Fictitious Play Approach to a Mobile Unit Situation Awareness Problem

Theodore J. Lambert • Hua Wang

Department of Industrial and Operations Engineering, University of Michigan, 1205 Beal Avenue, Ann Arbor, Michigan 48109-2117, USA

Department of Electrical Engineering and Computer Science, University of Michigan, 1301 Beal Avenue, Ann Arbor, Michigan 48109-2122, USA

lambertt@umich.edu • huaw@umich.edu

Until now the fictitious play approach to optimization has only been demonstrated on a dynamic traffic routing problem; therefore, it is necessary to apply this method to other problems in order to demonstrate its effectiveness as a heuristic optimization method. We used the large scale situational awareness simulation developed for the Multidisciplinary University Research Initiative (MURI) on "Low-Energy Electronic Design for Mobile Platforms" to test the fictitious play approach, since we already possessed bench mark solutions from a simulated annealing approach previously applied. We found that the fictitious play approach yielded similar solutions to simulated annealing and required comparable computational effort, while they both outperformed pure random search. This demonstrates the effectiveness of a fictitious play approach to optimization for the large scale situational awareness simulation, providing additional evidence as to fictitious play's value as an optimization heuristic. (*Programming: Integer: Heuristic ; Simulation; Military*)

---

1. Introduction

While the concept of fictitious play has been around for some time (Brown 1951), its application to optimization problems has just begun. Garcia et al. (2000) applied the fictitious play approach to optimization on a large-scale dynamic traffic network and found favorable results. Lambert et al. (2002) formalized this approach, calling it *sampled fictitious play* and providing a rigorous theoretical foundation. Before any optimization heuristic can be considered successful it must be tested on many real world problems.

In this paper we apply sampled fictitious play to the large scale situation awareness simulation developed for the Multidisciplinary University Research Initiative (MURI) on "Low-Energy

Electronic Design for Mobile Platforms". Stark et al. (2002) approached this problem with simulated annealing, the results of which were used as a benchmark for the solutions found by sampled fictitious play. We also implemented pure random search in order to provide evidence supporting the need for these heuristic methods within this problem. We found that sampled fictitious play performed as well as simulated annealing, yielding similar results, while both simulated annealing and sampled fictitious play outperformed pure random search by as much as 25%. This provides evidence of the effectiveness of the sampled fictitious play algorithm as an optimization heuristic.

2. Fictitious Play

Let Γ be a finite common interest game in strategic form with the set of players $N = \{1, 2, \ldots, n\}$. We denote the finite set of strategies of player $i \in N$ by $\mathcal{Y}^i$, and let $\mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \cdots \times \mathcal{Y}^n$. Denote the *payoff function* by $u : \mathcal{Y} \to \mathbb{R}$, where $\mathbb{R}$ denotes the set of real numbers.

For $i \in N$, let Δ^i be the set of mixed strategies of player i. That is,

$$\Delta^i = \left\{ f^i : \mathcal{Y}^i \to [0,1] : \sum_{y^i \in \mathcal{Y}^i} f^i(y^i) = 1 \right\}.$$

Each $f^i \in \Delta^i$ can be viewed as an assignment of probabilities, or *beliefs*, to the elements of $\mathcal{Y}^i$; in particular, with a slight abuse of notation we identify the pure strategy $y^i \in \mathcal{Y}^i$ with the extreme point of Δ^i which assigns a probability 1 to y^i. Set $\Delta = \Delta^1 \times \Delta^2 \times \cdots \Delta^n$.

We extend u to be the payoff function in the mixed extension of Γ. That is, for any $f \in \Delta$,

$$u(f) = u(f^1, f^2, \ldots, f^n) = \sum_{y \in \mathcal{Y}} u(y^1, y^2, \ldots, y^n) f^1(y^1) f^2(y^2) \cdots f^n(y^n). \tag{1}$$

Note that we have assumed players choose their strategies independently.

Let $g \in \Delta$, and let $\varepsilon \geq 0$. We say that g is an *ε-equilibrium* if for each $i \in N$

$$u^i(g) \geq u^i(f^i, g^{-i}) - \varepsilon \text{ for all } f^i \in \Delta^i,$$

where $(f^i, g^{-i}) = (g^1, \ldots, g^{i-1}, f^i, g^{i+1}, \ldots, g^n)$. A *Nash equilibrium* is a 0-equilibrium, and will be simply referred to as an *equilibrium*.

Denote by K the set of all equilibria of Γ, and denote by $\|\cdot\|$ the Euclidean norm on the Euclidean space that may be viewed as containing Δ. For $\delta > 0$ set

$$B_\delta(K) = \{g \in \Delta : \min_{f \in K} \|g - f\| < \delta\}.$$

A *belief path* is a sequence $(f(t))_{t=1}^{\infty}$ in Δ. We say that the belief path $(f(t))_{t=1}^{\infty}$ *converges to equilibrium* if each accumulation point of $(f(t))_{t=1}^{\infty}$ is an equilibrium point; that is, if for every $\delta > 0$ there exists an integer T such that $f(t) \in B_{\delta}(K)$ for all $t \geq T$. In other words, a belief path that converges to equilibrium is eventually arbitrarily close to some equilibrium of Γ.

A *path* in $\mathcal{Y}$ is a sequence $(y(t))_{t=1}^{\infty}$ of elements of $\mathcal{Y}$. To each path $(y(t))_{t=1}^{\infty}$ we naturally associate a belief path $(f_y(t))_{t=1}^{\infty}$ by letting

$$f_y(t) = \frac{1}{t}\sum_{s=1}^{t} y(s) \text{ for every } t \geq 1.$$

In the previous equation, the $y(s)$ should be viewed as elements of Δ.

We now formally define a fictitious play process. For $i \in N$ and for $f \in \Delta$, let

$$v^i(f) = \max\{u^i(g^i, f^{-i}) : g^i \in \Delta^i\}.$$

That is, $v^i(f)$ is the value of player i's best response to the other players' strategies f^{-i}. Notice from the definition of $u^i(f)$, that $v^i(f)$ can always be attained by an extreme point of Δ, i.e., $\max\{u^i(g^i, f^{-i}) : g^i \in \Delta^i\} = \max\{u^i(y^i, f^{-i}) : y^i \in \mathcal{Y}^i\}$. A path $(y(t))_{t=1}^{\infty}$ is a *fictitious play process* if for every $i \in N$,

$$u^i(y^i(t+1), f_y^{-i}(t)) = v^i(f_y(t)) \text{ for every } t \geq 1. \tag{2}$$

Notice that, as defined by (2), $y^i(t+1)$ is a best response of player i to the mixed strategies of the other players, as represented by the beliefs $f_y^{-i}(t)$.

We define the function $\bar{U}_k^i(\cdot, f_y^{-i}(t)) : \mathcal{Y}^i \to \mathbb{R}$ by

$$\bar{U}_k^i(y^i, f_y^{-i}(t)) = \sum_{j=1}^{k} \frac{u^i(y^i, Y_j^{-i}(t))}{k} \tag{3}$$

where the $Y_j^{-i}(t)$ are iid random vectors drawn from the distribution given by $f_y^{-i}(t)$. Then $\bar{U}_k^i(y^i, f_y^{-i}(t))$ is a sample mean (with sample size k) of player i's utility when playing or using y^i. Let $\bar{u}_{k_t}^i(y^i, f_y^{-i}(t))$ denote the realization of $\bar{U}_{k_t}^i(y^i, f_y^{-i}(t))$. If the "best response" of each player is chosen based on sample means instead of the actual means, i.e., $y^i(t+1)$ is chosen so that $y^i(t+1) \in \operatorname{argmax}\{\bar{u}_{k_t}^i(y^i, f_y^{-i}(t)) : y^i \in \mathcal{Y}^i\}$ for some $k_t \in \{1, 2, \ldots\}$, we will call the stochastic process $(y(t))_{t=1}^{\infty}$ a *sampled fictitious play process*.

Sampled Fictitious Play Algorithm

Initialization: Set $t = 1$ and select $y(1) \in \mathcal{Y} = \mathcal{Y}^1 \times \mathcal{Y}^2 \times \ldots \times \mathcal{Y}^n$ arbitrarily; set $f_y(1) = y(1)$.

Iteration $t \geq 1$: Given $f_y(t)$, find

$$y^i(t+1) \in \underset{y^i \in \mathcal{Y}^i}{\operatorname{argmax}} \{\bar{u}^i_{k_t}(y^i, f_y^{-i}(t))\},\ i = 1, \ldots, n, \tag{4}$$

where $\bar{u}^i_{k_t}(y^i, f_y^{-i}(t))$ is the realization of $\bar{U}^i_{k_t}(y^i, f_y^{-i}(t))$ as defined by (3). Set $f_y(t+1) = f_y(t) + \frac{1}{t+1}(y(t+1) - f_y(t))$, increment t by 1 .

The following theorem guarantees that the previous algorithm will generate a Nash equilibrium of the game.

Theorem 1 *Let Γ be a finite game in strategic form with identical payoff functions. Then every sampled fictitious play process $y(t)$ with sample sizes $k_t = \lceil Ct^\beta \rceil$ for $\beta > \frac{1}{2}$ and $C > 0$, will converge in beliefs to equilibrium with probability 1.*

Proof: See Lambert et al. (2002).

By applying the sampled fictitious play algorithm we will generate a Nash Equilibrium of the game, this being our surrogate for the optimal solution to the optimization problem.

3. Situation Awareness Problem Description

In the Multidisciplinary University Research Initiative (MURI) on "Low-Energy Electronic Design for Mobile Platforms," we try to solve a situational awareness problem. In this problem, a number of mobile nodes desire to keep track of the location of each other over some time duration. The nodes operate with batteries and thus have a finite energy constraint. The transmission of information by a node requires a certain amount of energy, as does the processing of any received signal. The goal of the design is to minimize the mean absolute error of the position estimates. There is a plethora of parameters that could be considered for optimization, but in order to develop a systematic and computationally tractable design methodology, we divide the problem into interacting design layers, namely, device layer, processing layer, and network layer as illustrated in Figure 3, and perform the optimization over a small set of parameters.

3.1 Device Layer

At the device layer, we assume each node has an omni-directional dipole antenna and a small power amplifier. We capture the operation of the amplifier and the coupling among the device layer and other higher layers through three parameters: the total consumed power P_{total}, the output power P_{out}, and the AM-to-AM voltage characteristics (Borich et al. 1998). We characterize the relation between the average amplifier output power and the energy constraint E_{ct} for transmitting a packet by

$$P_{out} = g_1(E_{ct}). \tag{5}$$

This relation is tabulated for use by higher layers. In certain situations it is possible that the actual consumed energy at the transmitter, E_{ta}, is less than the constraint on the consumed energy at the transmitter. In this case we define a function

$$E_{ta} = g_2(E_{ct}) \tag{6}$$

that maps the energy constraint to the actual energy.

3.2 Processing Layer

Figure 1 shows the basic block diagram of the processing layer. The channel encoder adds redundancy to a block of input information to protect it from channel errors. The output of the channel encoder is interleaved, modulated, and spread in bandwidth. The resulting signal is amplified by a power amplifier (PA) and transmitted. At the receiver the inverse operations are performed to recover the block of information. Even though each of these operations consumes power, we focus on the energy being consumed by the power amplifier, the demodulator, and the channel decoder because these elements consume much more energy than other elements in the system. We have covered the performance-energy tradeoff of the amplifier in Section 3.1; therefore, we emphasize the tradeoff between the demodulator and decoder in detail here.

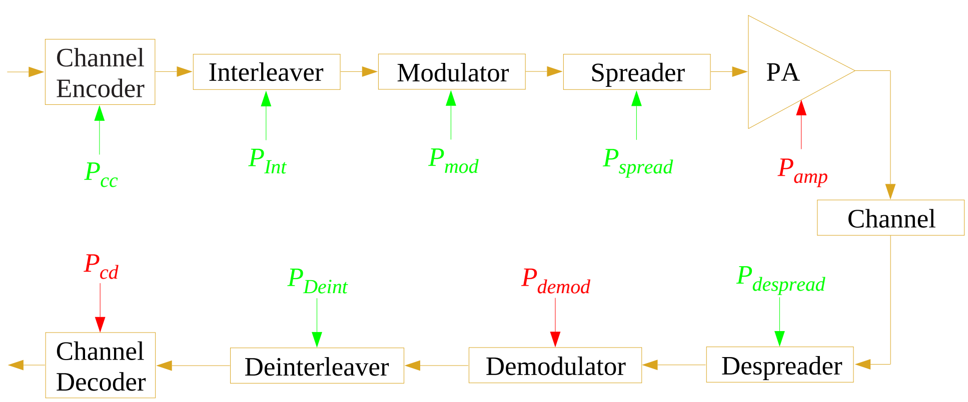


Figure 1: Processing layer block diagram

The amount of energy consumed while performing coding and modulation operations depends on the number of bits used to represent the data. The larger the number of bits used, the better the performance. However, the more bits used in the representation, the more energy consumed by the demodulator and decoder (Rabaey 1995). We let N_E denote the number of quantization bits used in the demodulator for data and coefficients, and N_D denote the number of quantization bits used in the decoder.

The performance measure that couples the processing layer with the network layer is the packet error probability, P_e. In general, P_e depends on the energy constraint E_{ct} for the transmitter to send a packet, the energy constraint E_{cr} for the receiver to process a packet, the received signal-to-noise ratio SNR, the number of bits of quantization used in the demodulator N_E, and the number bits of quantization used in the decoder N_D. Since N_E and N_D affect only the performance of the processing layer, we locally optimize P_e with respect to N_E and N_D for given E_{ct}, E_{cr}, and SNR

$$P_e = \min_{N_D, N_E} f(N_D, N_E, SNR, E_{ct}, E_{cr}) = g_3(SNR, E_{ct}, E_{cr}). \tag{7}$$

We therefore generate a parameterized version of P_e with respect to E_{ct}, E_{cr}, and SNR, and build a performance table for these parameterized versions of P_e. In addition, in order to calculate the actual energy needed to demodulate and decode signals, we model each individual algorithm using digital circuits (Hong et al. 1999, 2000). Because of the integer constraint on quantization bits used in the demodulator and decoder, the actual energy consumed by the receiver, E_{ra}, may be less than the constraint on energy E_{cr}. Let $N_E^*(SNR, E_{ct}, E_{cr})$ and $N_D^*(SNR, E_{ct}, E_{cr})$ be the optimum number of quantization bits in the demodulator and decoder respectively. Thus the actual energy consumed by the receiver is a function of the constraint on the energy and signal-to-noise ratio,

$$E_{ra} = g(N_E^*, N_D^*) = g_4(SNR, E_{ct}, E_{cr}). \tag{8}$$

The network layer (and global optimization) utilizes the table of P_e as a function of E_{ct}, E_{cr}, and SNR for calculating its own global performance.

3.3 Network Layer

We consider a network of nine nodes moving according to a specific mobility model. Each node attempts to keep track of the positions of all the other nodes by means of communication and estimation. We present the mobility models, the propagation models, the communication protocols, and the estimation schemes used by the nodes.

For the mobility model, we consider a region of size 6 km $\times$ 6 km and a group of nine nodes initially deployed in a zone as shown in Figure 2. Each node in the network moves to a new location at the end of every T_m seconds, where $T_m = 1$. All nodes travel at average speed v m/s, where $v = 1$, toward the common destination located at $\underline{G} = (6000, 6000)$ m. At each step, each node's motion is subject to a random disturbance in x and y coordinates.

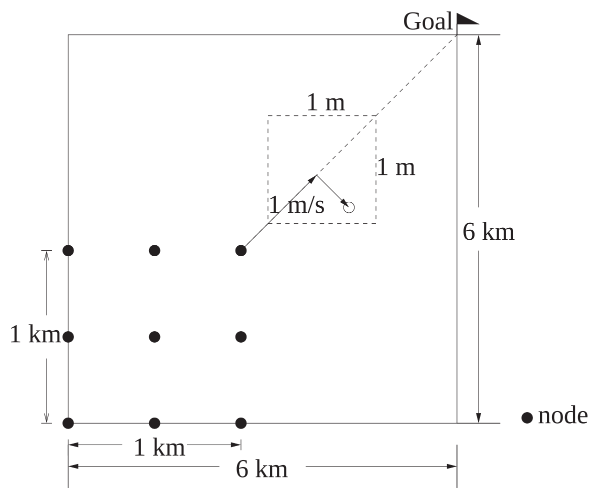


Figure 2: Mobility Model

The transmitted signal from each node experiences propagation loss and fading. We assume a two-path propagation model from the transmitter to the receiver, which consists of a direct path and a path reflected off the ground with 180 degree phase change at the reflection point. The cumulative effect of this model resulted in an attenuation A between received power and transmitted power, which is usually proportional to the fourth power of the distance between the transmitter and the receiver.

The transmission protocol is that each node transmits its position information packets every T seconds, where T is a design parameter. The medium access control is Time Division Multiple Access (TDMA), where each node is assigned a transmission slot of duration T/N, where $N = 9$ in our case. The slot duration is much larger than a packet duration. In a given slot, each packet transmission is followed, with probability q, by a retransmission, and so forth, until the slot ends. The retransmission probability q is considered as a design parameter because more complex automatic retransmission request (ARQ) schemes are not well-suited to the broadcast environment under consideration. The energy used for each packet transmission or retransmission is upper bounded by E_{ct}. The packet may be received by many other nodes, each of which consumes a certain amount of energy to process the packet, which is upper bounded by E_{cr}. When a node receives a packet, it does not send back any acknowledgment, nor does it forward the packet it receives to other nodes. As a consequence, every packet in the transmission protocol travels only one hop. In summary, we choose T, q, E_{ct}, and E_{cr} as the design parameters at the network layer that affect global performance.

Each node in the network estimates the other nodes' positions every T_e seconds, where $T_e = 2$. Since according to the mobility model the nodes move toward the goal in a straight line subject to noise, the new estimate is the extrapolation toward the goal of the position contained in the packet that was last received correctly, by an amount proportional to the product of velocity and time.

The estimation error of node j's position made by node i at time kT_e is defined as

$$\underline{e}_k^{(i,j)} = \underline{w}_k^{(j)} - \hat{\underline{w}}_k^{(i,j)}, \tag{9}$$

where $\underline{w}_k^{(j)}$ is the actual position of node j at time kT_e, and $\hat{\underline{w}}_k^{(i,j)}$ is the estimate of node i on the position of node j at time kT_e. For the purpose of optimization, we use mean absolute error as the performance metric,

$$J^{(i)} = E\left[\frac{1}{K(I-1)} \sum_{j=1, j\neq i}^{I} \sum_{k=1}^{K} \left\| \underline{e}_k^{(i,j)} \right\| \right], \tag{10}$$

where KT_e is the time horizon under consideration. In the above equation, the expectation is with respect to the mobility, the noise in the receiver, and the randomness in retransmission. The overall network performance measure is given by the average of the position estimation error contributed by all the nodes in the network:

$$J = \frac{1}{I} \sum_{i=1}^{I} J^{(i)}. \tag{11}$$

The goal is to minimize J over the parameters that affect global performance subject to a constraint on the energy used by each node. Let $E^{(i)}$ denote the energy used by node i over the time horizon KT_e. The constraint on energy is

$$\max_{1\leq i\leq I} E^{(i)} \leq E. \tag{12}$$

The objective is to determine the design parameters

$$[T^*, q^*, E_{ct}^*, E_{cr}^*] = \arg \min_{\substack{[T,q,E_{ct},E_{cr}] \\ \max E^{(i)} \leq E}} J(T, q, E_{ct}, E_{cr}) \tag{13}$$

and the corresponding performance $J^* = J(T^*, q^*, E_{ct}^*, E_{cr}^*)$.

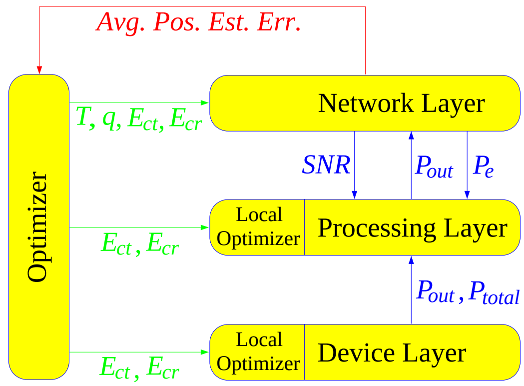


Figure 3: Coupling of different layers

4. Optimization Procedure

The parameters that describe the coupling among the layers are shown in Figure 3.

The optimization is, in part, simulation-based because we do not possess precise analytical expressions for the local and global optimization criteria we employ. The optimization program attempts to find the global minimum of the objective function J in equation (11).

The global optimization and simulation modules perform the following steps in attempting to find the globally optimal solution:

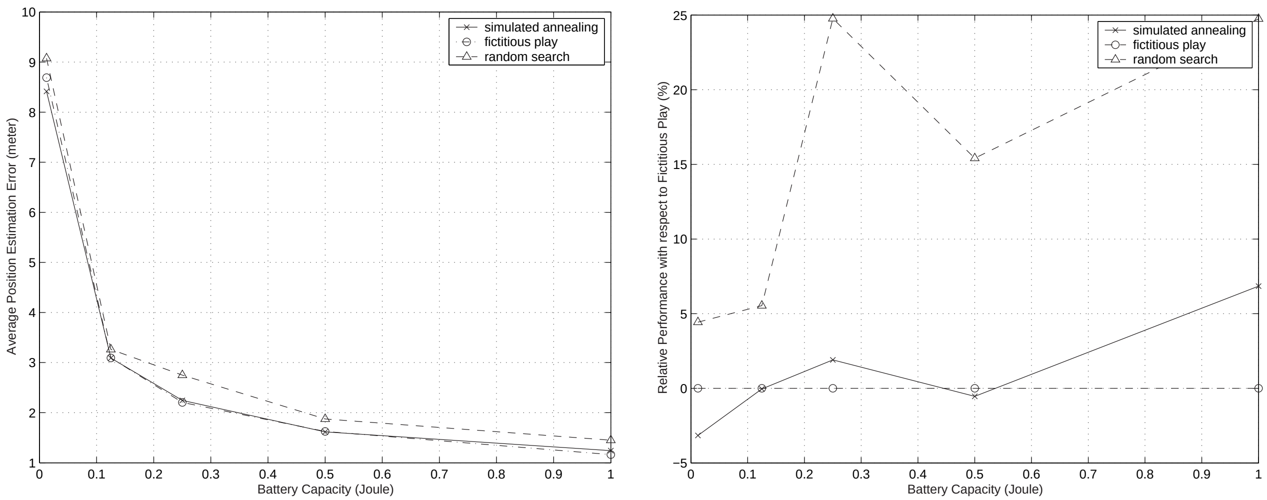
Step 1. The "optimizer" module determines the (new) parameters $[T, q, E_{ct}, E_{cr}]$, for which the network performance is to be evaluated.

Step 2. The "network simulator" module approximates the objective function in (11) for the given $[T, q, E_{ct}, E_{cr}]$ using Monte-Carlo simulation techniques. It returns the average position estimation error to the "optimizer" module.

Step 3. Steps 1 and 2 are repeated until a terminating condition is reached.

The "optimizer" module used in Step 1 is the sampled fictitious play algorithm, which has been discussed in Section 2. In Step 2, we implement the "network simulator" module in *OPNET*, a widely used network development and analysis tool (OPNET 2000). For the given parameters and interacting variables, the "network simulator" calculates the objective function, i.e., average position estimation error, through network simulation. As mentioned in Sections 3.1 and 3.2, the performance of the processing layer and device layer and interacting parameters has been tabulated offline so that the network layer can use them as function calls. In Step 3, the termination condition that we chose for our experiments was to stop after 30 iterations.

5. Results



(a) Normal Scale

(b) Percentage

Figure 4: Comparison of optimization results

Figure 4 is a comparison of the results found using sampled fictitious play, with $C = .01$ and $\beta = .51$, with the results found using simulated annealing and using random search. From the results, we can see the performance of simulated annealing and fictitious play algorithms are close, and they both outperform the random search algorithm by as much as 4-25%. The low percentage improvement occurs when the battery capacity is low, where the performance generally has a large variance. These results provide additional support to the usefulness of fictitious play as an optimization heuristic.

Acknowledgements

This research is supported in part by the National Science Foundation under Grants DMI-9713723, DMI-9900267, and DMI-0217283; the Department of Defense Research & Engineering (DDR&E) Multidisciplinary University Research Initiative (MURI) on “Low Energy Electronics Design for Mobile Platforms” and managed by the Army Research Office (ARO) under grant ARO DAAH04-96-1-0377; by the Department of Defense Augmentation Awards for Science and Engineering Research Training (ASSERT) on ”Optimization Algorithms for Low Power Mobile Platforms” and managed by the Army Research Office under grant ARO DAAG55-98-1-0155; and by the Horace H. Rackham School of Graduate Studies Faculty Grant.

References

Borich, V., J.H. Jong, J. East, W. Stark, 1998. Nonlinear Effects of Power Amplification on Multi-carrier Spread Spectrum Systems. *International Microwave Symposium* **1** 323-326.

Brown, G, 1951. Iterative Solution of games by Fictitious Play. *In Activity Analysis of Production and Allocation, Wiley*. 374-376

Garcia, A, D. Reaume, R.L. Smith, 2000. Fictitious Play for Finding System Optimal Routings in Dynamic Traffic Networks.*Transportation Research B* **34** 146-157.

Hong, S., W.E. Stark, 1999. Power Consumption vs. Decoding Performance Relationship of VLSI Decoders for Low-Energy Wireless Communication System Design. *Proceedings of IEEE International Conference on Electronics, Circuits and Systems*. 1593-1596.

Hong, S., W.E. Stark, 2000. Design and Implementation of a Low Complexity VLSI Turbo-Code Decoder Architecture for Low Energy Mobile Wireless Communications. *Journal of VLSI Signal Processing*. February, 2000. 43-57,

Lambert, T.J., M.A. Epelman, R.L. Smith, 2002. A Fictitious Play Approach To Large-Scale Optimization. Technical Report, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, Michigan, USA.

OPNET Technologies, 2000. *OPNET Modeler Manual*.

Rabaey, J.M.M., 1995. *Digital Integrated Circuits: A Design Perspective*. Prentice Hall PTR.

Stark W.E. , H. Wang, A. Worthen, S. Lafortune, D. Teneketzis. 2002. Low-Energy Wireless Communication Network Design. *IEEE Wireless Communications*. August, 2002. 60-72