LETTER Modeling of Electric Vehicle Charging Systems in Communications Enabled Smart Grids*

Seung Jun BAEK^{†a)}, Daehee KIM[†], Nonmembers, Seong-Jun OH[†], Member, and Jong-Arm JUN^{††}, Nonmember

SUMMARY We consider a queuing model with applications to electric vehicle (EV) charging systems in smart grids. We adopt a scheme where an Electric Service Company (ESCo) broadcasts a one bit signal to EVs, possibly indicating 'on-peak' periods during which electricity cost is high. EVs randomly suspend/resume charging based on the signal. To model the dynamics of EVs we propose an M/M/ ∞ queue with random interruptions, and analyze the dynamics using time-scale decomposition. There exists a trade-off: one may postpone charging activity to 'off-peak' periods during which electricity cost is cheaper, however this incurs extra delay in completion of charging. Using our model we characterize achievable trade-offs between the mean cost and delay perceived by users. Next we consider a scenario where EVs respond to the signal based on the individual loads. Simulation results show that peak electricity demand can be reduced if EVs carrying higher loads are less sensitive to the signal.

key words: queuing systems, time-scale decomposition, quality of service, plug-in hybrid electric vehicles

1. Introduction

As part of global efforts to reduce CO₂ emission, the migration to plug-in hybrid electric vehicles (PHEVs, EVs in short) in automotive industry has recently gained much impetus. Deployment of EVs on a large scale, however, will impose new demands on the current electrical grid infrastructure [1], i.e., it calls for intelligent management and estimation of the power consumed by EVs in future grids. As envisioned by the recent development of smart grids [2] and energy-aware 'smart' appliances, we assume that communication links are established between Electric Service Company (ESCo) and EVs. In our setup a one-bit signal denoted by U(t) is broadcast from ESCo to EVs. When U(t) changes from 0 to 1, EVs randomly decide whether to stop charging, which we will call *disconnecting*, during U(t) = 1. Such disconnected EVs resume charging when U(t) returns to 0. The assertion U(t) = 1 may indicate that the grid undergoes an 'on-peak' period, i.e., the unit cost of electricity is unusually high. Thus U(t) is an implicit control for the number of actively charging EVs in the system which is assumed to be proportional to the power consumed in the grid.

In this letter we consider a queueing model capturing

- ^{††}The author is with Electronics and Telecommunications Research Institute, Daejeon, Korea.
- *This work was supported by Ministry of Knowledge and Economy, Korea, 2006-S-601-01.

a) E-mail: sjbaek@korea.ac.kr

DOI: 10.1587/transinf.E94.D.1708

simple control of charging EVs on a large scale. EVs enter a grid requesting random amounts of electricity to recharge their batteries. Thus the grid can be viewed as a queue with infinite number of servers where the service offered by the queue (grid) corresponds to charging activities by EVs. We consider a queue subject to interruptions to model such randomized charging activities by EVs. Using time-scale decomposition, we propose simple estimates for the average EV population in steady state. Similar approaches [3] have been made to evaluate flow level performances of mobile users assuming wireless channels to be in either infinitely fast or slowly varying limit regimes. We leverage the results on queues with interruptions [4]. The work [5] models the EV demand systems as M/M/n queues, but does not consider the dynamically disconnecting scenarios as in smart grids.

Using the estimates we explore the trade-off between mean delay (sojourn time) experienced by users and the cost of electricity usage. Suppose the unit cost of electricity usage during on-peak periods is substantially higher than that in off-peak periods. As more users disconnect during onpeak periods and shift charging activity to off-peak periods, the mean cost will decrease. However such shifting of loads will incur an increase in the mean sojourn time of EVs. We propose an approximation to mean cost, and show how different degrees of trade-off between cost and delay can be achieved by controlling U(t). We find that for a system in which users are more sensitive to U(t) in the sense that more users disconnect at U(t) = 1 on average, a larger set of costdelay trade-off points can be achieved. This allows ESCo to offer various degrees of QoS to users. By contrast when there is a stringent constraint on delay, we show that it incurs less cost if users are less sensitive to U(t). We demonstrate these results using numerical evaluation of the derived expressions for average cost and delay. Finally we consider the case where the probability of disconnecting depends on the residual amount of electricity to charge at EVs. We model such responsiveness via an 'elasticity' function that maps residual work (demand) to the likelihood of disconnecting in response to control (price) signals. We consider several simple elasticity functions. By simulation we show that peak demands can be reduced if EVs with larger loads disconnect with lower probability. Presumably owners of EVs with nearly depleted batteries will be less willing to disconnect. This demonstrates that, such natural behavior by users helps reducing peak electricity demand on the grid.

Manuscript received December 8, 2010.

Manuscript revised April 6, 2011.

[†]The authors are with College of Information and Communications, Korea University, Seoul, Korea.

2. Model

2.1 Time-Scale Decomposition

We assume that EVs enter the grid according to a homogeneous Poisson process with rate λ . We denote the amount of electricity to be charged by a random variable (RV) S which is exponentially distributed with mean μ^{-1} . The system is modeled as an M/M/ ∞ queue with utilization $\rho = \lambda \mu^{-1}$. Once an EV is in the system it is charged at a rate of unit power. A binary signal U(t) is a continuous time Markov process where the transition rate from 0 to 1 (resp. 1 to 0) is given by α (resp. β). When U(t) = 0 EVs charge their batteries normally: when U(t) changes to 1 EVs independently disconnect with probability (w.p.) p. Newly arriving EVs during U(t) = 1 immediately disconnect w.p. p at the arrival. Note *p* quantifies the responsiveness of users to U(t), e.g., higher p implies higher sensitivity to the signal U(t). The system corresponds to a queue with random interruptions which turns out to be difficult to analyze. Instead we consider the following two cases of time-scale decom*position* as follows. The first case is where U(t) changes at a much faster rate than those of arrivals and service times. The other limiting case is where the variation of U(t) is relatively slow: specifically 'off-peak' period, i.e., the duration of U(t) = 0, is significantly longer than dynamics of EVs. Denote the total number of EVs in the system in steady state in a normal regime, in the regimes of infinitely slow and fast varying U(t) by RVs N, N^s and N^f respectively.

We first consider the case where U(t) has a long offpeak period, i.e., α vanishes, as follows.

Proposition 2.1: Suppose λ , μ and β are fixed. As α tends to 0, N^s is distributed according to $N_1 + N_2$ where N_1 and N_2 are independent random variables where $N_1 \sim \text{Poisson}(\rho)$ and the probability mass function (PMF) of N_2 is given by, up to the first order of α ,

$$\mathbb{P}(N_2 = n) = \frac{\beta}{\alpha + \beta} \mathbb{P}(M_1 = n) + \frac{\alpha}{\alpha + \beta} \mathbb{P}(M_2 = n) \quad (1)$$

where the random variable M_1 (resp. M_2) has negative Binomial distribution with parameters $(\alpha\mu^{-1}, \beta(p\lambda + \beta)^{-1})$ (resp. $(\alpha\mu^{-1} + 1, \beta(p\lambda + \beta)^{-1}))$.

Proof: Consider a M/M/ ∞ system that independently tags arriving EVs as class D EVs w.p. p and as class C EVs otherwise, where only class D EVs disconnect at U(t) = 1 w.p. 1. Class C EVs do not disconnect. As $\alpha \rightarrow 0$ this system becomes stochastically equivalent to ours which can be regarded as having two M/M/ ∞ queues where on their arrivals EVs are independently split and routed w.p. 1 - pto a normal queue (class C) and w.p. p to a queue with Markov modulated service interruptions by U(t) (class D). Let us denote the RVs representing the number of class C and D customers in the system by N_C and N_D respectively. Thus N^s converges in distribution to $N_C + N_D$ where $N_C \sim \text{Poisson}((1 - p)\rho)$ and from [4], N_D is distributed as $\hat{N}_1 + N_2$ where $\hat{N}_1 \sim \text{Poisson}(p\rho)$ and N_2 , which is independent of \hat{N}_1 , has the PMF given by (1). If we define $N_1 := N_C + \hat{N}_1$, since N_C and \hat{N}_1 are independent, we have that $N_1 \sim \text{Poisson}(\rho)$.

The decomposition turns out to be the simpler for the case where U(t) varies infinitely fast: every EV sees an 'average' service rate modulated by U(t), as if every customer is served at a rate that is uniformly reduced by a factor of $1 - \gamma p$ where $\gamma \triangleq \alpha (\alpha + \beta)^{-1}$ is the duty cycle of U(t). Thus $N^f \sim \text{Poisson}(\rho(1 - \gamma p)^{-1})$.

Corollary 1: The mean values for N^s and N^f are given by, up to the first order of α ,

$$\mathbb{E}[N^s] = \rho + \left(\frac{\alpha}{\mu} + \frac{\alpha}{\alpha + \beta}\right) \frac{p\lambda}{\beta}, \quad \mathbb{E}[N^f] = \frac{\rho}{1 - \gamma p} \quad (2)$$

Proof: We obtain $\mathbb{E}[N^s]$ by directly taking the expectations of $N^s \sim N_1 + N_2$, or see (3.8) of Corollary 1 from [4]. Also since $N^f \sim \text{Poisson}(\frac{\rho}{1-\gamma p})$, we get the above $\mathbb{E}[N^f]$. \Box

Later we simulate the population of EVs versus p under different timescales to assess the quality of the estimates (2).

2.2 Trade-Off between Cost and Delay

We investigate how to achieve various degrees of trade-off between cost and delay. In this section we assume that the system is in the regime of slowly varying U(t). Since U(t) is simply an indicator signal and can be controlled by ESCo, ESCo may set U(t) to 1 during only a *fraction* of the actual on-peak period as follows. If ESCo detects the beginning of an on-peak period, U(t) is asserted to 1. However ESCo may set U(t) to 0 *prior* to the termination of the on-peak period in order to reduce delay.

We construct U(t) and on-peak periods to capture such a scheme as follows. Consider sequences of RVs X_i , Y_i and Z_i for $i \in \mathbb{Z}_+$ which are generated i.i.d. from $X \sim \text{Exp}(\beta)$, $Y \sim \text{Exp}(\alpha)$ and $Z \sim \text{Exp}(\eta - \alpha)$ respectively where $\eta \triangleq (\tau - \beta^{-1})^{-1}$ for some $\tau \in [\beta^{-1}, \alpha^{-1} + \beta^{-1}]$. We associate the duration of the *i*-th on-peak period, off-peak period, U(t) = 1 and U(t) = 0 with $X_i + W_i$, $Y_i - W_i$, X_i and Y_i respectively where $W_i \triangleq \min(Y_i, Z_i)$. It is easy to verify that such a construction renders U(t) a Markov process. From the definition, the mean duration of on-peak periods is given by $\mathbb{E}[X + W] = \mathbb{E}[X] + \mathbb{E}[\min\{Y, Z\}] = \beta^{-1} + \eta^{-1} = \tau$. A time diagram illustrating the mean values of on/off-peak periods and U(t) is shown in Fig. 1. We assume $\alpha^{-1} + \beta^{-1}$ is normalized to 1, i.e., on-peak periods recur on a regular basis. We assume that τ is a fixed and β is chosen by ESCo.

Class *C* EVs, as defined in the proof of Proposition 2.1, will be charging for the whole duration of τ regardless of U(t). Class *D* EVs will stay disconnected for β^{-1} on average, however will be charging for the remaining on-peak period of duration $\tau - \beta^{-1}$ on average: see Fig. 1. Thus the average total time spent by actively charging EVs during onpeak periods, denoted by E_{on} , is approximated as follows:

Ì

$$\mathcal{E}_{\text{on}} \approx \mu^{-1} \{ \mathbb{E}[N_C] \cdot \tau + \mathbb{E}[N_D | U(t) = 0] \cdot (\tau - \beta^{-1}) \} \quad (3)$$



Fig.1 A time diagram for actual on-peak/off-peak periods and U(t).

where N_C and N_D are defined in the proof of Proposition 2.1. Since the dynamics of Class *C* EVs are independent of U(t), we have that $\mathbb{E}[N_C] = (1 - p)\rho$. The expression in the bracket of (3) approximates the average total number of actively charging EVs during on-peak period. Similarly the total time spent by EVs during off-peak periods is approximated as follows:

$$E_{\text{off}} \approx \mu^{-1} \{ \mathbb{E}[N_C](1-\tau) + \mathbb{E}[N_D|U(t) = 0](1-\tau) \}$$
(4)

Denote the rate of cost of electricity usage during on-peak and off-peak periods by c_{on} and c_{off} respectively where $c_{on} \ge c_{off}$. The mean cost K is defined by $K = c_{on}E_{on} + c_{off}E_{off}$.

Proposition 2.2: In steady state the mean number of class *D* EVs conditional on U(t) = 0 and U(t) = 1 is respectively

$$\mathbb{E}\left[N_D|U(t)=0\right] = p\rho\left(1+\frac{\alpha}{\beta}\right),\tag{5}$$

$$\mathbb{E}\left[N_D|U(t)=1\right] = p\left(1+\frac{\alpha}{\beta}\right)\left(\rho+\frac{\lambda}{\alpha+\beta}\right).$$
(6)

Proof: Let us define

$$p_0(n) = \mathbb{P}(N_D = n, U(t) = 0)$$

 $p_1(n) = \mathbb{P}(N_D = n, U(t) = 1).$

Also define

$$G_0(z) = \sum_{k=0}^{\infty} p_0(k) z^k, \qquad G_1(z) = \sum_{k=0}^{\infty} p_1(k) z^k.$$

Note the dynamics of class D EVs is that of M/M/ ∞ queue with Markov modulated interruptions. Thus we leverage the result from [4, Eq. (4.2)] such that

$$G_1(z) = \frac{\alpha}{p\lambda - p\lambda z + \beta} G_0(z).$$
(7)

By differentiating both sides of (7) by z we have that

$$G_1'(z) = \frac{\alpha}{p\lambda - p\lambda z + \beta}G_0'(z) + \frac{p\alpha\lambda}{(p\lambda - p\lambda z + \beta)^2}G_0(z)$$

Thus

$$G_1'(1) = \frac{\alpha}{\beta}G_0'(1) + \frac{p\alpha\lambda}{\beta(\alpha+\beta)}$$
(8)

where we have used $G_0(1) = \sum_{k=0}^{\infty} p_0(k) = \mathbb{P}(U(t) = 0) =$

 $\beta(\alpha + \beta)^{-1}$. By using $\mathbb{E}[N^s] = \mathbb{E}[N_D] + \mathbb{E}[N_C] = \mathbb{E}[N_D] + (1 - p)\rho$ and from the definition $\mathbb{E}[N_D] = G'_0(1) + G'_1(1)$,

$$G'_{0}(1) + G'_{1}(1) = \mathbb{E}[N_{D}] = \mathbb{E}[N^{s}] - (1 - p)\rho$$
$$= p\rho + \left(\frac{\alpha}{\mu} + \frac{\alpha}{\alpha + \beta}\right)\frac{p\lambda}{\beta}$$
(9)

where we have used (2) for $\mathbb{E}[N^s]$. From (8) and (9) we can solve for $G'_0(1)$ and $G'_1(1)$. Finally from $G'_0(1) = \mathbb{E}[N_D|U(t) = 0]\mathbb{P}(U(t) = 0)$ and $G'_1(1) = \mathbb{E}[N_D|U(t) = 1]\mathbb{P}(U(t) = 1)$, we get the results.

From (3), (4) and (5), *K* is given by:

$$K = \mu^{-1} (1 - p) \rho \{ c_{\text{on}} \tau + c_{\text{off}} (1 - \tau) \} + \mu^{-1} p \rho \left(1 + \frac{\alpha}{\beta} \right) \{ c_{\text{on}} \{ \tau - \beta^{-1} \} + c_{\text{off}} (1 - \tau) \}.$$
(10)

By Little's result, the mean delay denoted by T is given by

$$T = \lambda^{-1} \mathbb{E}[N^s] = \frac{\rho}{\lambda} + \left(\frac{\alpha}{\mu} + \frac{\alpha}{\alpha + \beta}\right) \frac{p}{\beta}.$$
 (11)

With other parameters fixed, by substituting α with $(1 - \beta^{-1})^{-1}$ in (10) and (11) we will treat *T* and *K* as functions of β . By numerical methods we characterize (*T*, *K*) pairs which are achievable by varying β for different values of *p* in Sect. 3.

2.3 Disconnecting Based on Residual Battery Capacity

We consider the case where the probability of disconnecting by an EV depends on the residual amount of electricity to fully charge the battery. We assume that the full battery capacity is 1 and that *S* is uniformly distributed on [0, 1]. Let $f : [0, 1] \rightarrow [0, 1]$ map *S* to the probability of disconnecting. We would like to reduce 'peak demands' defined by $\mathbb{P}(N_a > b)$ for some large *b* where N_a is an RV representing the number of actively charging EVs in steady state. We consider several types of $f(\cdot)$ and evaluate its impact on the peak demands using simulation in Sect. 3.

3. Numerical and Simulation Results

We present numerical results on the accuracy of the derived estimates (1). We consider three cases 'slow', 'fast' and 'moderate' which represent the relative timescales of U(t)to those of arrival/departure processes. Figure 2 shows the mean population of EVs against the probability of disconnecting where E[N] denotes the simulated average number of EVs in the system. The parameters for each timescale are denoted by 4-tuple $(\lambda, \mu, \alpha, \beta)$ and they are (1, 0.2, 0.5, 1), (1, 0.2, 10, 10) and (1, 0.3, 1, 1) for 'slow', 'fast' and 'moderate' regimes respectively in Fig. 2. In 'slow' regime offpeak periods are only 2 times longer than interarrival times and the same as service times on average. Nonetheless in Fig. 2 we observe that, in 'slow' regime, the estimate $\mathbb{E}[N^s]$ is close to $\mathbb{E}[N]$. Also in 'fast' regime we see that $\mathbb{E}[N^f]$ well approximates $\mathbb{E}[N]$. In 'moderate' regime where all



Fig. 2 Estimates of $\mathbb{E}[N]$ under different timescales.



Fig. 3 Trade-off curves for delay and electricity cost with varying *p*. The curves grouped by $\tau = 0.25$ and 0.75 are normalized by the maximum delay and cost values among the simulated range of values in that group.

the parameters are similar, our estimates remain good.

Figure 3 is a plot of the pairs (T, K) obtained by varying β with different values of p. The parameters are: $c_{on} = 2$, $c_{\text{off}} = 1, \mu = \mu_0$ and $\lambda = \lambda_0$ where $\mu_0^{-1} := 150$ min. and $\lambda_0^{-1} := 1.5$ min. There are two sets of curves for $\tau = 0.25$ and 0.75. The curves exhibit trade-off relations between delay and cost. When $\tau = 0.75$, suppose there exists some constraint on T, say 0.4, then the lowest cost can be achieved when p = 0.3 with $K \approx 0.855$. This shows that for lower p, one may get a lower cost subject to a delay constraint. This is because, for larger p with a constraint on T, ESCo has to keep the mean duration of U(t) = 1 small in order to meet delay constraint, however this causes more class D EVs to charge during on-peak period, which will increase K. However for small p, there is a limitation in cost saving, e.g., when p = 0.3 in Fig. 3, K cannot be lowered below 0.855. By contrast if p is large and constraints on T are loose, one can achieve a larger set of (T, K), e.g., p = 0.9 case in Fig. 3. A similar relation holds for the case where $\tau = 0.25$.

Finally we examine what elasticity function is effective in reducing peak demands. As candidate elasticity functions



Fig. 4 CDF of the number of actively charging EVs.

we consider f(x) = 0.5, f(x) = x and f(x) = 1 - x. The parameters of simulations are $\mu = \mu_0$, $\lambda = \lambda_0$, $\alpha = 0.5$ and $\beta = 0.5$. Figure 4 shows the cumulative distribution function (CDF) of the number of actively charging customers. From the tail of the distribution, we see that f(x) = 1 - x incurs the lowest peak demands. Decreasing $f(\cdot)$ implies that the customers with larger demands are less likely to disconnect, which seems reasonable since a newly plugged EV with a nearly depleted battery would not want to postpone the charging process. Such a natural behavior by users based on QoS is indeed beneficial to the overall system performance as well. The intuition is that, decreasing $f(\cdot)$ attempts to balance the variability in the sojourn times of EVs in the system caused by varying workloads offered to the system – this results in lighter tail in the CDF as in Fig. 4.

4. Conclusion

We proposed a queuing model with random interruptions for EV charging systems for smart grids. We derived simple estimates of the number of EVs using time-scale decomposition. From the estimates we characterized achievable tradeoffs between delay and electricity cost. We studied scenarios where EVs disconnect based on the residual amount of electricity to be charged. It is shown that the decreasing elasticity function is effective in reducing peak demands.

References

- R. Pratt, M. Scott, and M. Warwick, "Potential impacts of high penetration of plug-in hybrid vehicles on the U.S. power grid," DOE/EERE PHEV Stakeholder Workshop, June 2007.
- [2] M. Amin and B. Wollenberg, "Toward a smart grid: power delivery for the 21st century," IEEE Power and Energy Magazine, vol.3, no.5, pp.34–41, 2005.
- [3] T. Bonald, S.C. Borst, and A. Proutiere, "How mobility impacts the flow-level performance of wireless data systems," Proc. IEEE INFO-COM, 2004.
- [4] M. Baykal-Gursoy and W. Xiao, "Stochastic decomposition in M/M/∞ queues with Markov modulated service rates," Queueing Systems, vol.48, pp.75–88, 2004.
- [5] R. Garcia-Valle and J. Vlachogiannis, "Electric vehicle demand model for load flow studies," Electric Power Components and Systems, vol.37, pp.577–582, 2009.