Coordinated Path Planning for a UAV Cluster

Jan De Mot Vishwesh Kulkarni Sommer Gentry Vladislav Gavrilets Eric Feron

Abstract— **Open problems in coordinated target acquisition by a UAV cluster are addressed. Original results on the spatial separation of the UAVs under an optimal policy are presented. Extensions to cover optimal policies for *pre-specified* UAV spatial separations are outlined. These results help determine the number and spatial distribution of the UAVs needed to perform a target acquisition task efficiently. Analytical and simulation results are presented.** ■

Keywords— **unmanned aerial vehicles (UAVs), dynamic programming.**

I. Introduction

Distributed self-organizing vehicle networks, such as a cluster of Unmanned Aerial Vehicles (UAVs), are posing exciting navigation and coordination questions (see [9], [1], [7], [5], and [8]). Several key issues must be resolved to ensure that a deployed cluster performs its task efficiently. We investigate the following two:

1. *number of the UAVs* needed to maximize the efficiency.
2. *coordination policy* needed to maximize the efficiency.

We address these issues in a *coordinated target acquisition* setting wherein the cluster is to acquire a set of targets in a partially known, uncertain dynamic environment. In these operations, some *a priori* terrain information is available and on-board sensor observations are used to reduce its uncertainty. We are interested in situations where these observations are direction-dependent. It is assumed that the UAVs share this information and coordinate accordingly. Real world communication (see [12] and [13]) imposes restrictions on how far, or how close, the UAVs can be from one another. So, it is vital to characterize:

1. separation of the UAVs under an optimal policy, and
2. optimal policy when the separations are given *a priori*.

Novelty of our work is in answering these unexplored questions.

II. Terminology and Notation

We map the operation terrain into *regions* and associate each of those with a *payoff*, which is the difference between its vantage value and its associated exposure threat. The terrain traversal is mapped onto a graph traversal with a node on the graph denoting a region. The cost to go from a node n_1 to a node n_2 is the summation of the actual cost to go from the region represented by n_1 to the region represented by n_2, and the payoff of the region represented by n_2. For example, if n_2 denotes a particularly hazardous region then the link cost is taken to be very high even if the actual cost to go is very low.

The authors are with the Massachusetts Institute of Technology, Cambridge, MA. Emails: jdemot@mit.edu, vishwesh@mit.edu, sommerg@mit.edu, gavrick@mit.edu, feron@mit.edu.

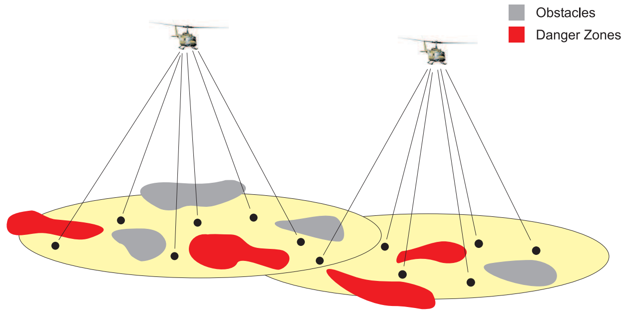

Fig. 1. Coordinated target acquisition application employing UAVs. The UAVs supervise ground vehicles that traverse a terrain full of obstacles and danger zones. The number and spatial distribution of the UAVs, and the coordination policy are important issues.

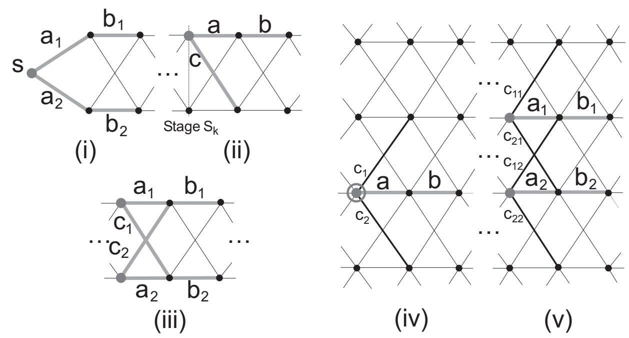

Fig. 2. The lookahead link terminology. Cases (i) and (ii) illustrate the nomenclature for the source node and for an arbitrary node, respectively, in a finite stage $\mathcal{G}_2$ when two agents are on the same node. Case (iii) illustrates the nomenclature for an arbitrary node in a finite stage $\mathcal{G}_2$ given two agents on different nodes. Cases (iv) and (v) illustrate the nomenclature for an arbitrary node in an infinite stage $\mathcal{G}_m$ when the agents are together and when they are separate, respectively.

The transition graph $\mathcal{G}_m$ of interest has a source s and a destination t connected through a possibly infinite number of stages S_i having m nodes each; at each stage, the first node is connected to the last node only once. Every node has a favorable lookahead direction having a two-link lookahead with link cost to the next stage denoted a and the link cost from the next stage to its next stage denoted b. Every node has an unfavorable lookahead direction offering a lookahead of at most a single link, its cost denoted c (see Figure 2 for illustrated problem specific details). The only exceptional nodes are the nodes at the last stage, which have only one path available to the target, which offers a single link lookahead, and the source node s for which every direction has a lookahead of two links. If two agents are on the same stage, then we say that they are in *configuration* $\mathcal{C}_0$ if they are on the same node, in $\mathcal{C}_1$ if they are on

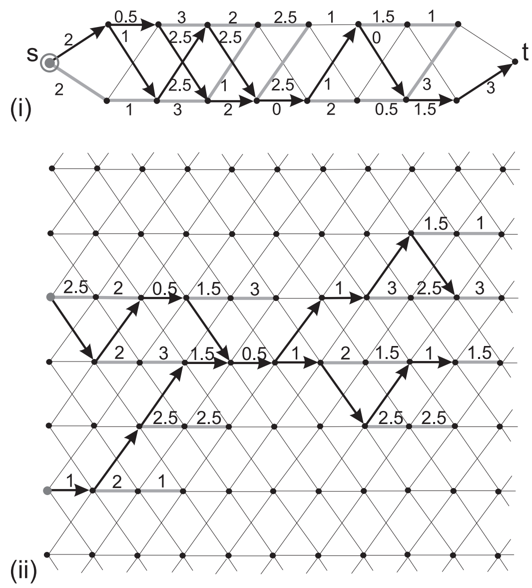

Fig. 3. (i) and (ii) illustrate some numerical results for Problem 1 and Problem 2, respectively. The bold arrows show the path of progression and the medium dark lines show the lookahead link costs at the traversed nodes. In the first stage of (i), the agents split with one of them choosing a higher cost link, which illustrates that an optimal policy encourages some UAV separation.

adjacent nodes and in $\mathcal{C}_{\ell+1}$ ($\ell \geq 1$) if the number of nodes in between them equals ℓ.

III. Problem Formulation

Problem 1: Let a cluster of two agents be stationed at the node s in $\mathcal{G}_2$. Let each of the agents have a lookahead of two links in the favorable direction and a lookahead of one link in the unfavorable direction. Then, find the path for each agent from s to t such that the expected aggregate cost is minimized. ∎

Problem 2: Let a cluster of two agents be stationed at some stage in an infinite stage $\mathcal{G}_m$. Let each of the agents have a two link lookahead in the favorable direction and a zero link lookahead in the unfavorable direction. The cost information available about only the lookahead links is deterministic and the costs of the rest of the links are independent uniformly distributed random variables. Then, find the path for each agent such that the expected aggregate cost is minimized. ∎

IV. Main Results

We implemented dynamic programming solutions to Problem 1 and Problem 2 (see Appendix) in MATLAB. An eight stage network was used for a numerical study of the Problem 1 solution. In both the cases, the unknown link costs were assumed to be independent and uniformly distributed over $\{0, 0.5, \ldots, 3\}$.

Efficiency of the target acquisition increases with the number of UAVs employed, as shown in Table 1. These efficiency results are for the infinite horizon case of $\mathcal{G}_7$ with

TABLE I

UAV Clusters and Incurred Costs

Number of Vehicles	Policy	Cost per Vehicle
1	Random Motion	7.50
1	Optimal Policy	5.19
2	Optimal Policy	4.20

the discount factor α taken to be 0.8. Four different initial position configurations are possible, viz. $\mathcal{C}_0$ through $\mathcal{C}_3$. The average cost was computed by choosing the initial position configurations using a uniform probability distribution.

Simulation results for a set of link cost scenarios are shown in Figure 3. They indicate that, for the topology considered, a cluster of two agents does not prefer to proceed towards the target with more than a single node separation. Indeed, we have the following result.

Lemma 1: [Cluster Separation Lemma]
Given a two agent cluster in $\mathcal{G}_m$ ($m \leq 7$) using optimal paths derived as the solution to Problem 2, configurations $\mathcal{C}_0, \mathcal{C}_1$ and $\mathcal{C}_2$ do not evolve into configurations $\mathcal{C}_\ell$ having $\ell > 2$. Furthermore, the configuration $\mathcal{C}_3$ always evolves into either $\mathcal{C}_2$ or $\mathcal{C}_1$ unless it is optimal for the agents to choose links a_1 and a_2. □

Proof: See Appendix.

Thus once the agents are in $\mathcal{C}_2$, they do not separate any further. Intuitively, adding another node in each stage should not cause the separation to grow. Therefore, we believe that Lemma 1 is strong enough to hold for any $\mathcal{G}_m$ and conjecture the following result.

Conjecture 1: [Cluster Separation Conjecture]
Given a two agent cluster in $\mathcal{G}_m$ using optimal paths derived as the solution to Problem 2, configurations $\mathcal{C}_0, \mathcal{C}_1$ and $\mathcal{C}_2$ do not evolve into configurations $\mathcal{C}_\ell$ having $\ell > 2$. Furthermore, a configuration $\mathcal{C}_\ell$, where $\ell > 2$, does not evolve into a configuration $\mathcal{C}_\nu$ such that $\nu > \ell$.

V. Discussion

Lemma 1 says that using an optimal policy, the cluster agents do not separate widely in acquiring the targets and gives bounds on the separation. The agents derive advantage from staying together because only thereby can an agent exploit the observations of another. Physical communication constraints often specify that the UAVs stay within a certain range. Optimal policies subject to this constraint can be obtained by simply penalizing the UAV separations in the cost-to-go functions. Preliminary work suggests that these findings apply to N-agent clusters as well.

The coordinated target acquisition problem can be viewed as some special instances of finite time partially observable *Markov Decision Processes* (MDPs) with the process state comprising the link costs from the current agent location(s) to the next stage and all the link costs from that stage to its next stage (see [11], [3], and [10]). However, it seems so far only Boutilier [4] has outlined extensions of standard MDPs to multi-agent coordination problems. The partially observable setting makes our work more broadly applicable than those extensions.

VI. Conclusion and Future Directions

We have solved the coordinated path planning problem for a two-vehicle cluster acquiring a target. Given an N agent cluster, our dynamic program has $O(N)$ state variables; its memory requirements and computational complexity grow exponentially with the number of states. Hence to address the case of clusters having a large number of agents, it is necessary to resort to neurodynamic programming [2] and go for approximate dynamic programming solutions rather than the exact ones.

Our separation and coordination policy results will be useful for UAV cluster networks in deciding which information is to be communicated, as well as how it is to be communicated.

Many situations demand that the cluster has a cluster-head agent which acts as a router for the rest of the agents. Some situations require that the agents take turns in becoming a cluster-head [6] and the rotation comes with an energy overhead. In these cases, positioning of the cluster-head is important, as is a characterization of the bounds on the number and the spatial distribution of the agents that can be supervised efficiently. We can readily account for these cases by incorporating a different link cost for the cluster-head agents since they burn more energy, and by adding a cost overhead when the rotation takes place. Lemma 1 and its generalizations are extremely useful in this context.

VII. Appendix

A. Solution to Problem 1

Let $J^*_{T,k}(a,c,b)$ denote the expected aggregate cost incurred in reaching t from S_k if the platoon members are at the same node of S_k with a, c, and b being the lookahead link costs. For symmetry reasons, $J^*_{T,k}$ is independent of node location in the stage S_k. Let $J^*_{S,k}(a_1,a_2,c_1,c_2,b_1,b_2)$ denote the expected aggregate cost incurred in reaching the target t from S_k if the platoon members are at different nodes of S_k with a_1, a_2, c_1, c_2, b_1, and b_2 being the lookahead costs. Let $I_T \doteq (a,c,b)$, $I_S \doteq (a_1,a_2,c_1,c_2,b_1,b_2)$, $I_0 \doteq (a_1,a_2,b_1,b_2)$. Let $\nu_S \doteq (\zeta_2,\phi_1,\phi_2,\psi_1,\psi_2)$. Then, the minimal aggregate expected cost $J^*_0(a_1,a_2,b_1,b_2)$ to reach t from s is computed by the following finite horizon reverse dynamic program:

$$
\begin{aligned}
J^*_{T,N}(s) &= 2a\\
J^*_{S,N}(a_1,a_2) &= a_1+a_2\\
J^*_{T,k}(I_T) &= \min\{f^T_{aa,k}(I_T), f^T_{ac,k}(I_T), f^T_{cc,k}(I_T)\}\\
J^*_{S,k}(I_S) &= \min\{f^S_{aa,k}(I_S), f^S_{cc,k}(I_S),\\
&\qquad f^S_{ac,k}(I_S), f^S_{ca,k}(I_S)\}\\
&\qquad \ldots 0\le k\le N-1\\
J^*_0(I_0) &= \min\{f_{a_1a_1,0}(I_0), f_{a_2a_2,0}(I_0), f_{a_1a_2,0}(I_0)\}\\
f^T_{ss,k}(a,c,b) &\doteq 2a+E_{\phi,\psi}[J^*_{T,k+1}(b,\phi,\psi)|b]\\
f^T_{sc,k}(a,c,b) &\doteq a+c+E_{\nu_S}[J^*_{S,k+1}(b,\nu_S)|b]\\
f^T_{cc,k}(a,c,b) &\doteq 2a+E_{\zeta,\phi,\psi}[J^*_{T,k+1}(\zeta,\phi,\psi)]
\end{aligned}
$$

Here, $f^\gamma_{\alpha\beta,k}$ with $\gamma\in\{S,T\}$, $\alpha,\beta\in\{a,c\}$ and $1\le k\le N-1$ denote the expected aggregate cost-to-go from S_k to t when the platoon is in configuration γ, the decision at S_k is to move to S_{k+1} using α and β and using an optimal policy from then onwards. A similar definition holds for $f_{a_1a_1,0}$, $f_{a_2a_2,0}$ and $f_{a_1a_2,0}$. The optimal agent policy is then determined as follows. The value functions $J^*_{T,k}(\cdot)$ and $J^*_{S,k}(\cdot)$ are precomputed for each node in the state transition graph. After arriving at a node, the agents choose the link based on the lookahead link cost and the value function at the downstream node following that particular link.

B. Solution to Problem 2

Let $J^*_T(a,b)$ denote the expected aggregate discounted cost incurred in moving infinitely in the direction of t ($N=\infty$), given that the platoon members are at the same node of the current stage and given that a and b are the costs of the links in the lookahead region as shown in Figure 3. Let $J^*_{S,j}(a_1,a_2,b_1,b_2)$ denote the expected aggregate discounted cost incurred in moving infinitely in the direction of t, given that the platoon members are separated by j nodes with $0\le j\le\lfloor\frac{m}{2}\rfloor-1$ where the lookahead link costs are a_1, a_2, b_1, and b_2 as shown in Figure 3. Let

$$
\mathcal{L}\doteq\{c_1c_1,c_1c_2,aa,ac_2,c_2c_2,c_1a\}
$$
$$
\mathcal{L}_1\doteq\{c_{11},a_1,c_{12}\},\quad \mathcal{L}_2\doteq\{c_{21},a_2,c_{22}\}.
$$

Let the discount factor be ε ($0\le\varepsilon<1$). Then, the Bellman equation value iteration for the infinite horizon discounted cost dynamic program yields the following optimal cost-to-go functions:

$$
\begin{aligned}
J^*_T(a,b) &= \min_{\ell\in\mathcal{L}}\{f^T_\ell(a,b)\}\\
J^*_{S,j}(a_1,a_2,b_1,b_2) &= \min_{\alpha\in\mathcal{L}_1,\ \beta\in\mathcal{L}_2}\{f^{S,j}_{\alpha\beta}(a_1,a_2,b_1,b_2)\},
\end{aligned}
$$

where

$$
\begin{aligned}
f^T_{c_1c_1}(a,b) &\doteq E_{c_1,\phi,\psi}[2c_1+\varepsilon J^*_T(\phi,\psi)],\\
&\vdots\\
f^T_{c_1a}(a,b) &\doteq a+E_{c_1,\phi_1,\psi_1,\psi_2}[c_1+\\
&\qquad \varepsilon J^*_{S,0}(\phi_1,b,\psi_1,\psi_2)|b]\\
f^{S,0}_{c_{11}c_{12}}(a_1,a_2,b_1,b_2) &\doteq E_{c_{11},c_{21},\phi_1,\psi_1,\psi_2}[c_{11}+c_{12}+\\
&\qquad \varepsilon J^*_{S,0}(\phi_1,b_1,\psi_1,\psi_2)|b_1]\\
&\vdots\\
f^{S,0}_{c_{11}a_2}(a_1,a_2,b_1,b_2) &\doteq a_2+E_{c_{11},\phi_1,\psi_1,\psi_2}[c_{11}+\\
&\qquad \varepsilon J^*_{S,1}(\phi_1,b_2,\psi_1,\psi_2)|b_2]
\end{aligned}
$$

Here, f^T_δ and $f^{S,j}_{\alpha\beta}$ denote the expected aggregate discounted cost incurred in moving infinitely in the direction

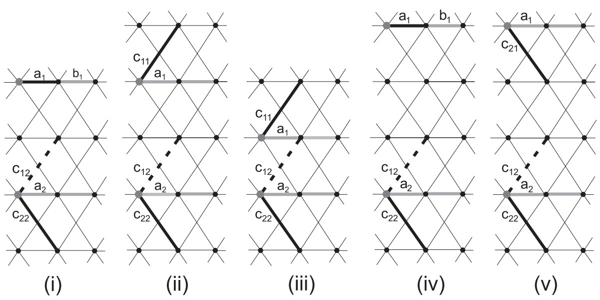

Fig. 4. (i), (ii), and (iii) illustrate the situations in which separation between two agents grows beyond that in $\mathcal{C}_2$. (iv) and (v) illustrate the situations in which the $\mathcal{C}_3$ separation stays unchanged even if one of the agents fails to choose an a_i link. The dashed lines indicate the preferred policy μ_1 chosen by one of the agents which avoids such separations. The solid lines indicate μ_2.

of t using δ and α, β respectively in the current stage and using an optimal policy from then onwards. The optimal value functions J_T^* and $J_{S,j}^*$, $\forall j$, allow to calculate the optimal path for the platoon given the link costs in the lookahead region.

C. Proof of Lemma 1

We present the proof for $m = 7$. The arguments can be used for the other cases as well. There are only three situations in which there exists a policy that leads to the configuration $\mathcal{C}_3$ from $\mathcal{C}_\ell$ where $\ell < 3$ (see Fig. 4 for the illustration and terminology). In each of the three cases, we will show that policy μ_1 is preferable to the policy μ_2.

In the Fig. 4 (i) configuration, μ_1 is preferable to μ_2 if

$$a_1 + E_{c_{22},\phi_2,\psi_1,\psi_2}[c_{22} + J_{S,2}^*(b_1, \phi_2, \psi_1, \psi_2)|b_1] > a_1 + E_{c_{12},\phi_2,\psi_1,\psi_2}[c_{12} + J_{S,0}^*(b_1, \phi_2, \psi_1, \psi_2)|b_1],$$

i.e., the expected values of c_{12} and c_{22} being equal, if

$$E_{\phi_2,\psi_1,\psi_2}[J_{S,2}^*(b_1, \phi_2, \psi_1, \psi_2)|b_1] > E_{\phi_2,\psi_1,\psi_2}[J_{S,0}^*(b_1, \phi_2, \psi_1, \psi_2)|b_1]. \quad \text{(A.1)}$$

Let J_k denote the approximate value function obtained after k value iteration steps. Although J^* is not precisely known, its value iteration upper bounds and lower bounds, say $\overline{J}_k$ and $\underline{J}_k$, are known [2] and those can be used to check validity of the inequality (A.1). Omitting the subscripts k for notational simplicity, it follows that if

$$E_{\phi_2,\psi_1,\psi_2}[\underline{J}_{S,2}(b_1, \phi_2, \psi_1, \psi_2)|b_1] > E_{\phi_2,\psi_1,\psi_2}[\overline{J}_{S,0}(b_1, \phi_2, \psi_1, \psi_2)|b_1] \quad \text{(A.2)}$$

holds for some k, then the inequality (A.1) holds as well. Using numerical data, (A.2) can be verified to hold for all b_1. Similarly, in the other two configurations, the following conditions need hold for some k:

$$E_{\phi_1,\phi_2,\psi_1,\psi_2}[\underline{J}_{S,2}(\phi_1, \phi_2, \psi_1, \psi_2)] > E_{\phi_1,\phi_2,\psi_1,\psi_2}[\overline{J}_{S,1}(\phi_1, \phi_2, \psi_1, \psi_2)], \quad \text{(A.3)}$$

$$E_{\phi_1,\phi_2,\psi_1,\psi_2}[\underline{J}_{S,2}(\phi_1, \phi_2, \psi_1, \psi_2)] > E_{\phi_1,\phi_2,\psi_1,\psi_2}[\overline{J}_{S,0}(\phi_1, \phi_2, \psi_1, \psi_2)]. \quad \text{(A.4)}$$

The above two inequalities can be verified to hold using numerical data.

It now needs to be established that the configuration $\mathcal{C}_3$ always evolves into either $\mathcal{C}_2$ or $\mathcal{C}_1$ unless it is optimal for the agents to choose links a_1 and a_2. There are only two situations in which there exists a policy that keeps $\mathcal{C}_3$ in $\mathcal{C}_3$ should the agents choose not to take the straight links a_i (see Fig. 4 for the illustration and terminology). Using the arguments employed earlier, it follows that μ_1 is preferable over μ_2 if (A.4) and

$$E_{\phi_1,\psi_1,\psi_2}[\underline{J}_{S,2}(\phi_1, b_1, \psi_1, \psi_2)|b_1] > E_{\phi_2,\psi_1,\psi_2}[\overline{J}_{S,1}(b_1, \phi_2, \psi_1, \psi_2)|b_1] \quad \text{(A.5)}$$

hold for some k. Using numerical data, the inequalities (A.4) and (A.5) can be verified to hold. QED.

VIII. Acknowledgment

This work has been financially supported by Navy-ONR Young Investigator Award Grant N000-99-1-0668 and Air Force Research Laboratory Grant F33615-01-C-1850.

References

[1] E. T. Bair. A Perspective on Army UAV Needs and Initiatives. *Program Executive Office, Intelligence, Electronic Warfare and Sensors*, September 1999.
[2] D. Bertsekas and J. Tsitsiklis. *Neuro-Dynamic Programming.* Athena Scientific, Belmont, MA, 1996.
[3] D. Blei and L. Kaelbling. Shortest Paths in a Dynamic Uncertain Domain. In *IJCAI Workshop on Adaptive Spatial Representations of Dynamic Environments*, pages 81–87, 1999.
[4] C. Boutilier. Multiagent systems: Challenges and opportunities for decision-theoretic planning. *AI Magazine*, 20(4):35–43, 1999.
[5] US DoD. Advanced Concept Technology Demonstration list for 2002 Announced. *US DoD, News Release no. 103-02.*
[6] W. Heinzelman, A. Chandrakasan, and H. Balakrishnan. An Application-Specific Protocol Architecture for Wireless Microsensor Networks. *IEEE Transactions on Wireless Communications*, 2002 (*To appear*).
[7] C. Hellman and E. Clark. Surveillance Platforms. *Terrorism Project - Center for Defense Information*, October 2001.
[8] S. P. Howard. Special Operations Forces and Unmanned Aerial Vehicles: Sooner or Later? *School of Advanced Airpower Studies, Air University, Maxwell Air Force Base, AL*, June 1995.
[9] Major R. C. Nolan II. The Pilotless Air Force? A Look at Replacing Human Operators with Advanced Technology. *The Research Department - Air Command and Staff College*, 1997.
[10] L. Kaelbling, M. Littman, and A. Cassandra. Planning and Acting in Partially Observable Stochastic Domains. *Artificial Intelligence*, 101:33–94, 1998.
[11] J. De Mot, S. Gentry, V. Kulkarni, and E. Feron. Stochastic Shortest Path Planning with Direction-dependent Lookahead. *IEEE Conference on Decision and Control*, 2002 (*submitted*).
[12] T. Rappaport. *Wireless Communications: Principles and Practice.* Prentice Hall, Upper Sadle River, NJ, 1996.
[13] R. Steele, C. Lee, and P. Gould. *GSM, cdmaOne and 3G Systems.* John Wiley and Sons, New York, NY, 2001.