Optimal Perimeter Patrol via Max-Plus Probability*

Ben G. Fitzpatrick, Kristin Holmbeck-Cannell, Matthew Laffin, and Yun Wang, Member, IEEE

Abstract— This paper addresses a perimeter patrol problem involving control of unmanned air vehicles. Around the perimeter of a protected area are placed a number of unmanned ground stations. These stations send alert signals which are investigated by the unmanned air vehicles. We present an approach to optimization of vehicle tasking based on max-plus probabilistic models and computational schemes.

I. INTRODUCTION

The use of and interest in unmanned air vehicles (UAVs) as important components in military operations continues to grow. The US Air Force’s UAV plan [1] indicates a goal of moving from “man in the loop” to “man on the loop,” requiring a fairly high level of autonomous operation. Significant improvements in autonomy are necessary to achieve this ambitious program. Of great interest is pushing the envelope of current capability to allow unmanned autonomous systems to make time-critical decisions without the need for human oversight or control.

Requirements of intelligence, surveillance and reconnaissance (ISR) missions mesh very well with UAV capabilities and man on the loop operational goals. One important consideration is that the complications of autonomous and organic combat collaboration are not an issue. Moreover, sending UAVs into uncertain and unprotected spaces clearly offers advantages for troop safety. Organizing search patterns and delivering data for human assessment then become key planning problems.

We consider in this paper a stochastic control problem derived from the Air Force Research Laboratory’s Cooperative Operations in Urban TERRain (COUNTER) project [2, 3]. UAVs and unmanned ground sensors (UGSs) are networked to communicate wirelessly. These UGSs are positioned in the battlespace at locations of interest (which we call stations), and they transmit alerts when they detect a potential target in their field of regard. These sensors may be seismic, chemical, or some other relatively inexpensive and simple technology. When an UGS puts an alert onto the network, the UAVs can then travel to and inspect the station by capturing video. Depending on the capabilities of the UAV and the network, the video may be transmitted via the network or delivered physically to a human for examination.

The COUNTER subproblem we consider is one of a collection of UGSs stationed around the perimeter of a protected area. The UGSs generate alerts signaling a potential incursion attempt into the protected region. The UAVs patrol this perimeter, searching stations at which UGSs have generated alerts. This form of human/UAV collaboration balances nicely the abilities of computational autonomy and operator intervention. The decisions of where to send and when to retask are handled by the automation, while the human focuses on the visual imagery to detect intrusions.

We develop a Markov chain model of the dynamics of alerts arriving and UAV alert servicing. We structure the solution process as a stochastic optimal control problem, which is solved with dynamic programming. A novel feature of the approach is that we use max-plus probabilistic models, leading to additional computational efficiencies. Max-plus probability strikes a balance between traditional expectations and more conservative worst-case risk averse solutions.

The remainder of the paper is organized as follows. In Section II, we construct a mathematical model of the patrol problem. Section III contains the stochastic control problem construction. A max-plus probability revision is the subject of Section IV, and in Section V we provide some illustrative results and timing information. We wrap up with some remarks in Section VI.

II. DEFINING THE PERIMETER PATROL PROBLEM

Our model begins with m UGSs and n UAVs. The UGSs are positioned at fixed locations on the perimeter of a protected region, which is modeled as a simple closed curve surrounding the area of interest. The curve is assumed to be parameterized in polar coordinates ($r(\theta), \theta, \theta \in [0, 2\pi]$).

The arrival rate of alarms follows a Poisson process with rate $\alpha$. We assume that at any given time, at most one station can receive an alert. An alert arrival is equally likely to occur at any of the stations occupied by UGSs, leading to an effective individual arrival rate of $\alpha/n$. We model this process with indicator variables of alert occurrence. That is, we denote by $Y(t) = (Y_1(t), Y_2(t), \ldots, Y_m(t))$ the vector of alerts that occur at time $t$. By assumption, at most one entry in the vector can be 1. If all are zero, no alert occurs at time $t$.

This perimeter patrol problem is illustrated in Figure 1. UAVs follow the path around the area of interest, visiting UGS stations that generate alerts. The circles denote UGSs, with the red filled circle denoting an alarmed state. The UAVs fly patrol along the perimeter. When a UAV arrives at an alarmed UGS, the UAV’s sensor is engaged so that it can investigate the UGS’s vicinity.

*Research supported in part by Air Force Office of Scientific Research grants FA9550-09-1-0524, FA9550-10-1-0499, and FA9550-12-C-0035.

Ben G. Fitzpatrick is with the Department of Mathematics, Loyola Marymount University, Los Angeles, CA 90045 USA (phone: 310-338-7892; fax: 310-216-1628; e-mail: bfitzpatrick@lmu.edu) and Tempest Technologies, Los Angeles CA 90045 USA (phone: 310-338-7892; fax: 310-216-1628; e-mail: bfitzpatrick@tempest-tech.com).

Kristin Holmbeck-Cannell and Yun Wang are with Tempest Technologies, Los Angeles CA 90045 USA (phone: 310-216-1677; fax: 310-216-1628; e-mail: hohlmbeck, yun.wangl@tempest-tech.com).

Matthew Laffin, formerly of Tempest Technologies, is with KLA-Tencor, Milpitas, CA 95035; e-mail: matthew.laffin@kla-tencor.com.
An additional step in the modeling effort is the discretization of the continuous perimeter into $N+1$ discrete locations, $0,1,...,N$. The distance between any two locations is computed taking into account the wrapping:

$$
dist(i,j) = \min\{ |i-j|, N-|i-j| \}. \quad (1)$$

The stations, by which we mean the locations at which the UGSs reside, are denoted $(g_1,...,g_m) \in S^m=\{0,1,...,N\}^m$. These are fixed problem parameters. The locations of the UAVs, denoted by $V=(V_i,...,V_s) \in S^m$, are state variables. The dynamics for this component of the state are given by

$$v(t+h) = v(t) + u(t) \quad (2)$$

where $h$ denotes the time step required to move a UAV from one location in the state space to an adjacent location. The control variable $u$ satisfies the constraint $u \in \{-1,0,1\}$, in which -1 and 1 denote movement to an adjacent location. The control action $u=0$ denotes a loiter move for the UAV, which is allowed only when the UAV is at a station $v(t)=g_j$ for some $j$. Additionally, we constrain the UAVs so that they only loiter at UGS positions that have generated alerts. Thus, we need an additional state variable to keep up with the alert state.

In [4], a model of alerts involves a running indicator $A_i(t)$ governed by the dynamic equation

$$A_i(t+h) = \begin{cases} 
  0 & v_i(t) = i, u_i(t) = 0 \\
  A_i(t) \lor Y_i(t), & \text{otherwise} 
\end{cases} \quad (3)$$

The symbol $\lor$ denotes Boolean “or.” Note that the alarm state is zeroed out upon a loiter by one of the UAVs. In addition the model of [4] keeps up with the length of time each alert waits to be serviced. We denote by $T_i(t)$ the number of time steps an alert at station $i$ has been waiting to be serviced by a UAV. This state variable has dynamics

$$T_i(t+h) = \min\{ A_i(t)(T_i(t)+1), \bar{T} \}, \quad (4)$$

where $\bar{T}$ denotes the maximum wait time, another parameter to be specified. Finally, we define a dwell state

$$d(t+h) = (d(t)+1)(1-|u(t)|) \quad (5)$$

to keep up with the length of time the UAV spends at a given station.

Taking the position, alert status, waiting times, and dwell times together as the system state $x$, we have a dynamical system that can be written as

$$x(t+h) = f(x(t), u(t), Y(t)), \quad (6)$$

in which $f$ encapsulated the dynamics, $u$ is the control, and $Y$ is the alert arrival process. With the dynamics specified, we turn to the control problem.

### III. Stochastic Control for Perimeter Patrol

The optimization posed in [4] takes a quality of service approach. The cost function implemented therein balances information gained by loitering against the waiting times of alerts. We define the objective functional $J$ by

$$J(u) = E\left[ \sum_{k=1}^{\infty} \beta^k \left( H(d(k)) + (1-|u(k)|) - \beta \max_i T_i(k) \right) \right] \quad (7)$$

an infinite horizon time horizon discounted reward structure. In this cost functional, parameters $\beta$ and $\gamma$ denote the discount factor and penalty weight, respectively. The function $H(d)$ denotes the information obtained by loitering for $d$ dwell periods, forming the reward component. The penalty component (a negative reward or cost) is a scaled maximum waiting time: we seek a patrolling and dwelling strategy that avoids long waits.

The information function is meant to model an operator’s ability to extract information from the data the UAV collects. In order to task the UAVs autonomously, the UAV needs to have a model that describes what is gained by loitering. The information model is described in detail in [3], treating the operator as a sensor with a confusion matrix with correct and false detection of threats and nuisances. The Shannon information can be computed as a function of the number of dwells:

$$H(d) = \log \frac{P_{a}(d)}{pP_{na}(d)} \left( \frac{1}{pP_{na}(d)} + (1-p) \right)$$

$$+ p(1-p)\log \frac{1-p}{p(1-p)P_{na}(d)} \left( \frac{1}{p(1-p)P_{na}(d)} + (1-p) \right)$$

$$+ (1-p)\log \frac{P_{na}(d)}{1-p(1-p)P_{na}(d)} \left( \frac{1}{1-p(1-p)P_{na}(d)} + (1-p) \right)$$

in which $p$ denotes the a priori probability that an alert is a threat (not a nuisance), $P_{TR}(d)$ denotes the dwell dependent conditional likelihood of recognizing an existing threat, and $P_{NR}(d)$ denotes the dwell dependent conditional likelihood of recognizing an existing nuisance. The functions $P_{TR}$ and $P_{NR}$ both increase with $d$ towards an asymptotic probability (that can be strictly less than one) so that there is a diminishing return for dwelling many times. The graph of an example information function, following the parametric choices of [3], is illustrated in Figure 2.

With the information function in hand, we may turn to the problem of stochastic control. Maximizing the discounted infinite horizon reward $J$ over UAV control plans attempts to increase information gain from dwells against the waiting...
times of unvisited stations. The infinite horizon discounted 
reward, coupled with bounded information function and wait 
times, leads to a time-independent value function and a full 
state feedback control.

$$G(x,u) = \prod_{k=1}^{w} P[\text {an alarm at UGS } k \text { is not missed }]$$

$$= \prod_{k=1}^{w} (1 - q)^{v} I(T_{k} > 0) + \sum_{j=1}^{u} \{ l - P_{ug}(d_{j}) \} I(v_{j} = k) p$$  (10)

which includes the dispersion probability for a station that is 
waiting and the probability of missing a detection during 
dwells.

We make an approximation at this point, in order to 
reduce the state space. The waiting-time state variable is very 
expensive to maintain from a computational perspective, so 
we approximate the dispersion term of the missed alarm 
probability with the following:

$$(1 - q)^{v} I(T_{k} > 0) = (1 - q)^{M}$$  (11)

This rather coarse waiting time approximation is really an 
approximation of the last visit time by one of the UAVs. The 
approximation has the effect on the optimization of keeping 
M small by reducing dwells or maintaining smaller distances 
between UAVs and UGSs. The other terms, modeling the 
information gain from dwelling, tend to keep dwell time 
high, attaining a similar balancing objective to that of the 
previous cost structure. The advantage to this approximation 
is that we no longer need to keep the waiting times as state 
variables, reducing the dimensionality dramatically.

A second innovation on previous approaches to the 
perimeter patrol problem is the use of max-plus probability, 
which we discuss in Section IV.

IV. MAX-PLUS STOCHASTIC CONTROL

The max-plus algebra, as discussed in [5,6,7,8,9,10], 
involves a redefinition of arithmetic operations, 
computational and analytical benefit. We consider the real 
numbers, augmented by $-\infty$ : $\mathbb{R}^{-} = \mathbb{R} \cup \{ -\infty \}$. On this 
set, we define two operations, $\oplus$ and $\otimes$, by

$$a \oplus b = \max \{ a,b \},$$
$$a \otimes b = a + b.$$  (12)

It is well known that $\mathbb{R}^{-}$ forms a commutative semi-ring 
under these operations. The additive identity is $-\infty$ while 
the multiplicative identity is 0. Except for the additive 
identity, every element has a multiplicative inverse, 
suggesting that one might be able to extend the structure to a 
field structure. However, addition in this semi-ring is 
idempotent, meaning that $a \oplus a = a$. It is important to note 
that the only rings satisfying additive idempotency are 
trivial; that is, the only element is the additive identity. 
Thus, extending the semi-ring to a ring (and hence a field) is 
not a possibility.

From these basic operations, we can build standard linear 
algebraic objects, such as matrices and vectors. If we 
consider an $n \times n$ array, $A$, of elements of $\mathbb{R}^{-}$ and a column
vector, \( x \), of \( n \) elements of \( \mathbb{R}^+ \), define the max-plus matrix-vector product \( y = A \otimes x \) by

\[
y_j = \max_{i=1}^n (A_{ij} \otimes x_j) = \max_j \{ A_{ij} \cdot x_j \}.
\] (13)

To illustrate the application of max-plus algebraic structure, we consider a discrete time nonlinear control problem. We begin with a dynamical system under control of the form

\[
x(t + 1) = f(x,u), \quad x(0) = x_0
\] (14)

with a control objective given by

\[
J(u, x_0, t_0) = \sum_{t=0}^{t_f} L(x(t), u(t))
\] (15)

which is to be maximized of the set of admissible control functions, \( u \). The Bellman equation of dynamic programming is

\[
V(y, t) = \max_u \{ L(y, u) + V(f(y, u), t + 1) \}.
\] (16)

By inspection, this iteration is max-plus linear. Indeed, the purpose of introducing max-plus arithmetic is the linearity of value iteration.

In the case of stochastic control, the Bellman equation is not max-plus linear. The reason for this is that the expectation

\[
E(X) = \int X(\omega) dP(\omega)
\] (17)

is not max-plus linear. Scalar multiples factor out of the expectation in max-plus arithmetic, but the expectation of the max-plus sum is not the max-plus sum of expectations:

\[
\int (X(\omega) + Y(\omega)) dP(\omega) \neq \int X(\omega) dP(\omega) \oplus \int Y(\omega) dP(\omega)
\]

\[
\max(X(\omega), Y(\omega)) dP(\omega) \neq \max\left\{ \int X(\omega) dP(\omega), \int Y(\omega) dP(\omega) \right\}.
\] (18)

To treat the UAV planning problems of interest, we need to introduce probabilistic models compatible with max-plus arithmetic [11, 12, 13, 14]. A max-plus probability space is a triplet \((\Omega, F, P)\) of a set \(\Omega\) called the sample space, a sigma-field \(F\) of subsets of \(\Omega\), and a max-plus probability measure \(P\) that satisfies the following:

\[
P(\Omega) = 0,
\]
\[
P(\emptyset) = -\infty,
\]
\[
\left\{ A_i \right\}_{i=1}^n \subset F, \text{ mutually exclusive} \Rightarrow \bigoplus_{i=1}^n P(A_i) = \max_i \{ P(A_i) \} = P\left( \bigcup_{i=1}^n A_i \right).
\] (19)

These criteria are completely analogous to the standard definition of probability measures in “plus-times” arithmetic. Note that max-plus probabilities have some unintuitive properties. For example,

\[
\max \{ P(A^c), P(A) \} = 0.
\] (20)

For the probability measure \(P\) to have a density \(p\), we must have

\[
P(A) = \max_{x \in A} \{ p(x) \}.
\] (21)

Max-plus probabilities can be derived from standard probability measures through a large deviations approach. Max-plus arithmetic can be viewed as a limit of log-plus arithmetic:

\[
a \oplus_{\epsilon} b = \epsilon \log \left( \exp(a/\epsilon) + \exp(b/\epsilon) \right),
\]

\[
a \oplus_{\epsilon} b = \epsilon \log \left( \exp(a/\epsilon) \exp(b/\epsilon) \right).
\] (22)

Max-plus operations are the limits of the respective log-plus operations as \(\epsilon \to 0\). Log-plus operations enter into probabilistic modeling in the following way. We consider a family of probability measures \((P_\epsilon)_{\epsilon>0}\) on the measurable space \((\Omega, F)\). If this family obeys a large deviation principle, then the limit

\[
Q(A) = \lim_{\epsilon \to 0} \epsilon P_\epsilon(A)
\] (23)

exists and is a max-plus probability measure [MC]. As a simple example, consider a family of Gaussian densities

\[
p_\epsilon(x) = \frac{1}{\sqrt{2\pi\epsilon}} \exp \left( -\frac{1}{2\epsilon} |x - \mu|^2 \right)
\] (24)

with vanishing variance \(\epsilon \to 0\). Then the measure \(Q\) is characterized by the quadratic density

\[
q(x) = -\frac{1}{2} |x - \mu|^2.
\] (25)

This process gives us a method of constructing max-plus measures from exponential density families [11, 13].

Expectations in max-plus probability are defined by the following:

\[
E_{\oplus} (X) = \max_{\omega} \{ X(\omega) + P(\omega) \}
\] (26)

with which we can define objective functions for max-plus stochastic control problems.

Discrete time max-plus stochastic control involves a dynamical system

\[
x(t + 1) = f(x, u, y), \quad x(0) = x_0
\] (27)

in which \(y\) denotes a max-plus random disturbance. The cost functional is given by

\[
J(u, x_0, t_0) = E_{\oplus} \left[ \max_{t \leq t_f} \{ L(x(t), u(t)) \} \right]
\]

\[
= \max_{t \leq t_f} \left[ L(x(t, \omega), u(t)) + p(\omega) \right]
\] (28)

where the max-plus probability density \(p\) is the model for the random process \(y\). This process weighs the uncertainty inherent in the noise with the reward for the state. A similar risk aversion criterion for state estimation in adversarial games can be found in [15].
The Bellman equation of dynamic programming is then

\[
V(x,t) = E^\omega \left[ \max_u \left\{ L(x,u) + V(f(x,u,\omega), t+1) \right\} \right] \tag{29}
\]

a max-plus linear equation. With these tools prepared, we return to the perimeter patrol problem.

V. MAX-PLUS SOLUTION OF PERIMETER PATROL PLANNING

The running cost for the max-plus approach uses the function

\[
G(x,u) = \prod_{k=1}^m R_k(x,u)
\]

\[
R_k(x,u) = \left[ (1-q)^{M(x,u)} + \sum_{j=1}^p (1-P_{M,j})I(v_j = k)p \right.
+ \sum_{j=1}^p (1-P_{M,j})I(v_j = k)(1-p) \right]. \tag{30}
\]

In the max-plus formulation, we take the log of this quantity for the running cost:

\[
L(x,u) = \sum_{k=1}^m \log \left( R_k(x,u) \right). \tag{31}
\]

The probability model for the Poisson/geometric alert arrival is also logged in order to produce the objective

\[
J(u,x_0,t_N) = \max_{\alpha} \max_{t_N \in S} \left[ L(x(t,\omega),u(t)) + p(\omega) \right] \tag{32}
\]

to which we may apply the max-plus iteration

\[
V(x,t) = \max_{\alpha} \left\{ L(x,u) + V(f(x,u,\omega), t+1) + p(\omega) \right\}, \tag{33}
\]
to determine the value and the optimal control.

To illustrate, we consider the example computed and illustrated in [4]'s paper, with 1 UAV, 4 UGSs, and a perimeter discretized with \(N=15\) locations. For the original information gain/waiting time penalty formulation, we take \(\lambda=0.9\), \(\beta=0.0013\). Dwell times are capped at a maximum of 5, and wait times saturate at 15 delays. Alert arrivals occur at a rate of once every 15 time steps. Thus, the wait time constraint and arrival rate are selected to correspond to one UAV lap around the perimeter. The arrival rate model governs the stochasticity in the dynamics and the expectation in the objective functional. The dispersion probability we use is 0.6.

We implement three algorithms to compute the value function and control: the standard value iteration method, a state-aggregated value iteration derived from [4], and the max-plus stochastic control method. Each of these has a different state space, so we do not attempt to compare value functions. Rather, we run Monte Carlo simulations of the process, generating random alerts according to the Poisson arrival model, and we apply the feedback control derived from the dynamic programming solution.

In Figures 3 and 4, we see that the distribution of dwell times and the alert waiting times over the course of the simulation are quite similar for all three methods. We also compare the Monte Carlo averaged objectives for the three techniques, in Figure 5, using the objective specified for the standard stochastic control problem. The state aggregation and max-plus control algorithms show performance remarkably close to the computational optimum.
Computation, so we compare only the state aggregated and max-plus algorithms. Figures 6 and 7 show the dwell time distribution and the wait time distribution for the state aggregation method of [4] and the max-plus method we have constructed. Figure 8 shows values of the objective functional evaluated at a sampling of states and averaged over the 1200 realizations of the simulated process.

Computations were carried out on a 3 Ghz Dell PC with 16Gb RAM running Windows 7 PC and Matlab 2011a.

VI. CONCLUSION

We have developed a new dynamic model and objective functional for designing UAV perimeter patrol plans. Using max-plus probability, we have implemented an efficient value iteration that is linear in the max-plus algebra. The combination of the state space reduction and the Bellman equation simplification make for a very efficient computational approach.

ACKNOWLEDGMENT

The authors gratefully acknowledge the support of the Air Force Office of Scientific Research. We also thank the Air Force Research Laboratory Control Sciences branch, in particular D. Doman, D. Casbeer, P. Chandler, K. Krishnamoorthy, and M. Pachter for a number of helpful and stimulating interchanges on COUNTER and the perimeter patrol problem.

REFERENCES