput or battery output power as the optimization and assessment metrics instead of an exact analytical battery SoH deterioration model.

When it is difficult or perhaps impossible to achieve explicit and precise system modelling, machine learning offers a potent tool for the agent (i.e., decision-maker) to "learn" how to "act" optimally. The agent has the ability to observe the status of the environment and respond accordingly. As a result of the action, the agent will receive a reward. The agent is motivated by the reward and attempts to "learn" from its prior experiences in order to derive a policy, which is a mapping from each potential state to an action. The HEV energy management issue has been tackled using reinforcement learning, therefore the HEV energy management strategy doesn't depend on any information of the future driving profile. The driver behaviour has been learned using an inverted reinforcement learning technique, although that is not our main concern. In the proposed work, we look at the HEV energy management issue with an emphasis on lowering the running costs of a HEV, which include the cost of fuel and amortised battery replacement (i.e., battery purchase plus installation cost). We provide a nested learning system in which both the best course of action (which includes choosing the gear ratio and whether to utilise ICE or EM to drive the car) and the maximum range that a battery SoC can support are dynamically learnt. More specifically, the outer-loop learning process modifies the battery SoH degradation from a HEV's perspective while the inner-loop learning process controls the HEV components' operating modes. The suggested HEV energy management does not rely on flawless and accurate system modelling due to the use of machine learning techniques (i.e., HEV component modelling and driving profile modeling.) The proposed nested learning framework for HEV energy management differs from the reinforcement learning-based framework in that I it incorporates the amortised battery replacement cost; and (ii) it uses two nested learning processes, where the inner-loop learning process is essential for minimising fuel consumption and the outer-loop learning process is essential for minimising the amortised battery replacement cost. The proposed HEV management philosophy reduces operational costs by up to 48%, according to experimental data.

II. System Description:

Although the goal of this work is to create a smart HEV controller that learns from its experience to determine the best energy management strategy, it is still necessary to comprehend the fundamentals of HEV operation. We discuss the parallel HEV configuration, which is typical of the majority of the literature work on HEV energy management, as an example without losing the generality of the discussion. A parallel HEV can operate in one of five different ways depending on the energy flow: One of the following scenarios: the vehicle is propelled solely by the ICE, solely by the EM, simultaneously by the ICE and EM, simultaneously by the ICE and EM while driving the EM to charge the battery pack, and finally, solely by the EM when the vehicle is braking (i.e., regenerative braking mode.)

HEV Component Analysis :

1) Internal Combustion Engine (ICE):

The fuel efficiency of an ICE is determined using the quasi-static ICE model as follows: ICE(TICE,ICE) = TICE • ICE/(m f •Df). (1) In (1), TICE and ICE stand for the ICE's torque (in N•m) and speed (in rad/s), which reflect the ICE's operation point, respectively. Depending on the ICE operation point, mf represents the fuel consumption rate (in g/s) of the ICE. The fuel energy density (measured in J/g) is Df. The fuel consumption rate contour map for a sample ICE in the ICE speed-torque plane is shown in Figure 1(a).

2) Electric Motor (EM):

The EM can be used as a generator to charge the battery pack or as a motor to move the vehicle. The equation for the EM's efficiency is EM(TEM, EM) = (TEM • EM)/Pbatt. (TEM • EM) TEM 0 Pbatt TEM 0 where Pbatt is the output power of the battery pack, EM and TEM are the torque and speed of the EM, respectively. Pbatt > 0 indicates that the battery pack is discharging when the EM is operating as a motor; when the EM is operating as a generator, Pbatt 0 indicates that the battery pack is charging while the TEM is negative. The efficiency contour map of the EM as a motor or generator is shown in Figure 2. The following restrictions should be adhered to for an EM to operate safely and without incident 0 ≤ ωEM ≤ ω max EM , (4) T min EM (ωEM) ≤ TEM ≤ T max EM (ωEM).

3) Vehicle Tractive Force:

When the driver presses the brake or accelerator pedal, the vehicle tractive force FT R is generated to support the vehicle's speed and acceleration. R = a+Fg + FR + FAD FR = m• g • cos CR = m• g • sin FAD = 0.5; ; CD; AF; v where m denotes the vehicle's mass, a denotes its acceleration, Fg denotes the force caused by road slope, FR denotes rolling friction, FAD denotes air drag, denotes the road slope angle, CR denotes rolling friction, denotes air density, CD denotes air drag, AF denotes the vehicle's frontal area, and v denotes its speed. The tractive force FT R can be derived using given v, a, and. Following that, Twh = FT R • rwh, wh = v/rwh, and wheel speed wh are connected to FT R, v, and wheel radius rwh. The required power to move the object, pdem, is satisfied by pdem = FT R • v = Twh •wh.

III. A Nested Learning FrameWork For HEV Energy Management:

The running costs of a HEV, including fuel costs and amortised battery replacement costs, are what we are trying to reduce in this work. In order to accomplish this, we suggest a nested learning framework for HEV energy management, in which the best ways to move the car and the limits on how much the battery's SoC can change are simultaneously learned by inner-loop reinforcement learning and outer-loop adaptive learning. The outer-loop adaptive learning process is essential for minimising the amortised battery replacement cost, while the inner-loop reinforcement learning process is essential for minimising fuel consumption.

A. Motivation :

The following justifies our usage of reinforcement learning in the inner loop. I The inner-loop HEV energy management tries to optimise an expected cumulative return rather than an immediate reward; the reinforcement learning likewise aims to optimise the overall fuel consumption during a driving journey rather than the instantaneous fuel consumption rate at each time step. (ii) Different HEV operation modes are needed during a trip due to variations in the vehicle's speed, power requirements, and battery charge level; the reinforcement learning agent responds differently based on the situation at hand. (iii) The inner-loop HEV energy management only knows the current vehicle speed and power demand values, as well as the current temperature. It has no prior knowledge of the entire driving journey. The present state and reward are all that the reinforcement learning agent needs to learn the best course of action; it is not necessary for it to be aware of previous system inputs or intricate system modelling. However, we also take into account battery SoH degradation in the inner loop by incorporating the battery capacity fading term into the reward of the reinforcement learning, so that the inner loop itself can be used as an independent HEV energy management framework for lowering the overall operating cost. The inner loop is the key to minimising fuel consumption.

The battery pack SoC is clamped by a defined lower bound and higher bound in the earlier work on HEV energy management, which uses a fixed battery SoC range. The resulting HEV energy management tactics may then have a tendency to consume up the battery energy available even for a few very brief urban journeys, which could seriously degrade the battery SoH. Regenerative braking can supply a sizable amount of energy to the battery during urban excursions. Using all of the battery's energy is not always essential.**Inner-Loop Reinforcement Learning Process:**

1) Reinforcement:

Background information on learning: In reinforcement learning, the decision-maker is referred to as the agent, and everything around him or her is referred to as the environment. Each discrete time step in a series, where t = 0, 1, 2,... The agent observes the environment's state st S at each time step t and acts at A on the basis of that observation, where S and A are the sets of potential states and actions, respectively. The agent finds the environment in a new state st+1 and receives a numerical reward rt+1 one time step later, partially as a result of the action made.

A policy of the agent is a mapping from each state s to an action a that identifies the action(s) the agent will select while the environment is in state s. An agent's ultimate objective is to identify the best course of action so that V (s) = E (X k=0 k •rt+k+1 | st = s) is maximised for each state s S. The expected return when the environment begins in state s at time step t and continues to follow policy is represented by the value function V (s). The discount rate, which has the value 0 1, is a parameter that makes sure the infinite sum (Pk=0 k •rt+k+1) converges to a finite value. What's more, reflects the haziness of the future. RT + k+1 is the reward received at time step t +k +1.

2) State Space:

We establish the parameters for the state space of the inner-loop reinforcement learning, where pdem, the amount of power required to move the HEV, v, the speed of the car, and q, the amount of charge in the battery pack, are. Under different conditions, various measures ought to be taken. For instance, the HEV controller should charge the battery by using the EM as a generator if the power demand is negative, meaning the car is braking. On the other side, if the power requirement has a very high positive value, discharging the battery is the appropriate course of action to power the EM, which helps the ICE move the car.

S = s = [pdem, v,q] T |pdem ∈ Pdem, v ∈ V ,q ∈ Q ,

An agent for reinforcement learning ought to be capable of observing a state. The current power demand level pdem and vehicle speed level v can be determined in the real inner-loop reinforcement learning implementation by utilising sensors to measure the driver-controlled pedal motion. Nevertheless, since the battery pack terminal voltage varies with the charging/discharging current and is not a reliable indicator of q, the charge level q cannot be determined from online measurement of terminal voltage. The agent requires the Coulomb counting method, which is commonly implemented using a dedicated circuit, to observe the charge level q.

The finite sets of power demand levels, vehicle speed levels, and battery pack charge levels are designated as Pdem, V, and Q in, respectively. The definition of these finite sets requires discretization. In particular, Q is established by discretizing the [qmin,qmax] range of stored charge into a finite number of charge levels: Q = {q1,q2,...,qN},

where qN = qmax and qmin = q1 q2... qN. In the charge-sustaining energy management for regular HEVs, qmin and qmax are typically 40% and 80% of the battery pack nominal capacity, respectively; 0% and 80%, respectively, in the chargedepleting energy management for PHEVs. Since qmax is typically fixed in the HEV control, we will adjust qmin value to modulate the battery SoH degradation during the outerloop adaptive learning process. Action Area 3 A finite number of actions, each represented by the battery pack's discharge current and the gear ratio value, constitute the action space of inner-loop reinforcement learning, according to our definition:

A = a = [i,R(k)]T |i ∈ I,R(k) ∈ R

where the agent performs action a = [i,R(k)]T, which involves discharging the battery pack with current I and selecting the kth gear ratio. A finite number of current values in the range [Imax,Imax] are contained inside the set I. Please take note that I > 0 indicates that the battery pack is being discharged, while I 0 indicates that it is being charged. The permitted gear ratio values, which vary depending on the drivetrain architecture, are contained in the set R. Typically, there are four or five different gear ratios altogether.

As an alternative, we can construct a condensed action space Are, where action are = I is to discharge the battery pack with current I (and the gear ratio R(k) is chosen by resolving an optimization issue in a way that minimises the fuel consumption rate). The quantity of state-action pairs affects the complexity and rate of convergence of reinforcement learning algorithms. As a result, the smaller action space Are contributes to a four- to five-fold boost in convergence speed and a reduction in complexity. Yet, in order to solve the optimization problem, this condensed action space is dependent on HEV component modelling. For model-free control, we can either utilise the original action space, or we can use the reduced action space Are for reduced complexity and increased convergence.

Reward:

Instead, we can create a condensed action space Are, where action are = I is to drain the battery pack with current I (and the gear ratio R(k) is selected by finding a solution to an optimization problem that minimises the fuel consumption rate). The complexity and rate of convergence of reinforcement learning algorithms depend on the number of state-action pairs. As a result, the smaller action space Are helps to increase convergence speed by four to five times and reduce complexity. However, this condensed action space is dependent on HEV component modelling in order to solve the optimization problem. We can either use the original action space for model-free control or the reduced action space are for simplified control.

3) TD(λ)-Learning Algorithm:

Since the TD()-learning technique has a substantially greater convergence rate and performs better in non-Markovian environments, we use it to derive the best inner-loop reinforcement learning strategy. Each state-action combination (s,a) in this algorithm is given a Q value, indicated by Q(s,a), where state s is denoted by the power demand pdem, vehicle speed v, and battery charge level q, and action an is denoted by discharging the battery with current I and selecting the k-th gear ratio. The predicted discounted cumulative benefit of taking action an in state s is roughly represented by the Q(s,a) value. The following is a summary of the TD()- learning algorithm.

The initial Q values in the TD()-learning process are chosen at random. The agent first chooses an action based on the Q(s,a) values at each time step t for the current state st. The exploration-exploitation policy is used for action selection to reduce the chance of becoming stuck in a suboptimal solution, meaning the agent doesn't always choose the option that produces the highest Q(st,a) value for the current state st. The agent observes a new state st+1 and obtains a reward rt+1 after performing the chosen action at. The agent then modifies the Q(s,a) values for all the state-action pairings, which updates the eligibility e(s,a) of each state-action pair, depending on the observed st+1 and rt+1. A constant between 0 and 1 represents the eligibility e(s,a) of a state-action pair, which indicates the frequency with which the specific state-action combination has been encountered recently. We do not need to change Q values and eligibility e of all state-action pairs because the eligibility of the state-action pairs is used. As the eligibility of all other state-action combinations is at most M, which is insignificant when M is large enough, we simply preserve a list of the M most recent stateaction pairs.

4) Application Specific Improvement of the TD(λ):

Learning Algorithm: By accepting various HEV operating modes, we tweak the TD()-learning algorithm to enhance its performance and convergence rate in the HEV control scenario. In particular, in addition to the recorded Q values, the agent also considers the actual HEV operating mode when choosing an action for the present state. For instance, if the power demand is negative, as in the case of regenerative braking, the agent will undoubtedly select the maximum permitted charging current for the battery pack in order to maximise the recovery of kinetic energy. The agent will be more likely to use EM power to move the vehicle if the battery charge level is very high. Moreover, the agent is likely to utilise more ICE if the battery charge level is very low.

5) Complexity and Model-Free Analysis:

The TD()-learning algorithm has a temporal complexity of O(|A| + M] at each time step, where |A| is the total number of actions and M is the number of recently stored state-action pairings. Since |A| + M typically lies within a few hundred, the algorithm has a very low computational cost. The TD()-learning algorithm typically achieves convergence within L time steps, where L is around three to five times the number of state-action pairs. The TD()-learning method can converge in simulation after an hour of driving because to the application-specific improvement, which is substantially faster than the lifetime of a HEV. The makers may initialise the Q values to further accelerate the convergence rate.

Experimental Results:

We model a PHEV's operation using the model created in the vehicle simulator ADVISOR [1]. Table 1 provides an overview of the PHEV's important characteristics. We put our suggested policy to the test and contrast it with the rule-based and reinforcement learning (RL) policies. We utilise both real-world and test driving trip profiles, which have been created and made available by various initiatives and organisations like the U.S. Environmental Protection Agency and the European Union's MODEM (Modeling of Emissions and Fuel Consumption in Urban Areas project).

TABLE I. PHEV KEY PARAMETERS

Vehicle	Transmission	ICE
m = 1254 kg CR = 0.009 CD = 0.335 AF = 2 m2; rwh = 0.282 m	ρreg = 1.75 ηreg = 0.98 ηgb = 0.98 R(k) = [13.5; 7.6 5.0; 3.8; 2.8]	Peak power 41Kw peak eff. 34% EM peak power 56kW peak eff. 92%
	Battery Capacity 25A·h Voltage 240V	

The running costs of the PHEV during various driving excursions as simulated by Table II under the suggested, RL, and rule-based policies are shown. According to Table II, for instance, the suggested policy causes 0.0028% battery capacity fading and 344.17g of fuel consumption during the MODEM5713 driving trip, which translate to $0.76 in amortised battery replacement costs and $0.37 in fuel consumption costs, for a total operating cost of $1.13.

With an average gas price of $3/gallon in USA and a PHEV battery replacement costing $8,000, A PHEV's battery replacement typically costs between $10,000 and $12,000 [3] for a battery pack with an average capacity of 10kWh. For the 6kWh battery, we use a replacement cost of $8,000. To assess the battery capacity decreasing throughout each journey, we employ the whole cycle-decoupling approach [30]. Table II shows that, when compared to RL and rule-based policies, the proposed policy consistently delivers the lowest operational cost. The proposed policy reduces operational costs by up to 48% when compared to the rule-based policy and up to 47% when compared to the RL policy.

TABLE II. OPERATING COST OF THE PHEV IN DIFFERENT TRIPS USING THE PROPOSED, RL, AND RULE-BASED POLICIES.

Trip	Trip Proposed	RL	Rule
MODEM 5713 cost	0.0028%($0.76) +344.17g($0.37) =($1.13)	0.0045%($1.22) +310.56g($0.33 =($1.55)	0.0044%($1.18) +383.30g($0.41) =($1.59)
Hyzem Motorway cost	0.0018%($0.50) +1991.9g($2.16) 0 =($2.66)	0.0048%($1.28) +2001.9g($2.17) =($3.45)	.0050%($1.36) +2093.6g($2.27) =($3.63)
FTP75 cost	0.0027%($0.73) +311.40g($0.33) =($1.06)	0.0043%($1.16) +295.97g($0.32 =($1.48)	0.0048%($1.30) +623.73g($0.67) =($1.97
US06 cost	0.0028%($0.74) +414.17g($0.45) =($1.19)	0.0043%($1.17) +354.34g($0.38) =($1.55)	0.0036%($0.98) +321.02g($0.34) =($1.32)
UDDS cost	0.0032%($0.85) +298.48g($0.32) =($1.17)	0.0044%($1.19) +355.85g($0.38) =($1.57)	0.0048%($1.30) +630.22g($0.68) =($1.98)
OSCAR cost	0.0021%($0.57) +149.51g($0.16) =($0.73)	0.0043%($1.16) +222.75g($0.24) =($1.40)	0.0042%($1.12) +242.54g($0.26) =($1.38)

The following findings are also based on Table II: I With a PHEV, the amortised battery replacement cost makes up a significant amount of the overall running cost and, for some driving trips, is even higher than the cost of fuel. (ii) For shorter driving distances, the proportionate amortised battery replacement cost is more significant. (iii) In addition to lowering running costs, our suggested approach can greatly increase battery life. (iv) Despite the fact that the RL policy can lower fuel consumption when compared to rule-based policy, in some cases the operating cost from the RL policy is even greater because the RL policy does not account for the battery cost when optimising the fuel usage. (v) When maximising the battery life, the amortised replacement cost is not negligible.

V. Conclusions:

In order to reduce a HEV's running costs, the energy management problem for HEVs is examined in this research utilising a nested learning approach. While the outer loop modifies the battery SoH degradation globally, the inner loop controls the operation modes of the HEV components and is crucial to minimising fuel use. The suggested HEV energy management programme reduces operating costs by up to 48%, according to experimental findings.

References:

[1] ADVISOR 2003 documentation. National Renewable Energy Lab.

[2] High-performance battery monitor IC with coulomb counter, voltage and, temperature measurement. Texas Instruments.

[3] http://batteryuniversity.com/learn/article/hybrid electric vehicle.

[4] E. Alpaydin. Introduction to machine learning. MIT press, 2004.

[5] H. Banvait and et al. A rule-based energy management strategy for plug-in hybrid electric vehicle (phev). In ACC'09.

[6] A. G. Barto. Reinforcement learning: An introduction. MIT press, 1998.

[7] B. M. Baumann and et al. Mechatronic design and control of hybrid electric vehicles. Mechatronics, IEEE/ASME Trans, 2000.

[8] H. Borhan and et al. Mpc-based energy management of a power-split hybrid electric vehicle. Control Systems Technology, IEEE Trans, 2012.

[9] C. Chan. The state of the art of electric, hybrid, and fuel cell vehicles. Proceedings of the IEEE, 2007. [10] C.-C. Chan and et al. Electric, hybrid, and fuel-cell vehicles: Architectures and modeling. Vehicular Technology, IEEE Trans, 2010.

[11] S. Delprat and et al. Optimal control of a parallel powertrain: from global optimization to real time control strategy. In Vehicular Technology Conference, 2002.

[12] S. Ebbesen, P. Elbert, and L. Guzzella. Battery state-of-health perceptive energy management for hybrid electric vehicles. Vehicular Technology, IEEE Transactions on, 61(7):2893–2900, 2012.