A NOVEL TECHNIQUE FOR MODELING STATES TRANSITION ENERGY CONSUMPTION OF RADIO FREQUENCY TRANSCEIVERS IN WSN USING MARKOV CHAIN

John Adinya Odey
University of Calabar

Bassey Ele
University of Calabar

Anthony O Otiko
Cross River University of Technology

Prince Ana
Cross River University of Technology

A NOVEL TECHNIQUE FOR MODELING STATES TRANSITION ENERGY CONSUMPTION OF RADIO FREQUENCY TRANSCEIVERS IN WSN USING MARKOV CHAIN

[1]John Odey, [1]B.I. Ele, [2]A.O. Otiko, [2]Prince Ana

[1] Department of Computer Science University of Calabar, Calabar, Nigeria
[2] Department of Computer Science Cross River University,Calabar, Nigeria

ABSTRACT
Energy optimization remains an important consideration in the designs and implementations of wireless sensor networks. One of the often recommended techniques to low power designs in Wireless Sensor Networks is duty cycling between sleep and active states to optimize power consumption. However most suggested duty cycling techniques fail to appreciate energy consumptions during state transitions and the performance loss incurred due to overheads associated with state change.In this paper we present a new technique for modeling and accurately predicting the energy consumption during transition between various states of the radio frequency transceiver of a typical WSN node. Our Model enables predictions by some precise stochastic probabilities of the next state of the transceiver using the Markov Chain. The proposed technique is used to calculate the transition energy of the transceiver between states and enable decisions on when to switch between states or remain in a particular state in continuous time.

Keywords: Markov Chain, Low power WSN, Transceiver Transition Energy Consumption, Transition Matrices, Duty Cycling, Random Walk, Pascal Triangle.

1. INTRODUCTION

The development of wireless sensor networks (WSN) technology in the last decade has enable the ubiquitous applications of sensors from everyday activities to remote and mission critical areas. Despite the many advantages and characteristics of WSN, one considerable drawback is that battery powered WSN often requires frequent service intervals due to the limited capacity of the energy storage element. The maintenance cost to replace or recharge the hundreds of WSN nodes may exceed the system cost in a relatively short period of time (Wang W.S et al , 2011) . One of the consistent and pervasive concerns when designing WSN and WSN applications is how to manage power consumption to optimize the lifespan of such networks. Several energy optimization techniques in WSN centered on duty cycling are presented in literature and majority often fall in the categories of protocol design (Ahmad and Dutkiewicz, 2009) and dynamic power management policies(DPM)(Benini and Micheli, 1997).

The basic idea of low duty cycle protocols is to reduce is to reduce the time a node is idle or spend overhearing an unnecessary activity by putting the node in sleep state. DPM encompasses a set of techniques that achieve energy-efficient computation by selectively turning off (or re-ducing the performance of) system components when they are idle (or partially unexploited)(Benini and Micheli, 1997). The existing DPM policies can be broadly classified into three categories; Timeout Policies, Predictive Policies and Stochastic Policies.

The basic assumption of Timeout policies is that if the device remains idle for τ, then it should further stay idle for at least T_{be}. In (Karlin, 1994) the timeout policy proposed guarantees it would not consume more than twice the energy of an ideal offline policy. The authors in (Douglis et al., 1995), applied adaptive time out policy where *t* is increased or decreased arithmetically or geometrically on the ratio of performance delay and sleep time. The disadvantage of the timeout policies is that they waste energy while waiting for the timeout to expire.

In Predictive policies , the length of the upcoming idle period is pedicted. Thus one can make a decision immediately on whether to sleep or not depending upon the prediction being greater or less than T_{be}. (Hwang and Wu, 1997) proposes a policy where the upcoming idle period length is calculated by taking an exponential average of the predicted and actual lengths of the previous idle period. Some other predictive policies use an adaptive learning tree to make predictions (Chung et al., 1999).

In Stochastic policies models, minimizing power consumption and performance delays becomes stochastic optimization problems. Authors (Paleologo et al., 1998) and (Qiu and Pedram, 1999) modeled DPM as a Markov decision processes where services and inter request arrival time is modeled as an exponential distribution. Both view request and service as memoryless distributions where future states are dependent only on the current states.

In most practical instances, Power Managed Components (PMC) is modeled by a finite-state representation called power state machine (PSM) (Benini, 2000). (Odey and Li, 2012) proposes an energy consumption model, where energy consumption in WSN transceiver is modeled as a finite state machine of recognizable finite and transition states where states are the various modes of operation that span the tradeoff between performance and energy consumption.

Transitions between finite states of operation have a cost. In many cases, the cost is in terms of delay or performance loss if transition is not instantaneous, and the component is not operational during a transition. In most duty cycling and energy optimization techniques in WSN, what happens within the transition states still remains unpredictable and can serve as the flaw in an efficient energy consumption policy. For instance, if entering a low-power state requires power-supply shutdown, returning from this state to the active state requires a (possibly long) time for: 1) turning on and stabilizing the power supply and the clock; 2) reinitializing the system; and 3) restoring the context (Benini, 2000). An effective energy optimization must seek to maximize power savings while keeping performance degradation within acceptable limits. In other words, we need to decide when it is worthwhile (performance and power-wise) to transition to a low-power state and which state should be chosen (if multiple low-power states are available) (Benini, 2000).

In this paper as in (Paleologo et al., 1998) and (Qiu and Pedram, 1999), we propose an energy optimization model for the transition state of a transceiver in a WSN node as a Markovian process with stochastic abilities in finite increment of time.

2. Theoretical Foundation:

Definition 2.1.1 A *stochastic process* is a family of random variables $\{X_{(t)}, t > 0\}$ where t is the time parameter. The values assumed by the process are called the *states*, and the set of possible values is called the *state space*.

Definition 2.1.2 A (discrete-time) Markov chain with (finite or countable) state space X is a sequence $X_0, X_1, ...$ of X valued random variables such that for all states i, j, k_0, k_1, and all times n = 0, 1, 2,...,

$$P(X_{n+1} = j \,|\,|\, X_n = i,\ X_{n-1} = K_{n-1}, \ldots) = P(i, j) \quad (1)$$

where P(i, j) depend only on the states i, j and not on the time n or the previous states $Kn\text{-}1, n\text{-}2, \ldots$ The numbers P(i, j) are called the transition probabilities of the chain.

Proposition 2.1 If X_n is a Markov chain with transition probabilities P(x,y) then for every sequence of states $X_0, X_1, \ldots X_{n+m}$,

$$P(X_{m+i} = x_{m+i}\ \forall\ 0 < i \leq n \parallel Xi = xi\ \forall\ 0 \leq i \leq m) = \prod_{i=1}^{n} P(x_{m+i-1}, x_{m+i}). \quad (2)$$

Consequently, the n-step transition probabilities

$P^n(x,y) := P(X_{n+m} = y \parallel X_m = x)$ (3)

depends only on the time lag *n* and the initial and terminal states *x,y* , but not on *m.*

Proof. The first statement can be proved by a completely routine induction argument, using the definition of a Markov chain and elementary properties of conditional probabilities. The second follows from the first, by summing over all possible sequences x_{m+i} of intermediate states: the right side of equation (3) makes it clear that this sum does not depend on *m*, since the factors in the product depend only on the transitions x_i , x_{i+1} made, and not on the times at which they are made.

Definition 2.2.1 A state *j* is said to be accessible from state *i* if there is a positive-probability path from *i* to *j* , that is, if there is a finite sequence of states $k_0,k_1,...,k_m$ such that $k_0 = i$, $k_m = j$, and $P(k_t , k_{t+1}) > 0$ for each $t = 0, 1,...,m-1$. States *i* and *j* are said to communicate if each is accessible from the other. This relation is denoted by $i \leftrightarrow j$.

Fact 1. Communication is an equivalence relation. In particular, it is transitive: if *i* communicates with j and j communicates with k then i communicates with k .

It follows that the state space *X* is uniquely partitioned into communicating classes (the equivalence classes of the relation ↔). If there is only one communicating class (that is, if every state is accessible from every other) then the Markov chain (or its transition probability matrix) is said to be irreducible. In general, if there is more than one communicating class, then states in one communicating class C1 may be accessible from states in another class C2; however, in such a case no state of C2 can be accessible from a state of C1.

Definition 2.2.2 The period of a state i is the greatest common divisor of the set $\{n \in N: P_n(i,i) > 0\}$. If every state has period 1 then the Markov chain (or its transition probability matrix) is called aperiodic.

Note: If *i* is not accessible from itself, then the period is the *g.c.d.* of the empty set; by convention, we define the period in this case to be $+ \infty$.

Fact 2. If states *i, j* communicate, then they must have the same period. Consequently, if the Markov chain is irreducible, then all states have the same period. There is a simple test to check whether an irreducible Markov chain is aperiodic: If there is a state *i* for which the 1-step transition probability $P(i,i) > 0$, then the chain is aperiodic.

Fact 3. If the Markov chain has a stationary probability distribution π for which $\pi(i) > 0$, and if states *i, j* communicate, then $\pi(j) > 0$.

Definition 2.2.3 A state *i* is said to be *recurrent* if and only if, starting from *i*, eventual return to this state is certain. A recurrent state is said to be *positive recurrent* if and only if the mean time to return to this state is finite. A state *i* is said to be *transient* if and only if, starting from *i*, there is a positive probability that the process may not eventually return to this state.

Definition 2.2.4 . If the set of all states of a stochastic process *X* form a single communicating class, then *X* is *irreducible*.

Theorem 2.1

(1) If the Markov process is irreducible, then the limiting distribution $\lim_{t\to\infty} P_i(t) = P_i$, $i \in S$, exists and is independent of the initial conditions of the process, The limits $\{P_n | n \in S)$ are such that they either vanish identically (i.e., $P_i = 0$ for all $i \in S$) or are all positive and form a probability distribution (i.e., $P_i > 0$ for all $i \in S$, $\sum_{i \in S} P_i = 1$).

(2) The limiting distribution $\{P_i,\ i \in \mathbf{S}\}$ of an irreducible positive recurrent Markov process is given by the unique solution of the equation: $\boldsymbol{PG} = 0$ and $\sum_{i \in S} P_j = 1$ where $\boldsymbol{P} = (p_0, p_1, \ldots)$.

Definition 2.3 The random walk on an undirected graph is a Markov chain where the states are represented as the vertices of the graph, and a transition consists of choosing an edge through the vertex on which the marker sits (all edges being equally likely) and moving to the other end of this edge.

Theorem 2.2 : Let G be an undirected, connected, non-bipartite graph on n vertices. Then:

(1) There is a unique stationary distribution $\pi = (\pi(1), \ldots, \pi(n))$. Furthermore, all entries in π are non-zero.

(2) For all vertices i, j, we have $\lim_{t\to\infty} P^t_{i,j} = \pi(j)$. Note that the limit is independent of i. In words, this means that no matter where we start the random walk (i.e., regardless of our starting point i) we end up in state j (for large enough t) with the same probability $\pi(j)$.
(3). Let $h_{i,i}$ denote the expected number of steps for a random walk beginning at vertex i to return to i. Then $h_{i,i} = 1/\pi(i)$.

Let G = (X; E) be a connected graph with *n* nodes and *m* edges. Consider a random walk on G: we start at a node X_0; if at the *t*-th step we are at a node X_t , we move neighbor of X_t with probability $1/d(X_t)$. Clearly, the sequence of random nodes (X_t: t = 0;1; : : :) is a Markov chain. The node X_0 may be fixed, but may itself be drawn from some initial distribution P_0. We denote by P_t the distribution of X_t:

$P_t(i) = Prob(v_t = i)$:

We denote by $M = (P_{ij})_{i;j \in V}$ the matrix of transition probabilities of this Markov chain. So

$$P_{ij} = \begin{cases} \frac{1}{d(i)}, & \text{if } i,j \in E \\ 0, & \text{otherwise} \end{cases}$$

Let A_G be the adjacency matrix of *G* and let *D* denote the diagonal matrix with $(D)_{ii} = 1/d(i)$, then $M = DA_G$. If *G* is *d*-regular, then $M = (1/d)A_G$.

The rule of the walk can be expressed by the simple equation

$$P_{t+1} = M^T P_t,$$

(the distribution of the *t*-th point is viewed as a vector in R^V), and hence

$$P_t = (M^T)^t P_0 ,$$

It follows that the probability P^t_{ij} that, starting at *i*, we reach *j* in *t* steps is given by the *ij*-entry of the matrix M^t.

If G is regular, then this Markov chain is symmetric: the probability of moving to u, given that we are at node v, is the same as the probability of moving to node v, given that we are at node u. For a non-regular graph G, this property is replaced by time-reversibility: a random walk considered backwards is also a random walk. More exactly, this means that if we look at all random walks (X_0; : : : ; X_t), where X_0 is from some initial distribution P_0, then we get a probability distribution P_t on X_t. We also get a probability distribution Q on the sequences (X_0; : : : ; X_t). If we reverse each sequence, we get another probability distribution Q_0 on such sequences. Now time reversibility means that this distribution Q_0 is the same as the distribution obtained by looking at random walks starting from the distribution P_t. The probability distributions P_0; P_1; : : : are of course different in general. We say that the distribution P_0 is stationary (or steady-state) for the graph G if $P_1 = P_0$. In this case, of course, $P_t = P_0$ for all $t \geq 0$; we call this walk the stationary walk. A one-line calculation shows that for every graph G, the distribution

$$\pi(X) = \frac{d(v)}{2m}$$

is stationary. In particular, the uniform distribution on *X* is stationary if the graph is regular.

3. SYSTEM MODELING

From (Odey and Li, 2012) we model the energy consumption of the transceiver as an aggregation of energy consumption of finite states and the energy expended in transition between this states as shown in **figure 1**. **Figure 2** depicts the energy consumption of transceiver during transition between finite states. Our paper addresses the the energy optimization problem of the transition period t, representing one cycle of transition that captures both entry and exit of both Sleep and Idle states.This is the transition cycle that occur more frequently and are often the subject of most duty cyling and dynamic power optimization research where balance is sought between keeping the nodes in idle states or sleep states.

$$T_{dt} = t3 - t2 + t5 - t4 \quad (4)$$

Also the energy consumed in Sleep State is given as

$$E_{slp} = E_t + E_{Sleep} \quad (5)$$

where E_{sleep} is energy consumed in sleep state and E_t is the transition energy.

Let us consider a WSN transceiver node whose state transition delay time is T_{dt} (including shutdown and wake-up delays) when the transition energy E_t is expended in a single transition cycle. Suppose the power consumed in the Idle and sleep states is P_{idle} and P_s respectively. The minimum length of time, the transceiver stays in idle state to save power is called the break-even time T_{be}. It is also convenient to define the break even time T_{be} in the sleep state as the product of the transition delay time T_{dt} and the minimum length of time T_{ms} spent in the sleep state to save energy: $T_{be} = T_{dt} + T_{ms}$,

The transceiver breakeven point is the case in time when the energy spent in Idle State equates or balances- out the energy spent in the sleep state (E_{slp})

$$P_{Idle} \times T_{be} = E_t + P_s (T_{be} - T_{dt}) \quad (6)$$

or

$$T_{be} = (E_t - P_s \times T_{dt})/P_{Idle} - P_s \quad (7)$$

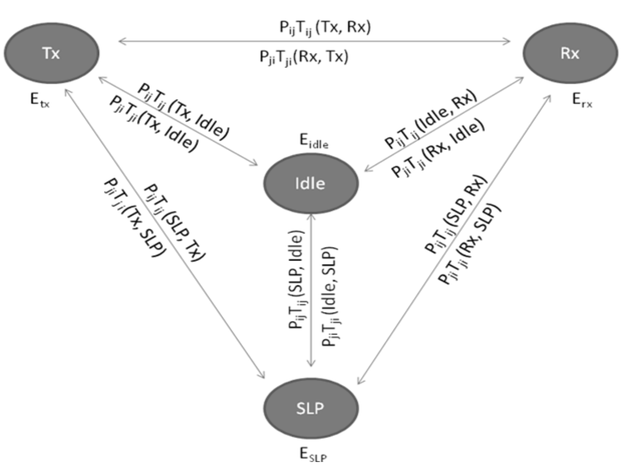

Figure 1: Transceiver Energy Consumption model.

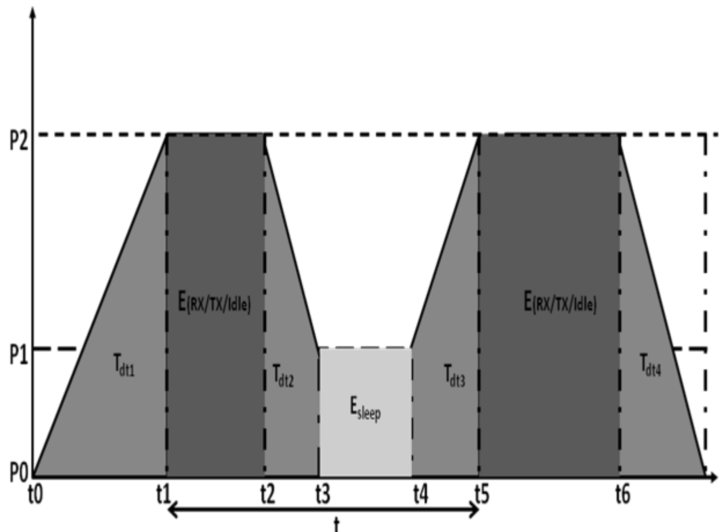

Figure 2: Transceiver Energy Consumption.

Let us model our transceiver energy consumption as a Markov chain of 2X2 matrix of Idle *(I)* and Sleep *(S)* states in discrete time represented as the vertices of a weighted graph with transition probabilities P(i,j) and an initial Probability distribution function u. Let us consider the transition to the next states as a one dimensional random walk along the weighted edge of a graph; the Pascal triangle (Figure 3) starting from the vertex. We shall imagine walking from the top vertex of the Pascal triangle down to the vertex below along the edges, i.e., at every vertex on our transition we are given a choice whether to proceed to the right *(I)* or to the left *(S)* vertices immediately below. A walk to the next vertex depends only on the current vertex and not on previous vertices. So that at every vertex we have two choices whether to step right or left to a new vertex with probability proportional to the weight of the corresponding edge.

As the walk is updated according to a 2×2 matrix of transition probabilities, the (*I*,*S*)th entry of which gives the probability that the model moves from vertex *I* to vertex *S* at any time change (It doesn't matter if the starting vertex is *I* or *S*, the powers of the transition matrix approach a matrix with constant columns as the power increases. The number to which entries in the *I*th column converge is the asymptotic fraction of time the transceiver spends in state *I*; represented by the invariant probability distribution of convergence. This invariant probability distribution of convergence is only possible if our random walk satisfies the conditions of irreducibility and aperiodicity stated in the theorems in section II. If the initial probability distribution of our transition energy model is 0.5, and the random walk through the edges of the Pascal triangle is modeled as a Bernoulli event with a binomial distribution, we still arrive at literary points of convergence. There exist some point in the transitions with equal median distribution both left and right, however the random walk will not achieve an invariant distribution.

To achieve the desired state of our transceiver, we can model a random walk through a favoured state (*S* or *I*) by skewing the initial probability distribution in favour of any particular state (though we will still arrive at the same points of convergence). In the random walk through the Pascal triangle, knowing the convergence point enables transition energy optimization decisions to be taken as this equates the breakeven time in our transceiver energy consumption.

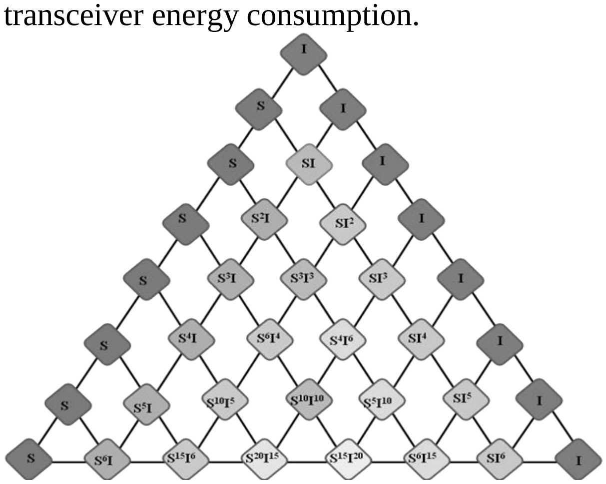

Figure 3: Pascal Triangle showing random walk paths.

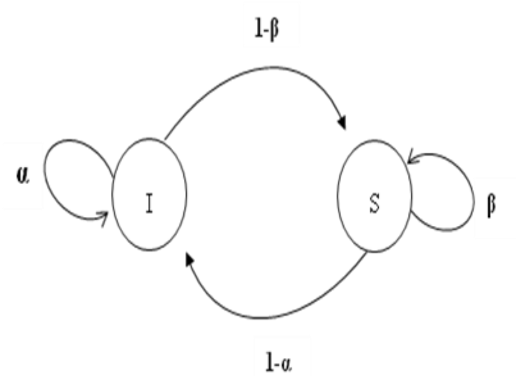

Figure 4: State diagram of a 2x2 matrix

Let us consider our models mathematically here as a 2x2 matrix represented in the state diagram below with a transition matrix *P*.

$P = \begin{bmatrix} \alpha & 1-\beta \\ 1-\alpha & \beta \end{bmatrix}$, with a ∈ [1,0] and b ∈ [0,1].

$$P - 1\cdot I = \begin{bmatrix} \alpha-1 & 1-\beta \\ 1-\alpha & \beta-1 \end{bmatrix}$$

$$det(P - 1\cdot I) = \left| \begin{bmatrix} \alpha-1 & 1-\beta \\ 1-\alpha & \beta-1 \end{bmatrix} \right|$$

$$det(P - 1\cdot I) = (\alpha - 1)\cdot(\beta - 1) - (1-\beta)\cdot(1-\alpha) = 0$$

$$P - \lambda\cdot I = \begin{bmatrix} \alpha-\lambda & 1-\beta \\ 1-\alpha & \beta-\lambda \end{bmatrix}$$

$$(\alpha - \lambda)\cdot(\beta - \lambda) - (1-\alpha)\cdot(1-\beta) = 0$$

$$\lambda^2 - (\alpha+\beta)\cdot\lambda - (1+\alpha+\beta) = 0$$

$$\frac{\lambda^2 - (\alpha+\beta)\cdot\lambda - [1-(\alpha+\beta)]}{(\lambda-1)} = \lambda_2 + [1-(\alpha+\beta)]$$

$$\lambda_2 = (\alpha+\beta) - 1$$

Since $0 < \alpha < 1$ and $0 < \beta < 1$, $0 < |\lambda 2| < 1$. Therefore $\lambda 1 = 1$ is the dominant eigenvalue. This fact will manifest itself when we demonstrate that the corresponding eigenvector is indeed the steady state vector Xss.
Now let us find the corresponding eigenvector.

$$P - 1\cdot I = \begin{bmatrix} \alpha-1 & 1-\beta \\ 1-\alpha & \beta-1 \end{bmatrix}$$

$$\begin{bmatrix} \alpha - 1 & 1 - \beta \\ 0 & 0 \end{bmatrix}$$

$(\alpha - 1) \cdot X_1 = (\beta - 1) \cdot X_2.$

$P \cdot V = V$

$$\begin{bmatrix} \alpha & 1 - \beta \\ 1 - \alpha & \beta \end{bmatrix} \cdot \begin{pmatrix} 1 - \beta \\ 1 - \alpha \end{pmatrix} = \begin{pmatrix} 1 - \beta \\ 1 - \alpha \end{pmatrix}$$

But the components of the vector $V = \begin{pmatrix} 1 - \beta \\ 1 - \alpha \end{pmatrix}$ must add to *1*. Otherwise, it cannot be a state vector.

$$V1 = \begin{bmatrix} \alpha \cdot (1 - \beta) \\ \alpha \cdot (1 - \alpha) \end{bmatrix} = \begin{bmatrix} \frac{(1-\beta)}{[2 - (\alpha+\beta)]} \\ \frac{(1-\alpha)}{[2 - (\alpha+\beta)]} \end{bmatrix}$$

If the steady state vector $\boldsymbol{X_{ss}}$ is the eigenvector corresponding to $\boldsymbol{\lambda = 1}$ and the steady-state vector can also be found by applying *P* to any initial state vector a sufficiently large number of times, *m*, then P^m must approach a specialized matrix.

Let the $\lim_{m \to N} P^m = Q$ for the matrix $P = \begin{bmatrix} \alpha & 1 - \beta \\ 1 - \alpha & \beta \end{bmatrix}$ where "N" is a very large positive integer.

$P \cdot X_{ss} = X_{ss}$

$Q \cdot X_0 = X_{ss}$

$$X_0 = \begin{pmatrix} \gamma \\ 1 - \gamma \end{pmatrix}$$

$$\begin{bmatrix} q11 & 1 - q22 \\ (1 - q11) & q22 \end{bmatrix} \cdot \begin{pmatrix} \gamma \\ 1 - \gamma \end{pmatrix} = \begin{bmatrix} \frac{(1-\beta)}{[2 - (\alpha+\beta)]} \\ \frac{(1-\alpha)}{[2 - (\alpha+\beta)]} \end{bmatrix}$$

$$(-q_{11} - q_{22} - 1) \cdot \gamma + (1 - q_{22}) = \frac{(1-\beta)}{[2 - (\alpha+\beta)]}$$

$$(1 - q_{11} - q_{22} - 1) \cdot \gamma + (q_{22}) = \frac{(1-\alpha)}{[2 - (\alpha+\beta)]}$$

For both equations above to be true for all values of $0 \le \gamma \le 1$, $q_{11} + q_{22} = 1$. Then we obtain these results.

$$q_{11} = \frac{(1-\beta)}{[2 - (\alpha+\beta)]} \qquad q_{22} = \frac{(1-\alpha)}{[2 - (\alpha+\beta)]}$$

$$Q = \begin{bmatrix} \frac{(1-\beta)}{[2 - (\alpha+\beta)]} & \frac{(1-\beta)}{[2 - (\alpha+\beta)]} \\ \frac{(1-\alpha)}{[2 - (\alpha+\beta)]} & \frac{(1-\alpha)}{[2 - (\alpha+\beta)]} \end{bmatrix} = \boldsymbol{(}X_{ss} \cdot X_{ss})$$

From equation 7, the breakeven time T_{be} for the Idle and Sleep state is given as

$T_{be} = (E_t - P_s \text{ x } T_{dt})/P_{idle} - P_s$

Scaling to achieve a scalar and Vector arithmetic, we arrive at

$\mathbf{Q} = (E_t - P_s \text{ x } T_{dt})/P_{idle} - P_s$

Therefore our Transceiver Transition Energy consumption Model is given as

$E_t = \mathbf{Q} \cdot (P_{Idle} - P_s) + P_s T_{dt}$

For example let us consider the *I* and *S* to be the Idle and Sleep states of our transceiver respectively, and if we assume that 70% of the time, the transceiver is in the Idle state and while the transceiver stays in the sleep state (*S*) 20% of the same time. We can determine the probability of a new transition to either the sleep or idle state after 20 transitions. Now after a period of iterative time transitions, the distribution of our sleep (*S*) state and Idle (*I*) states will be given as

$$P = \begin{bmatrix} p11 & p12 \\ p21 & p22 \end{bmatrix} = \begin{bmatrix} 0.7I & + & 0.2S \\ 0.3I & + & 0.8S \end{bmatrix} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix} \begin{pmatrix} I \\ S \end{pmatrix}$$

$$P1 = \begin{pmatrix} p11 \\ p21 \end{pmatrix} = \begin{pmatrix} 0.7 \\ 0.3 \end{pmatrix} \qquad P2 = \begin{pmatrix} p21 \\ p22 \end{pmatrix} = \begin{pmatrix} 0.2 \\ 0.8 \end{pmatrix}$$

The matrix $\mathbf{P} = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}$ is our transition matrix. Note that, the entries of each column vectors are positive and their sum is 1.

Now that we have the transition matrix, we need a state vector ***X*** such that P.X = X, that is, an eigenvector of *P* associated to the eigenvalue 1. We need a particular state vector, namely the initial state vector. Our model considers a transceiver which is either in one of two states (Idle and Sleep State) but not both at the same time. The last state vector reflects that.

$$X_0 = \begin{pmatrix} \text{"Probaility of tranceiver in Idle State"} \\ \text{"Probaility of tranceiver in Sleep State"} \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$X_1 = P \cdot X_0$

$X_2 = P \cdot P \cdot X_0 = P^2 \cdot X_0$

$X_k = P^k \cdot X_0$

$$X_{20} = P^{20} \cdot X_0 = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}^{20} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$$

Therefore, after 20 transitions, we will find out that the transceiver spent only 40% of time in the Idle state.

What happens after 50 transitions?

$$X_{50} = P^{50} \cdot X_0 = \begin{pmatrix} 0.7 & 0.2 \\ 0.3 & 0.8 \end{pmatrix}^{50} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$$

We arrive at the same result. In other words, the state vector converged to a steady-state vector.
In this case that steady-state vector is
$Xss = \begin{pmatrix} 0.4 \\ 0.6 \end{pmatrix}$.

Irrespective of the starting state, eventually equilibrium must be achieved. The transient or sorting-out phase takes a different number of iterations for different transition matrices, but eventually the state vector features components that are precisely what the transition matrix calls for. So, subsequent applications of P do not change the matured state vector.

4 MODEL SIMULATION

To model our transition energy model requires the convergence of a 2x2 matrix of markov chain, the random walk in a pascal triangle and the use of skewness in our probability distribution function to to favour either the sleep or Idle States. To simulate the random walk through a Pascal triangle, the popular Plinko probability (Phet, 20013) , the Mathematica Wolfram demonstration suite for "The Skew Normal Density Function" (Coutinho, 2013) and "the probability machine" (Wes64, 2013) were used to demonstrate the various components of our model.

We simulated a Pascal triangle of eleven rows with the initial distribution function of 0.5, 0.78 and 0.2 using the plinko probaility tool . An average of 100 balls is allowed to roll down the triangle from the top vertex to the base of the triangle. The balls randomly walks (transition) through each vertex either going left (Sleep State) or right (Idle state). And the probability distribution of the walks at the end of the eleventh row is shown below the triangle.

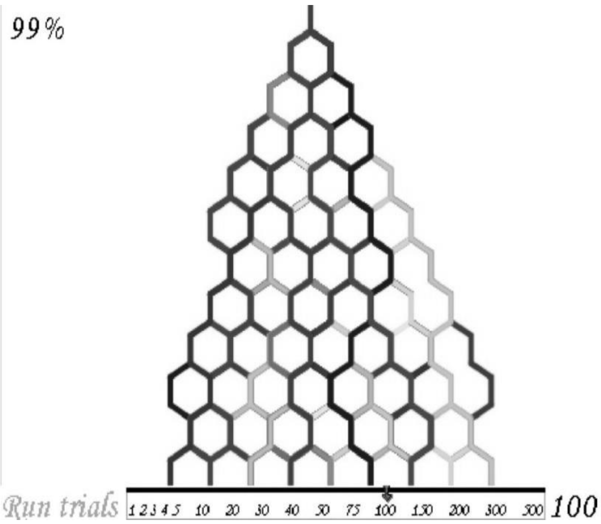

Figure 5: 100 trials-Random walk Path on Pascal triangle.

The probability machine captures this random walk of 100 balls choosing between the two states in each finite time increments with colourful walk paths (transition) as shown in **Figure 5**. The results of the experiment shows a normal distribution curve as shown by the histogram under the Pascal triangle in **Figure 6** when the probability of distribution is 0.5 and the Pascal triangle generated is shown in **Table 1.**

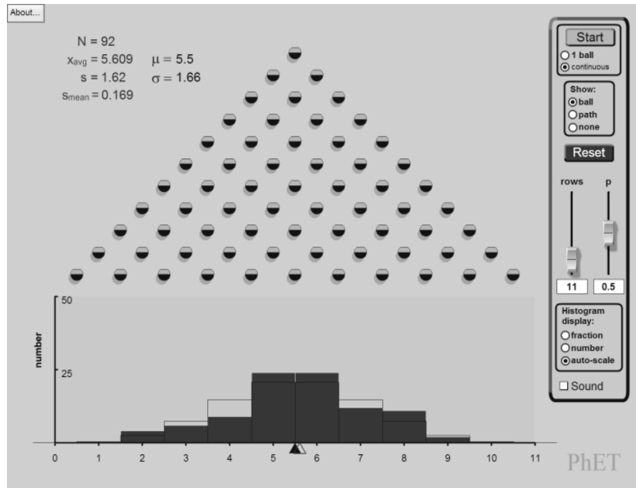

Figure 6: Pascal Triangle showing random walk paths for 0.5 initial distributions.

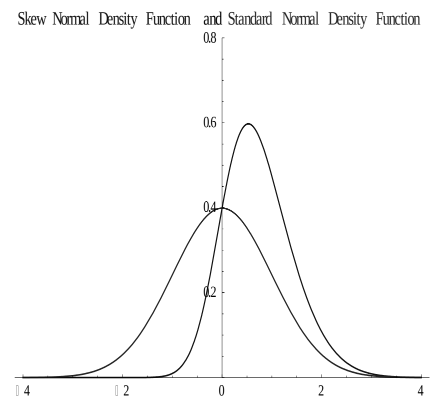

Figure 7: Random walk of a 2x2 matrix Probability distribution with 0.47 skewness.

Table 1: One dimensional walk through Pascal triangle.

Steps	−5	−4	−3	−2	−1	**0**	**1**	**2**	**3**	**4**	**5**
0						1					
1					$\frac{1}{2}$	0	$\frac{1}{2}$				
2				$\frac{1}{4}$	0	$\frac{2}{4}$	0	$\frac{1}{4}$			
3			$\frac{1}{8}$	0	$\frac{3}{8}$	0	$\frac{3}{8}$	0	$\frac{1}{8}$		
4		$\frac{1}{16}$	0	$\frac{4}{16}$	0	$\frac{6}{16}$	0	$\frac{4}{16}$	0	$\frac{1}{16}$	
5	$\frac{1}{32}$	0	$\frac{5}{32}$	0	$\frac{10}{32}$	0	$\frac{10}{32}$	0	$\frac{5}{32}$	0	$\frac{1}{32}$

In **Table 1** subsequent rows are found by adding half of each cell in a given row to each of the two cells diagonally below it. In fact, it is simply Pascal's triangle padded with intervening zeros and with each row multiplied by an additional factor of 0.5 [15].

Mean: $E(X) = \sqrt{\frac{2}{\pi}}\,\delta = 0.717$

Variance: $Var(X) = 1 - \frac{2\delta^2}{\pi} = 0.485$

Skewness: $\gamma_1 = \frac{(4-\pi)E(X)^3}{2\,Var(X)^{3/2}} = 0.469$

Kurtosis: $\gamma_2 = \frac{2(\pi-3)E(X)^4}{Var(X)^2} = 0.318$

where $\delta = \frac{c}{\sqrt{1+c^2}}$

Mean: $E(X) = \sqrt{\frac{2}{\pi}}\,\delta = -0.793$

Variance: $Var(X) = 1 - \frac{2\delta^2}{\pi} = 0.371$

Skewness: $\gamma_1 = \frac{(4-\pi)E(X)^3}{2\,Var(X)^{3/2}} = -0.945$

Kurtosis: $\gamma_2 = \frac{2(\pi-3)E(X)^4}{Var(X)^2} = 0.811$

where $\delta = \frac{c}{\sqrt{1+c^2}}$

Applying a skewness of 0.469 and -0.945 to the initial distribution produced the probability distributions shown in **Figure 7** and **Figure 8** respectively. This demonstrates how easily this energy model can be used to favour a transition to Sleep state or Idle state of the transceiver. Also in **Figure 7** and **Figure 8**, there appears a convergence point at point 4 on the graph irrespective of the whether the distribution is skewed or not. This demonstrate that our random walk will have points of convergence if modeled as a binomial distribution or have convergence or invariant / stationary probability distribution when modeled a markov chain of 2X2 matrix

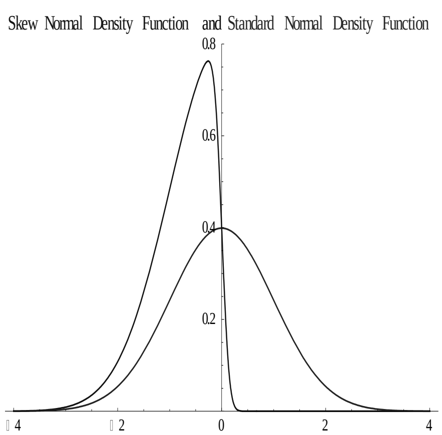

Figure 8: Random walk of a 2x2 matrix Probability distribution with -0.95 skewness.

5. CONCLUSION

This paper shed lights on how energy consumption can be maximized in the transceiver during state transitions. In this Paper we modeled the transition between the active and sleep states of the transceiver as a random walk of a 2X2 matrix elements on a weighted graph represented as the Pascal triangle. Irrespective of the starting state, we are able to arrive at a probability of consistent stationary distribution for the model. The usefulness of our model lies in the fact that a clearer picture of energy optimization in WSN deployment is presented, which enables transition energy optimization decisions to be taken as this model can skewed to favor a transceiver state of interest.

Details of the binomial distribution of state transitions and also the various properties and number patterns of the Pascal triangle as it relates to the transition energy consumption of the transceiver will be investigated in future research.

REFERENCES

Ahmad M.R, Eryk DutkiewiczE. (2009), 'A survey of low duty cycle MAC protocols in wireless sensor networks," Book Chapter in Wireless Sensor Network. Benini L. and De Micheli G., Dynamic Power

Benini L., Bogliolo A., and De Micheli, G.(2000) , "A survey of design techniques for system-level dynamic power management," IEEE Transactions on VLSI systems, June 2000, pp. 299-316.

Cameron P.J.(2003), “Markov chains and random walks” The Encyclopaedia of Design Theory.

Chung E.Y., Benini L., and Micheli G. D (1999). “Dynamic power management using adaptive learning tree.” In International Conference on Computer-Aided Design, pages 274-279, 1999.

Coppersmith D., Doyle, P., Raghavan p., Snir M.(1994), “ Random Walks on Weighted Graphs, and Applications to On-line Algorithms (Extended Abstract)” http://www.math.dartmouth.edu/~doyle/docs/catabs/catabs.pdf

Coutinho M.(2013), "The Skew Normal Density Function" from the Wolfram Demonstrations Project http://demonstrations.wolfram.com/TheSkewNormalDensityFunction/ Murilo Coutinho.

Douglis F., Krishnan P., and Bershad B.(1995), “ Adaptive Disk Spin-Down Policies for Mobile Computers. In Computing Systems, volume 8, pages 381-413.

Hwang C.H. and Wu A. C.(1997), “ A Predictive System Shutdown Method for Energy Saving of Event-Driven Computation”. In International Conference on Computer-Aided Design, pages 28-32, 1997.

Karlin A., Manesse M, McGeoch L and Owicki S(1994)., "Competitive Randomized Algorithms for Nonuniform Problems", Algorithmica, pp. 542-571, 1994.

Katz J. (2005)“Notes on Complexity Theory “Handout 8, http://www.cs.umd.edu/~jkatz/complexity/f05/lecture8.pdf Last updated: October, 2005.

Lovasz L. “Random Walks On Graphs(1993): A Survey” Bolyai Society Mathematical Studies, 2 Combinatorics, Paul Erd″Os Is Eighty (Volume 2) Keszthely (Hungary), 1993, Pp. 1–46.

Management :Design Techniques and CAD Tools, Kluwer Academic Publishers, 1997.

Odey J.A. and Li D.(2012), "Low Power Transceiver Design Parameters for Wireless Sensor Networks," Wirelesss Sensor Network, Vol. 4 No. 10, 2012, pp. 243-249. doi: 10.4236/wsn.2012.410035.

Paleologo G. A., Benini L., Bogliolo A., and Micheli G. D.(1998), “Policy Optimization for Dynamic Power Management.” In Design Automation Conference, pages 182-187, 1998.

Phet (2013), “Plinko Probability” http://phet.colorado.edu/sims/plinko-probability/plinko-probability_en.html. Accessed 16th February, 2013

Qiu Q., and M. Pedram M.(1999),” Dynamic Power Management Based on Continuous-Time Markov Decision Processes.” In Design Automation Conference, pages 555-561, 1999.

Wang W.S, O’Keeffe R, Wang N, Hayes M, O’Flynn B. and Mathuna S. C. O (2011)., “Practical wireless sensor networks power consumption metrics for building energy management applications,” in 23rd European Conference Forum Bauinformatik.

Weisstein, E. W(2013). "Random Walk--1-Dimensional." From MathWorld--A Wolfram Web Resource. http://mathworld.wolfram.com/RandomWalk1-Dimensional.html. Accessed 10th January, 2013.

Wes64, Scratch Projects, “Probability Machine” http://scratch.mit.edu/projects/Wes64/2463317