12TH ICCRTS

"Adapting C2 to the 21st Century"

Applying Spatial-Temporal Model and Game Theory to Asymmetric Threat Prediction

Assigned Topics: Modeling & Simulation (Track 3)
Assigned Paper Number: I-063

Mo Wei
Intelligent Automation, Inc.
15400 Calhoun Drive, Suite 400
Rockville, MD 20855
Tel: 301 294 5227
Fax: 301 294 5201
Email: mwei@i-a-i.com

Genshe Chen (Principal point of contact)
Intelligent Automation, Inc.
15400 Calhoun Drive, Suite 400
Rockville, MD 20855
Tel: 301 294 5218 (direct)
Fax: 301 294 5201
Email: gchen@i-a-i.com

Jose B. Cruz, Jr.,
The Ohio State University
205 Dresses Laboratory, 2015 Neil Ave
Columbus, OH 43202
Ph: (614)292-1588
Email: cruz.22@osu.edu

Leonard Haynes
Intelligent Automation, Inc.
15400 Calhoun Drive, Suite 400
Rockville, MD 20855
Tel: 301 294 5200
Fax: 301 294 5201
Email: lhayes@i-a-i.com

Martin Kruger
The Office of Naval Research
Email: Martin_Kruger@onr.navy.mil

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2007		**00-00-2007 to 00-00-2007**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Applying Spatial-Temporal Model and Game Theory to Asymmetric Threat Prediction	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Intelligent Automation, Inc.,15400 Calhoun Drive, Suite 400,Rockville,MD,20855**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Twelfth International Command and Control Research and Technology Symposium (12th ICCRTS), 19-21 June 2007, Newport, RI

14. ABSTRACT

Accurate predictions of enemy course of actions (ECOA) are important to the command and control optimization strategies in long-lasting battles. In most Command and Control (C2) applications, the existing techniques, such as spatial-temporal point models for ECOA prediction or Discrete Choice Model (DCM), assume that insurgent attack features/patterns, or at least the trends of behavior patterns, are static. However, this static assumption is no longer true for intelligent and organized insurgents in recent antiterrorism war. These insurgents sometimes deliberately violate probability theory predictions so they can apply surprise attacks to create more casualties and spread terror. In this paper, a new game theoretic framework is proposed for modeling dynamic changes of enemy behavior features and predicting future threats. This framework semantically combines several different approaches; namely, a feature prediction game, higher level hybrid data fusion, techniques to provide concrete spatial-temporal modeling and prediction, emotion analysis of adversary rationality and non-rationality, deception identification and modeling, hierarchical knowledge representation, and a non-zero sum stochastic adversarial Markov game. We mainly describe the modification of existing spatial-temporal point models, the fusion of dynamic game feature selection technique and dynamic cohesiveness feature selection technique, the ontology about selected/unselected features, and construction of probability predictions.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**47**	

Erik Blasch
AFRL/SNAA
Erik.Blasch@WPAFB.AF.MIL

ABSTRACT

Applying Spatial-Temporal Model and Game Theory to Asymmetric Threat Prediction

Accurate predictions of enemy course of actions (ECOA) are important to the command and control optimization strategies in long-lasting battles. In most Command and Control (C2) applications, the existing techniques, such as spatial-temporal point models for ECOA prediction or Discrete Choice Model (DCM), assume that insurgent attack features/patterns, or at least the trends of behavior patterns, are static. However, this static assumption is no longer true for intelligent and organized insurgents in recent anti-terrorism war. These insurgents sometimes deliberately violate probability theory predictions so they can apply surprise attacks to create more casualties and spread terror.

In this paper, a new game theoretic framework is proposed for modeling dynamic changes of enemy behavior features and predicting future threats. This framework semantically combines several different approaches; namely, a feature prediction game, higher level hybrid data fusion, techniques to provide concrete spatial-temporal modeling and prediction, emotion analysis of adversary rationality and non-rationality, deception identification and modeling, hierarchical knowledge representation, and a non-zero sum stochastic adversarial Markov game. We mainly describe the modification of existing spatial-temporal point models, the fusion of dynamic game feature selection technique and dynamic cohesiveness feature selection technique, the ontology about selected/unselected features, and construction of probability predictions.

1. Introduction

In recent years, many models and techniques have been applied to symmetric adversarial situations such as wargaming or criminal cases with plenty of prediction results. However, asymmetric adversaries faced by U.S. forces become more and more frequent and satisfying predications are still quite difficult to obtain. With a belief that the end justifies the means and not constrained by some basic morality, asymmetric adversaries are often able to engage surprise attacks and successfully disappear before U.S. forces have time to concentrate. Adversaries' ability to learn and adapt their attack patterns and the often unstructured data sources such as ontology information complicate this problem even more. Information that could be used to dynamically identify attack patterns and changes of patterns and predict when-where-how such attacks are most likely to happen will greatly benefit U.S. forces by reducing such surprise attacks to non-surprising events.

To successfully model and predict the adversaries' most likely future courses of actions (COAs), automated processing techniques that identify patterns of COAs and assess possible future threats are critical needs. By dynamically mining and analyzing terrorists and insurgents' preferences and features of actions, U.S. forces can locate dangerous time slots and locations more accurately, prepare counterstrike COAs, or at least mitigate the impact of attacks. At the same time, the new information identified can be automatically fused to the existing system and prediction performance will be improved.

Liu and Brown [11] applied extended discrete choice models to prediction of spatial probability of criminal activity, which produces much higher accuracies than traditional hot-spotting techniques. The fundamental technique utilized by Liu and Brown is to utilize more features in predictions. However, a key limitation of the approach is that the changes in the subjects' decision-making processes are not modeled. This is to say, the subjects' preferences are assumed static.

To solve such problem, a new framework for modeling dynamical changes of features and predicting future threats is proposed. In this framework, techniques that can provide concrete spatial-temporal modeling and prediction, higher level hybrid data fusion, emotion analysis of adversary rationality and non-rationality, deception identification and modeling, hierarchical knowledge representation, and a non-zero sum stochastic adversarial Markov game are combined semantically.

This paper is organized as follows. In Section II, we will summarize the technical approach, which includes problem description, the refinement of spatial-temporal models, the feature ontology, the fusion of different feature selection techniques, and production of probability predictions. Section III describes the experimental results and explanations. Section IV provides conclusions for the paper.

2. Technical Approach

2.1 Framework

The new framework for modeling dynamical changes of features and predicting future threats is shown in Fig. 1.

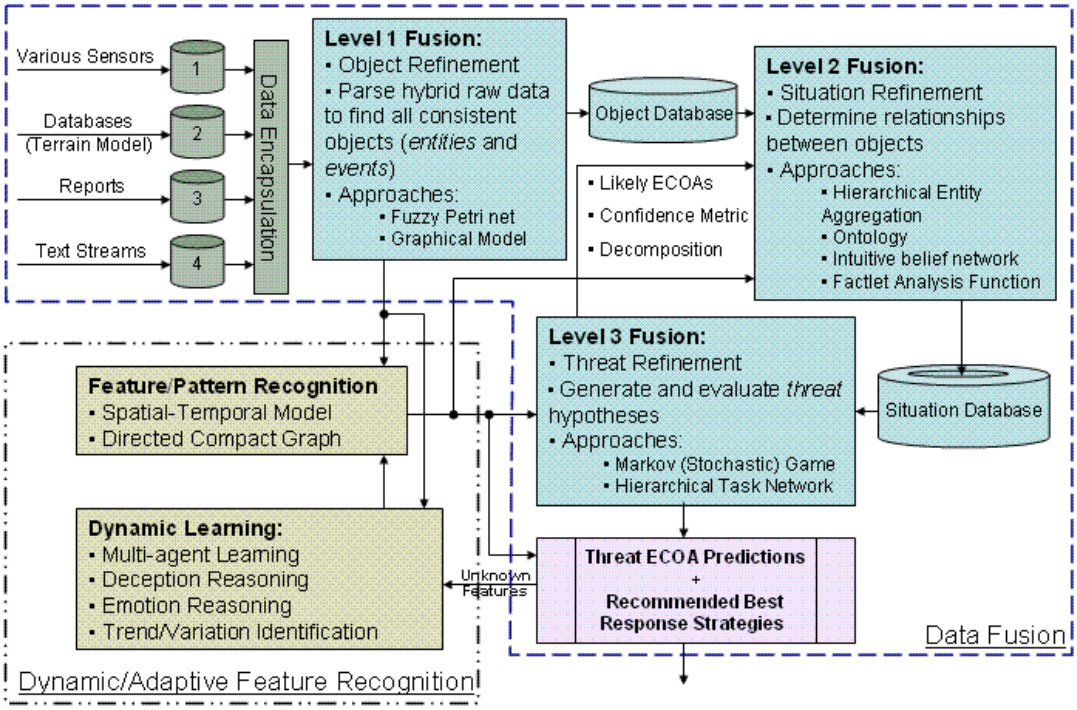


Fig. 1 Overall System Architecture

Inputs are fed into the data refinement (Level 0) and object assessment (L1) data fusion components. The refined objects and related pedigree information are used by a feature/pattern recognition block to generate primitive prediction of intents of adversary asymmetric threats. If the observed features are already associated with adversary intents, we can easily obtain them by pattern recognition. In some time-critical applications, the primitive prediction can be used before it is refined by relatively time-consuming high-level data fusion.

High-level situation and threat refinement (L2 and L3) data fusion based on a Graphical Markov game model, Hierarchical Entity Aggregation (HEA), Ontology, and Hierarchical Task Network are proposed to refine the primitive prediction generated in stage 1 and capture new unknown features, which will be associated to related L1 results in dynamic learning block. In the dynamic learning model, we take deception reasoning, emotion reasoning, trend/variation identification, and distribution model and calculation into account. Our approach to deception detection is heavily based on the application of pattern recognition techniques to detect and diagnose what we call out-of-normal conditions in the battle space. Out-of-normal conditions are enemy activities that are not deemed as a part of the normal expected evolution of systems. The results of dynamic learning or refinement shall also be used to enhance L2 and L3 data fusion. This adaptive process may be considered as process refinement (Level 4) data fusion.

In this paper, we mainly describe the modification of an existing spatial-temporal point-model, the fusion of dynamic game feature selection technique and dynamic cohesiveness feature selection technique, the ontology about selected/unselected features, and building of the probability predictions.

2.2 Technical Approaches

A typical scenario (Fig. 2), an urban warfare situation, is a good base on which we can illustrate our dynamic adaptive hierarchical game theoretic approach for modeling and prediction of asymmetric threat learning processes. Satellite pictures and available topographic information provide corresponding data.

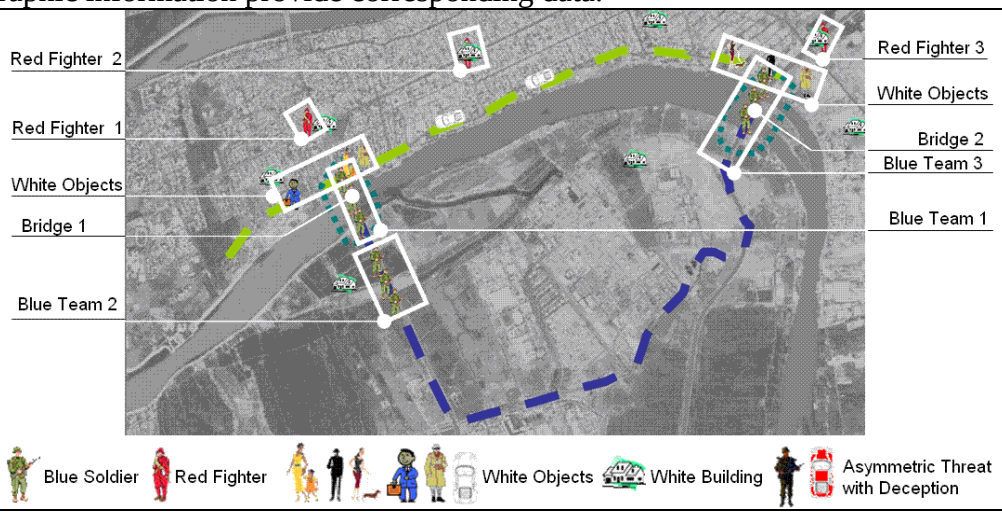


Fig. 1: A typical scenario of urban warfare

The blue force's mission is to try their best to secure the entire area, including the urban districts, bridges, mains roads and blocks, as shown in Fig. 1. The blue ground force consists of several teams of soldiers each with small arms. The red force (terrorist and/or insurgent forces) includes several armed fighters and some asymmetric adversaries hiding in and acting like the white objects (the civilians and vehicles). The red forces are equipped with small arms, mortar, or improved explosive devices (IEDs). We assume there is an asymmetry in total forces between blue side and red side. Blue side has more soldiers than red side. Moreover, the objectives of blue side and red side are asymmetric: the objectives of red side are to kill blue forces, destroy public properties, and cause terror without considering the loss of themselves and the consideration of collateral damage. Usually, they will attack weakly defended or undefended targets and retreat before blue side can concentrate for a counterstrike. However, in this situation, red force may impose some "surprise" attacks that might be chosen deliberately by adversaries to "violate" the predictions based solely on probability theory. In other words, these red forces may change their attack pattern and modify their behavior modes purposefully. Based on this scenario, we estimate the changes in enemy strategies even before such changes have been fully implemented. We give out a primitive estimation of ECOAs by following the pattern/feature recognition model. Based on such prediction, some associated best response strategies of blue side actions can be recommended.

If the primitive prediction is almost correct, there are two possible response strategies for the blue force, according to different goals. If at this time the blue force's purpose is to stop the red forces' actions, the recommended COA of blue force can publicly send out a message (for example, to focus more soldiers/policemen/ambulance on the suspected areas, to put a news on a newspaper that the red commander will read for sure, and/or

some other approaches) to intelligent red forces telling that their actions are in the control of blue side. As a consequence, probably the red forces will refrain from the actions. However, if the purpose of blue force is to set up a trap and catch the insurgents so that in the long run, the total number of attacks will go down; the blue force can only do soldier maneuvers secretly. In other words, in such cases not only the red might do deceptions, the blue might also perform deception actions. If the first guess is not correct (for example, the attack pattern is new and unknown), our game theoretic data fusion module and dynamic learning module will dynamically refine the primitive estimation and update the feature/pattern records.

To fuse possible features/patterns into prediction techniques and game feature prediction and selection, we must first identify different types of surprise attacks and prepare possibly related features. Only after knowing which type of attacks will happen at some next stage with an associated probability, a better resource allocation algorithm is possible. In this scenario, considering information from different resources (such as papers, newspapers, reports from Department of Defense/Navy/Marine/Army/Air Force), the most typical surprise attacks are stated in Table 1.

Table 1 Types of typical attacks

Index	Description
1	Gun Fighter/Mortar/Small Arms
2	IED (Improvised Explosive Device)
3	Kidnap/Hijack
4	Robbery/Stealing
5	“Dirty” bomber/Bio-attacks

Table 1 is based on common sense and numerous reports about the on-going anti-terrorism wars. Such classification is typical and will serve as the initial classification and will not necessarily be fixed. When the war progresses, new types of actions might be added and out-of-date types might be dropped (here out-of-date means for a long time period there is little or no attacks are observed). This is actually a kind of learning/adaptation and will be implemented dynamically via aggregation/clustering partitioning techniques, which combine our hierarchical aggregation technique [1-6] and the work of Milligan, Cooper, Mojena, etc [7-8]. According to different requirements, predictions about the overall probability that an “event” will occur, predictions about the specific probability that an specific attack (such as IED) will occur, and predictions about the relative probabilities that which kind of attack it will be if there will come an attack, can be produced.

In a broad sense, any possible attribute (or feature) might be related to another attribute, which means any attribute can serve as a potential feature or pattern. However, due to real world limits such as computation requirements, usually we can only choose some measurable, available, and “probably” related attributes/hypotheses and put them in a pool of “raw attributes.” In such raw attribute pool, there might still exist hundreds or even thousands of attributes, which would greatly exceed the computation capability of

existing computer systems since each attribute will serve as a dimension and when the number of attributes increases the computation will be daunting. As a result, before ingesting features into the system, a much smaller key feature set should be dynamically selected from the raw attribute pool. A subset of a typical raw attribute pool is stated in Table 2.

Table 2 The attributes

Index	Description
1	Population density per square mile
2	Religion hotness
3	Male people population density per square mile
4	Average family size
5	Young people (from 11 to 29) population density per square mile
6	Average salary per year
7	Average income per person per year
8	Average price of houses
9	Ratio of Children in school
10	Percentage of people who are once involved in drugs
11	Percentage of people who are once involved in crimes
12	Percentage of people who are in debt
13	Average debt per person
14	Average percentage of people who have children
15	Average age
16	Distance to nearest soldier/policemen station
17	Distance to nearest hospital
18	Distance to nearest fire department
19	Distance to nearest highway
20	Distance to nearest church/school/library
21	Distance to nearest park
22	The time difference from last time's attack
23	Distance to nearest occurred attacks
24	Morale of insurgents
25	Average of wellness of public utilities
26	Distance to nearest lake/river/sea
27	Distance to nearest desert/wood
28	Average expenditure on alcohol beverages, tobacco, and smoking
……	……

Note that the 24th feature "morale of insurgents" is not like other features and not easy to measure and quantify, although it is very important in almost every battle. Here according to theories [22-23] about sequential game and social networks by which trust and social organization can be developed, we choose to adapt the "anger formula" in [9-10] and model morale as follows

$$m(t) = m_0 + \sum_{i=1}^{N(t)} A_i(t_i) e^{-h_i(t-t_i)} \quad (1)$$

Where m_0 is the initial value assigned by the experienced human commanders, $N(t)$ is the total number of attacks occurred since the beginning of battle period until time t, h_i is the decay factor which is different for different types of attacks, and t_i is the average of starting time and ending time of the ith attack. A_i is calculated as

$$A_i(t_i) = a_i \frac{L_i^B - L_i^R}{L_i^R} \frac{R_i^B}{R_i^R} \frac{1}{T_i} \quad (2)$$

Where a_i is a coefficient pre-assigned by experienced commanders for different types of attacks. L_i^B is the loss of Blue for ith attack. R_i^B is the resources applied by Blue for ith attack. Since the superscript R is for Red, L_i^B and R_i^B can be explained accordingly. T_i is the time spent for ith attack. When the time goes on, an occurred attack will have less and less influence on the morale. As a result, and to save the computation resources, when its influence decays to a considerable sufficiently small, we will not calculate it any more in later time stages.

Similarly, Table 2 only serves as the initial raw attribute pool and the raw attribute pool will be dynamically updated, for no one knows exactly what would happen next stage. An abnormal insurgent might even choose to kill only the driver who drives a car with a even number on the plate, while such attribute is usually considered as “minor” and “irrelevant” to terrorism. When the war procedure progresses and when time is allowed, abnormal situations will be identified and studied so that new attributes can be added to the raw attribute pool.

The feature selection procedure is shown in Fig 3, which will automatically selects features/preferences/attributes for future event probability prediction. To avoid temporarily removing some known important features from the key feature set, the key feature set consists of two parts: Reserved Feature Subset and Selective Feature Subset. Reserved feature subset is composed of important features that should not be ignored at any time. Selective feature subset consists of the features automatically selected by the feature search algorithm stated in the “Feature selection” dashed block.

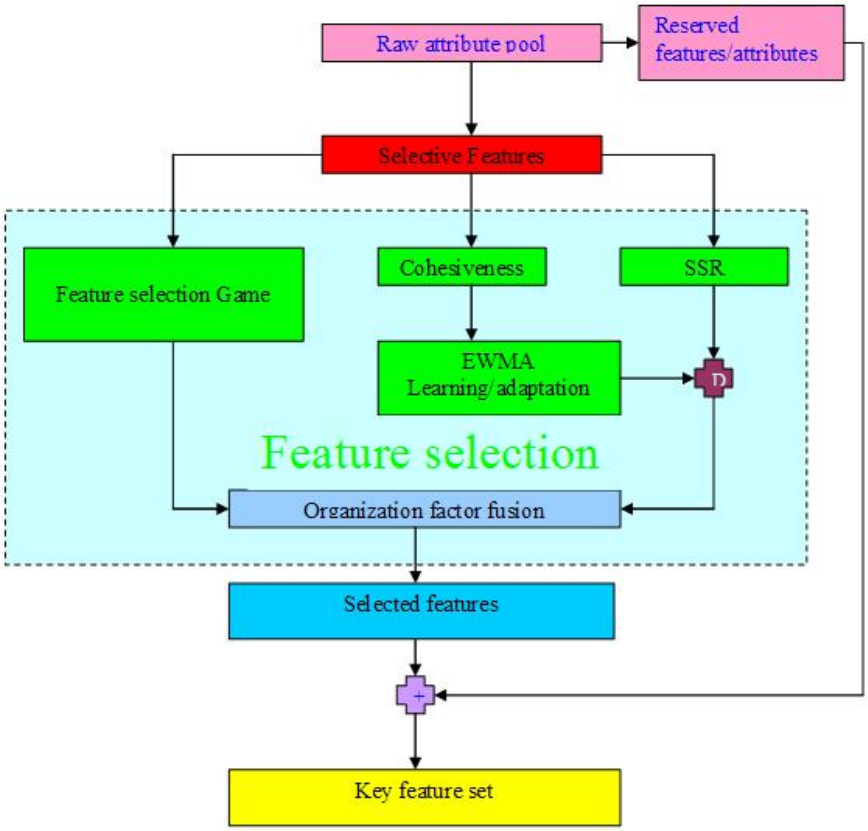


Fig. 3. Feature Selection

The feature selection game module will model the learning/rationalizing procedure of players and produce the probabilities that different features will be selected by players in next stage. For k th feature, its output is denoted as $P\ (k,t_m)$, which means the predicted probability that the k th feature is applied for attacks by Red at the time t_m. Details of feature selection game can be found in [21]. The SSR (Regression Sum of Squares) module will produce the importance describing a feature's capability to predict the whole existing data set if the prediction is solely based on this feature. The result of SSR calculation will be fused with the result of cohesiveness and EWMA adaptation and produce the probabilities as if the battle data is absolutely non-organizational (or random). Fusing the probabilities for rational situations produced by feature selection game and the non-organizational probabilities produced by SSR, cohesiveness, and EWMA (Exponentially Weighted Moving Average technique) learning/adaptation, the selected features can be prioritized.

For feature selection game module, see [21], of which we will not repeat the details here. In addition, SSR's definition and calculation can be found in every textbook about statistics theory. In the following paragraphs, we will focus on some technical details of our strategies about cohesiveness, EWMA, and fusion techniques.

For cohesiveness calculation, we modified the traditional algorithm so that it is more suitable for feature selection than the traditional approach [11]. Denote the distance

between two events i and j in the feature/attribute space as d_{ij}. The corresponding similarity is calculated as $a_{ij} = 1/(1+\alpha d_{ij})$, where α is the reciprocal of the average inter-event distance. The term used to measure feature cohesiveness is defined as

$$V_g = \frac{2\sum_{i=1}^{n-1}\sum_{j=i+1}^{n} b_{ij}}{n(n-1)} \tag{3}$$

Where b_{ij} can be defined as $b_{ij} = 4a_{ij}(1-a_{ij})$. n is the number of occurred events. Usually smaller values of V_g implies better capability of the feature set for defining the point pattern. Considering the prior distribution usually deviates from uniform distribution, the adjusted value of V_g, denoted as $V_g^{(k)} = V_g(E_k)/V_g(P_k)$, where $V_g(E_k)$ and $V_g(P_k)$ are the V_g scores for the event feature data set E_k and the prior feature data set P_k respectively, is a more realistic measurement of capability.

Note: we do not simply discard features as Liu and Brown [11] do for "some features that do not exhibit enough variation in the event feature data set". In [11], Liu and Brown calculate

$$r_k = \frac{\max_{x_{ik}, x_{jk} \in E_k} \left| x_{ik} - x_{jk} \right|}{\max_{x_{ik}, x_{jk} \in P_k} \left| x_{ik} - x_{jk} \right|} \tag{4}$$

and discard features which has sufficiently small r_k. This is because from the view of linear prediction or a traditional point-model, this feature might not be easy to apply. However, if about one feature, for a long time almost all events are fallen in a very narrow interval, it should be said that this feature has prediction capability since if the feature value is in that interval, the related event-occurring probability will be high otherwise is low. Actually in some sense, such features might be the most useful attributes since event data exhibit strong concentrations about it, which would produce predictions with high confidence.

Our approach is the combination of the game theory approach and the spatial-temporal prediction approach. As a result, it is applicable to apply such kinds of features. For this reason and to save calculation time, we choose to select features according to the overall result of the feature selection block in Fig. 3, thus will not discard a feature/attribute simply according to this r_k rule.

After calculating the V_g values, Exponentially Weighted Moving Average technique (EWMA) can effectively identify the on-going features and learn future trends. Before each step that will calculate the transition density functions, an additional step which will choose the key feature set specifically for this step is added and will produce I_g values for all features in the space corresponding to this time point. Denote the adjusted V_g value calculated at time point t_m for feature index k as $V_g^{(k)}(t_m)$. Define

$$z^{(k)}(t_m) = \lambda \sum_{j=0}^{\infty} V_g^{(k)}(t_j)(1-\lambda)^j \tag{5}$$

where λ is the exponentially weighting factor. When $z^{(k)}(t_m)$ goes out-of-control lines calculated via probability theory, it implies that there might be a pattern change or feature variation. Via simulations, a set of feasible λ can be determined according to different priorities such as fast detection of trend/variation, reducing false alarms, or optimal balancing. At the start-up stage, $z^{(k)}(t_1)$ is set as $V_g^{(k)}(t_1)$. The reason why we choose EWMA is that EWMA is very insensitive to the common assumption "normality" which is abused by many researchers. A second reason is that EWMA is notably good at supervising small trends and variations under situations with dynamics.

To fuse the results from cohesiveness (adjusted via EWMA) module and the SSR module, we apply the following formula

$$P^2(k,t_m) = SSR(k,t_m) / Z^{(k)}(t_m) \tag{6}$$

The higher the $P^2(k,t_m)$ is, the better the k th feature is for prediction at the time t_m. The character "D" in the fusion cross (right side in Fig. 2) means "division".

For the fusion between the fused output $P^2(k,t_m)$ and the feature selection game output $P^1(k,t_m)$, we apply the following formula (equation (7)). The higher $P(k,t_m)$ is, the higher the probability that this feature will be selected by Red in next stage.

$$P(k,t_m) = r \square P^1(k,t_m) + (1-r) \square P^2(k,t_m) \tag{7}$$

Where r is the organizational factor, which can be determined as

$$r = \frac{p_1}{p_1 + p_2} \tag{8}$$

Where p_1 and p_2 are the p-values (calculation of p-values can be found in any statistics textbook) of the following two hypothesis tests (HT), respectively

HT1: $\begin{cases} H_0: & \text{In records past features are selected by Red solely according to game} \\ H_a: & \text{In records past features are not selected by Red solely according to game} \end{cases}$

HT2: $\begin{cases} H_0: & \text{In records past features are selected by Red absolutely randomly} \\ H_a: & \text{In records past features are not selected by Red absolutely randomly} \end{cases}$

In other words, r is a coefficient reflecting the extent of the Red players' abiding by the results of games. This is because when Red consists of many non-organizational agents, there might not be significant planning and coordination between them, which makes the occurrences look like random results. In addition, when the number of agents and events increase, the results will be more and more like a random distribution. This is also the basis of existing traditional probability prediction approaches. However, if Red agents do have strict organization and do have perfect coordination and planning, no matter how large their number is, their attacks will not show enough randomness since each of their attacks is a result of a purposeful selection, not a result of random assignment. In addition, r is also a sign of how closely related the future pattern and existing pattern are. It somewhat likes a derivative factor (which has the capability of "predicting") in PID (Proportional-Integral-Derivative) controllers, although not exactly the same.

Selected features will be put into the inner core of the ontology which stores the features and the corresponding structures/relationships among them. A brief language-independent diagram about the dynamical feature ontology maintenance is illustrated in Fig. 4.

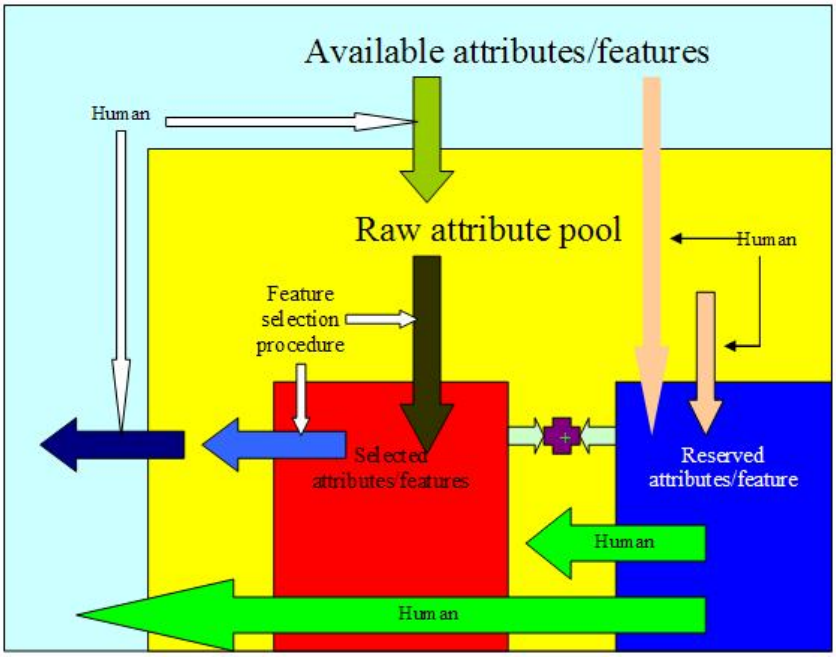


Fig. 4. Language-independent feature ontology illustration

Generally when inspecting a feature and checking whether it should be chosen or dropped, an inner-layer feature will have priority. This is because for most cases inner-layer features are more "important" and there are also some hardships when transferring from one feature set to another feature set, which results in some "adhesiveness" between different stages. For outer-layer features, they will be inspected in following situations:

(1) Orders from human commander (if any) [Level 5 fusion – User refinement].
(2) The system has enough leisure time after computing inner-layer features.
(3) Existing inner features can not model the behaviors accurately enough.

Note that although the system can run without human participation once initialized, a human commander always holds the highest authority for every step. If a human commander wishes to inspect a feature that is at the very outer layer, that feature would be checked before any inner-layer features. When there is no human guide, the system will dynamically update the outer-layer features in the later two cases.

Based on the dynamic feature ontology, spatial-temporal models can be built up. Kernel density function and estimation techniques [For details, see 11-17] are applied with an important correction. That is, when it is observed that the event initiators are intelligent and will not try to initiate a new event right after last occurred event at exactly the same location, we will choose the following formula (9) as the kernel density function instead of traditional Gaussian functions [14-15, 19] or Exponential functions [11-12].

$$f_{ig} = \begin{cases} w_1 e^{-\lambda_i^1 (D_{ig} - (D_{in} + l_i))}, & \text{when UpperLimit} \geq D_{ig} \geq (D_{in} + l) \\ w_2 e^{-\lambda_i^2 ((D_{in} + l_i) - D_{ig})}, & \text{when LowerLimit} \leq D_{ig} \leq (D_{in} + l) \end{cases} \quad (9)$$

Where subscript g stands for the map grid index, n is the index of past event, D is the feature value, i is the index of feature, and w_1, w_2, λ_i^1, λ_i^2 are subject to the following restriction:

$$\int_{LowerLimit}^{UpperLimit} f_{ig} dD_{ig} = 1 \quad (10)$$

l_i is the most likely distance between the location at which an event just occurred and the location at which a next location would happen. The initial value of l_i is assigned by experienced commander and later l_i can be automatically updated according to statistical event data obtained in the process of long-time battle. When the feasible feature interval is wide enough and l_i is not close to the ending points of the interval, to save calculation resources, the UpperLimit and LowerLimit can be set as $+\infty$ and $-\infty$, respectively.

On the basis of the kernel functions, spatial-temporal prediction model can be built and probability predictions can be produced. The predictions the prediction results of Markov game will be fused according to the following formula

$$P_e(c, t_m, x, y) = r \square P_e^1(c, t_m, x, y) + (1 - r) \square P_e^2(c, t_m, x, y) \quad (11)$$

Where $P_e^1(c, t_m, x, y)$ is the prediction result from Markov game, $P_e^2(c, t_m, x, y)$ is the prediction result from spatial-temporal model, c is the even type index, t_m is the time, x and y are the coordinates, r is determined similarly as the feature selection organizational factor. For the Markov gaming details, see [18].

3. Simulations and Experiments

The final comprehensive probability prediction results (probability maps) in a long duration battle (which can be divided to three time-continuous stages) can be demonstrated in following figures (Fig. 5, Fig. 6, Fig. 7), which are based on the scenario described in Fig. 2. Indices of these three probability prediction maps are arranged in time sequence. All the strategies discussed in Section 2 are fused according to Fig. 1 to produce the ECOA threat probabilities over city districts. Over the time horizon, newly occurred events are fed to the system to update the identified and/or predicted event features/patterns, and finally update the probability predictions.

In this simulation run, the corresponding results of dynamic feature selection are listed in the following Table 3. The feature indices are defined in Table 2. Note that the number of selected features might change according to time thus might not be the same.

Table 3 Feature selection results for different time stages

Fig. 5	Fig. 6	Fig. 7
Feature 1	Feature 1	Feature 1
Feature 7	Feature 7	Feature 7
Feature 16	Feature 16	Feature 23
Feature 23	Feature 24	Feature 24
Feature 24		

In this simulation, we assume feature 1 and feature 2 are in the reserved feature set thus they are always selected. At the second stage feature 23 is temporarily removed thus feature 16 (distance from soldier stations) has a stronger influences on the probability map. This is reflected in Fig. 6 in which soldier stations have larger and clearer "safe boundaries" compared to Fig. 5. Note that feature 24 (the morale of Red) is always selected. Since the Blue Force successfully assigns the soldier/policemen/weapon resources, in Fig. 7 the Red insurgents have lower morale, which is reflected as a general lower probability to have a event for most location. The general scope of the river, which is generally not a favorite site for attacks for various reasons, is also reflected in all three maps. However, it is still possible to have an attack on the river, which means it might occur on a bridge, a boat, etc.

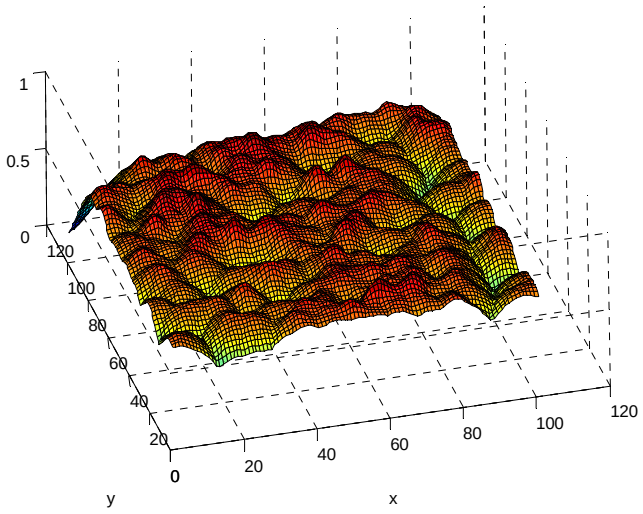


Fig. 5 Probability Map for stage 1

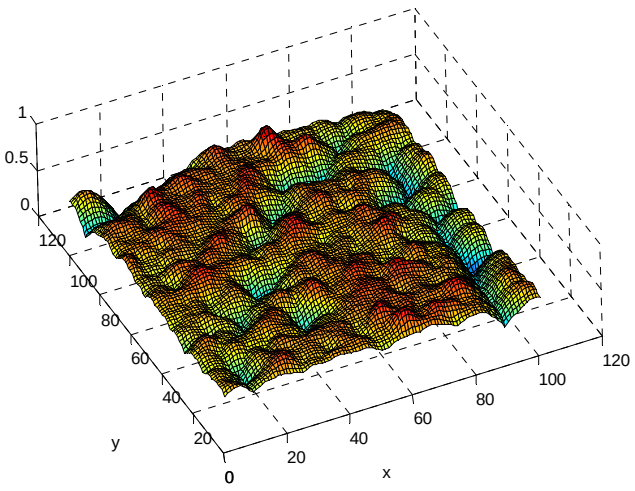


Fig. 6 Probability Map for stage 2

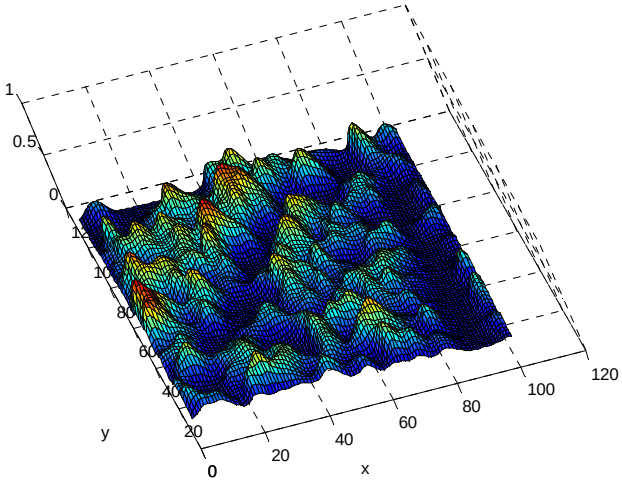


Fig. 7 Probability Map for stage 3

A simulation about event-occurring is shown in Fig. 8. We took the time scope of the last stage. The five kinds of marks in Fig. 8 correspond to different types of attacks, which is stated in the following Table 4.

Table 4 Symbols for attack types

Index	Description	Symbol
1	Gun Fighter/Mortar/Small Arms	Blue star
2	IED (Improvised Explosive Device)	Green circle
3	Kidnap/Hijack	Red Plus
4	Robbery/Stealing	Cyan square

5	“Dirty” bomber/Bio-attacks	Yellow triangle (down)

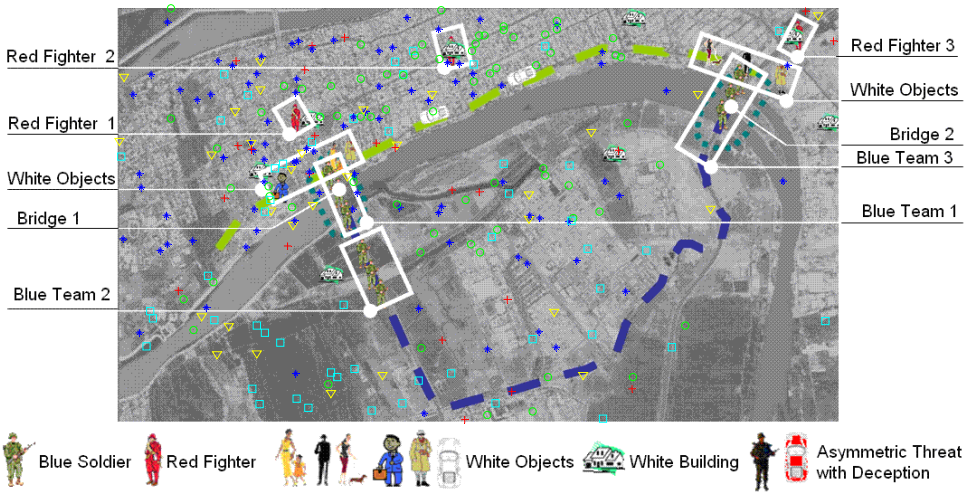


Fig. 8 The event occurrences illustration

4. Conclusions

In this paper, we refined existing spatial-temporal point-model prediction techniques, combine them with dynamic feature selection game model, and produce dynamic probability predictions. The features are stored in a language-independent ontology. We modeled and studied the learning/adaptation of behavior features and changing emotional factors in hostile environments. The advantages of our proposed algorithm and architecture are demonstrated in a typical asymmetric urban warfare scenario.

Acknowledgements

This research was supported by the US Navy under contract number N00014-06-M-0237. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Navy.

References

[1] Genshe Chen and Jose B. Cruz, Jr., "Genetic Algorithm for Task Allocation in UAV Cooperative Control," *AIAA Conference on Guidance, Navigation, and Control*, August 2003.

[2] Jose Cruz, Genshe Chen, Denis Garagic, Xiaohuan Tan, Dongxu Li, Dan Shen, Mo Wei, Xu Wang, "Team Dynamics and Tactics for Mission Planning," *Proceedings, IEEE Conference on Decision and Control*, December 2003.

[3] V.R. Saksena and J.B. Cruz, Jr., "An Approach to Decentralized Control of Large Scale Systems Using Aggregation Methods," *Proc. 23rdrd IEEE Conf. on Decision and Control*, Las Vegas NV, December 1984, pp. 198-201 (Invited).

[4] V.R. Saksena and J.B. Cruz, Jr., "Optimal and Near-Optimal Incentive Strategies in the Hierarchical Control of Markov Chains," *Proc. 1983 American Control Conf., San Francisco CA*, June 1983, pp.1103-1108 (Invited).

[5] J.B. Cruz, Jr., "Survey of Leader-Follower Concepts in Hierarchical Decision-Making," *Proc. Fourth International Conf. on Analysis and Optimization of Systems*, Le Chesnay, France, December 1980, pp. 384-396 (Invited).

[6] Jose B. Cruz, Jr., Genshe Chen, Dongxu Li, and Denis Garagic, "Target Selection in UAV Cooperative Control Under Uncertain Environment: Genetic Algorithm Approach," *WSEAS Transactions on Circuits and Systems*, Vol. 3, Issue 3, pp. 542-553, May 2004.

[7] Milligan, G. J., & Cooper, M. C., "An examination of procedures for determining the number of clusters in a data set," *Psychometrika*, 1985.

[8] Mojena, R., "Hierarchical grouping methods and stopping rules: An evaluation," *Computer Journal*, 1977.

[9] Mo Wei, Genshe Chen, Jose B. Cruz, Jr., Leonard S. Haynes, Martin Kruger, Erik Blasch, "Game Theoretic Modeling of Military Air Operations with Retaliatory Civilians," *IEEE Aerospace Conference 2007*, Accepted, to be presented at Big Sky, MT, March 2007.

[10] Mo Wei, Genshe Chen, Jose B. Cruz, Jr., Leonard Hayes and Martin Kruger, "Battlefield Modeling of Partially Emotional Civilians," *Decision and Support Systems*, Accepted in 2006.

[11] Liu, H. and Brown, D.E., "Criminal incident prediction using a point-pattern-based density model", *International Journal of Forecasting*, 2003.

[12] Fiksel, T., "Simple spatial–temporal models for sequences of geological events", *Elektronische Informationsverarbeitung* (1984), und Kybernetik, 20, 480–487, 1984.

[13] Block, C., "STAC hot-spot areas: A statistical tool for law enforcement decisions", In Block, C. R., Dabdoub, M., & Fregly, S. (Eds.), *Crime analysis through computer mapping*, Washington, DC: Police Executive Research Forum, p. 20036, 1995.

[14] J efferis, E., "A multi-method exploration of crime hot spots", *Presentation at the Annual Meeting of the Academy of Criminal Justice Sciences*, Albuquerque , NM , March 10– 14, 1998.

[15] Levine, N., "'Hot Spot' analysis using CrimeStat kernel density interpolation", *Presentation at the Annual Meeting of the Academy of Criminal Justice Sciences*, Albuquerque ,NM ,March 10– 14, 1998.

[16] Amir, M., *Patterns in forcible rape*, Chicago, University of Chicago Press, 1971.

[17] Baldwin, J. and Bottoms, A., *The urban criminal : A study in Sheffield*, London, Tavistock Publications, 1976.

[18] Genshe Chen, "A Game Theoretic Approach for Threat Prediction and Situation Awareness", *available at http://www.navysbir.com/05_s/31.htm*.

[19] Donald Brown, Jason Dalton, and Heidi Hoyle, *Spatial forecast methods for terrorist events in urban environments*, Springer-Verlag Berlin Heidelberg, 2004.

[20] Kutner, Nachtscheim, and Neter, *Applied Linear Regression Models*, 4th Edition.

[21] Mo Wei, Genshe Chen, Jose B. Cruz, Jr., Leonard Hayes, and Martin Kruger, "Game Theoretic Feature Selection in Crime Prediction," *SPIE Defense and Security Symposium*, Submitted in 2006.

[22] Kathleen M. Carley, "OrgAhead: A Simulation of Organizational Adaptation," *available at http://www.casos.cs.cmu.edu:80/projects/OrgAhead/index.html*.

[23] David M. Kreps and Robert Wilson, "Sequential equilibria," *Econometrica*, 50(4):863–894, 1982.

Applying Spatial-Temporal Model and Game Theory to Asymmetric Threat Prediction

Authors:
Mo Wei - Intelligent Automation, Inc.
Genshe Chen - Intelligent Automation, Inc.
Jose B. Cruz, Jr. – Cruz & Associates
Leonard Haynes - Intelligent Automation, Inc.
Martin Kruger - The Office of Naval Research
Erik Blasch - AFRL/SNAA

Presenter: Dr. Genshe Chen

Outline

- Motivations:
  - Why do we fuse game modeling to feature-related prediction
    - Why do we need features in prediction
    - Why do we need to apply game modeling
- Technical approach
  - The Model: How do we apply game theory
  - The Model: How do we use features in prediction
  - How do we fuse them
- Simulations and Explanations
- Conclusions

Why do we need features

- Very early prediction models (Model Type I)
  - Calculate the crime frequencies
- Later models (Model Type II)
  - Analyze possible crime preferences or features, such as pop. Density, income per cap, distance to police station, etc.
  - Fuse such analyses in prediction, typically
    - Statistically summarize features
    - Statistically apply features in probability models
  - Achieved great improvement on accuracy of city district crime predictions
- Thus features can greatly refine the predictions.

Why do we need game theory

- When the enemies are unorganized and non-intelligent, the occurrences of Course of Action (COAs) will be somewhat independent, which might make the whole scenario fit some probability models.
- However, if the enemy is a well-structured and has an intelligent organization, the scenario will be largely different.
  - Intelligent enemy's behavior might not have strong randomness.
  - The enemy might purposely choose COA time and site, perform such COA, calculate the loss and gain of last stage, then determine the next stage's action.
  - If necessary, they might even choose a different site for every stage, which will not display any traditional "geographical preference".
- That is, intelligent enemy might (suddenly) change preferences or behavior features

Why do we need game theory (continued)

- Model Type II assumes that the features are fixed once they are identified
  - If via past statistical data it is found that “the distance to gas station” is an effective feature, this feature will always be taken into account even if later terrorists change their pattern so that “the distance to a school” becomes the new effective feature for predicting their behaviors.
  - The reason why terrorists wish to change their behavior pattern is that they find that policemen already notice the old feature and prepare for it thus continuing old Course of Action (COA) pattern will bring too high risks.

Why do we need game theory (continued)

- Thus Model type II can not efficiently deal with possible changes of COA features.
  - Even if the model is modified such that after each time step the effective features should be chosen again, there will still be significant delay in identifying such changes of features, for the method of identifying effective features is based on statistical data to date.
  - Only after the changes happen long enough is it possible to detect such changes, not to mention predicting such changes.

Why do we need game theory (continued)

- Applying game theory can help predict possible changes of features
  - This is because the basic logic of game theory is to predict ahead via all available information, including past data and possible choices at current stage.
  - It does not need to wait for the enemy's change happening first thus no delay.
  - In addition, such prediction is often self-enforcing due to the properties of Nash solutions thus is more trustable
- Via game theory, it can be anticipated that surprise attacks will be reduced for many surprise attacks will not be surprise attacks under the new prediction technique fused with game theory.

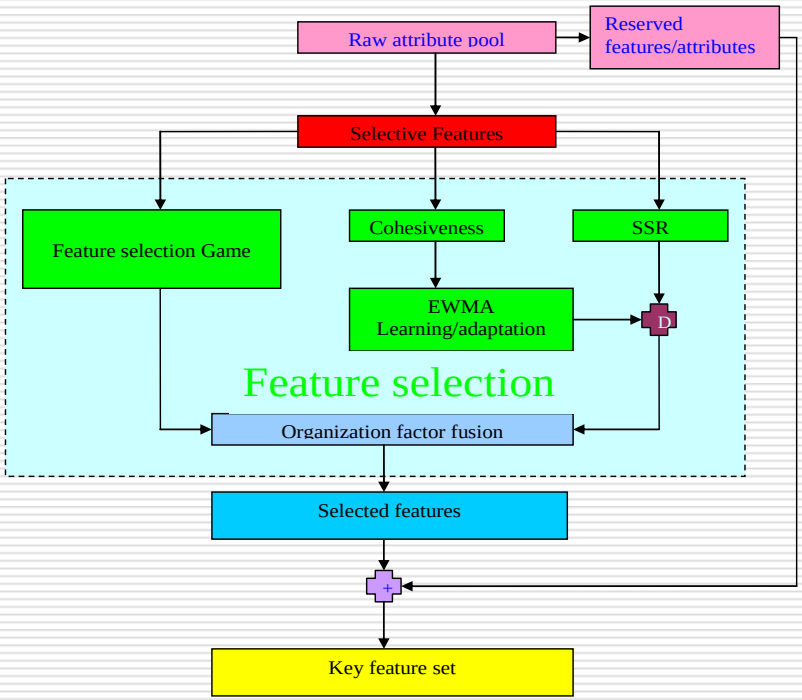
Our Technical Approach

- Our approach is a combination of game theory approach and spatial-temporal prediction approach.
  - Feature selection Game

    provides a prediction for the future active features that a player would choose
  - Kernel probability functions

    improve the model about possible actions of non-organizational insurgents.

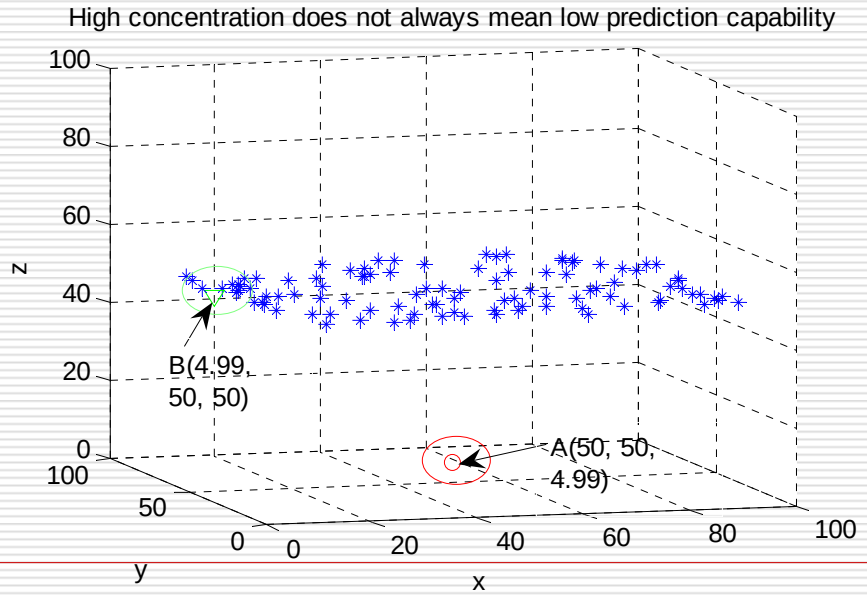
Advanced Hybrid Feature Selection Approach



- The feature selection procedure is used to automatically select features/preferences/attributes for future event probability prediction.
- To avoid temporarily removing some known important features from the key feature set, the key feature set consists of two parts:
  - Reserved Feature Subset
  - Selective Feature Subset
- Reserved feature subset is composed of the very important features which should not be ignored at any time.
- Selective feature subset consists of the features automatically selected by the feature search algorithm stated in the “Feature selection” dashed block.

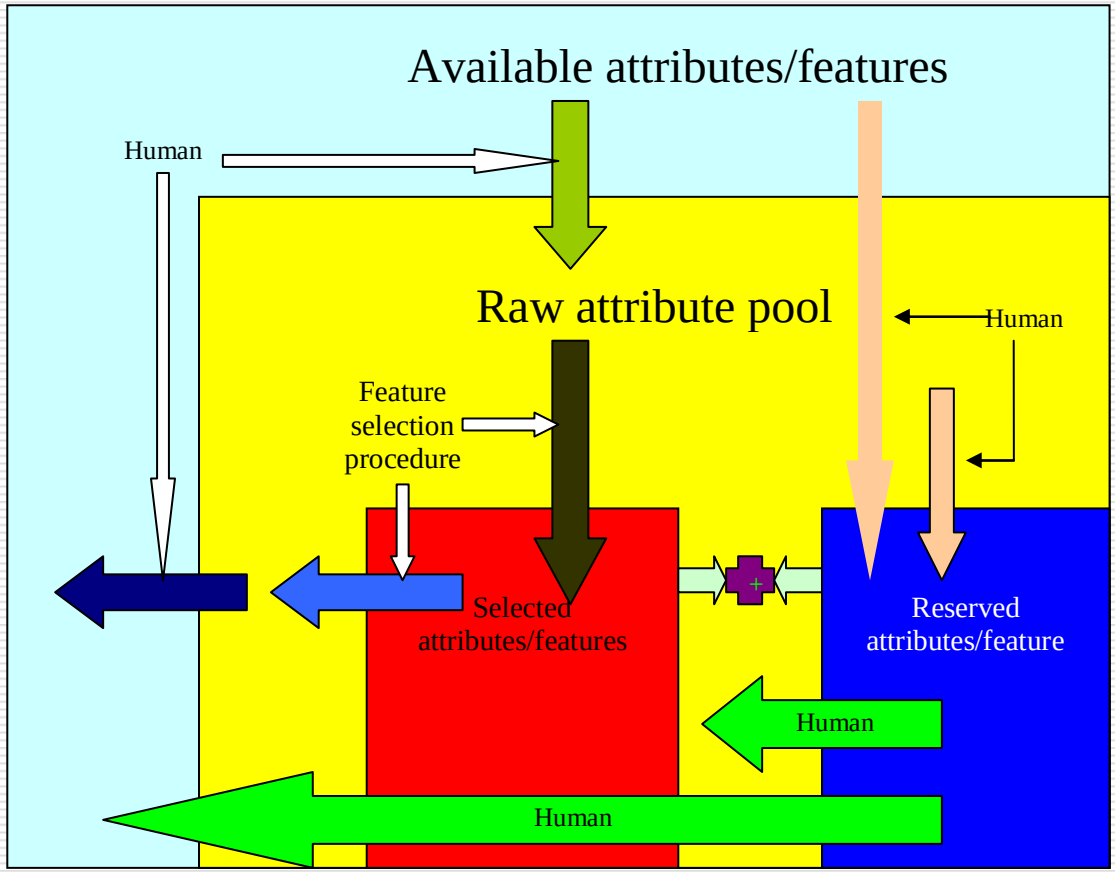
Cohesiveness Calculation

- For cohesiveness calculation we slightly modified the traditional algorithm so that it is more suitable for feature selection than traditional approach.
- We do not simply discard features as previous researchers did for “some features that do not exhibit enough variation in the event feature data set”.



Feature Storage

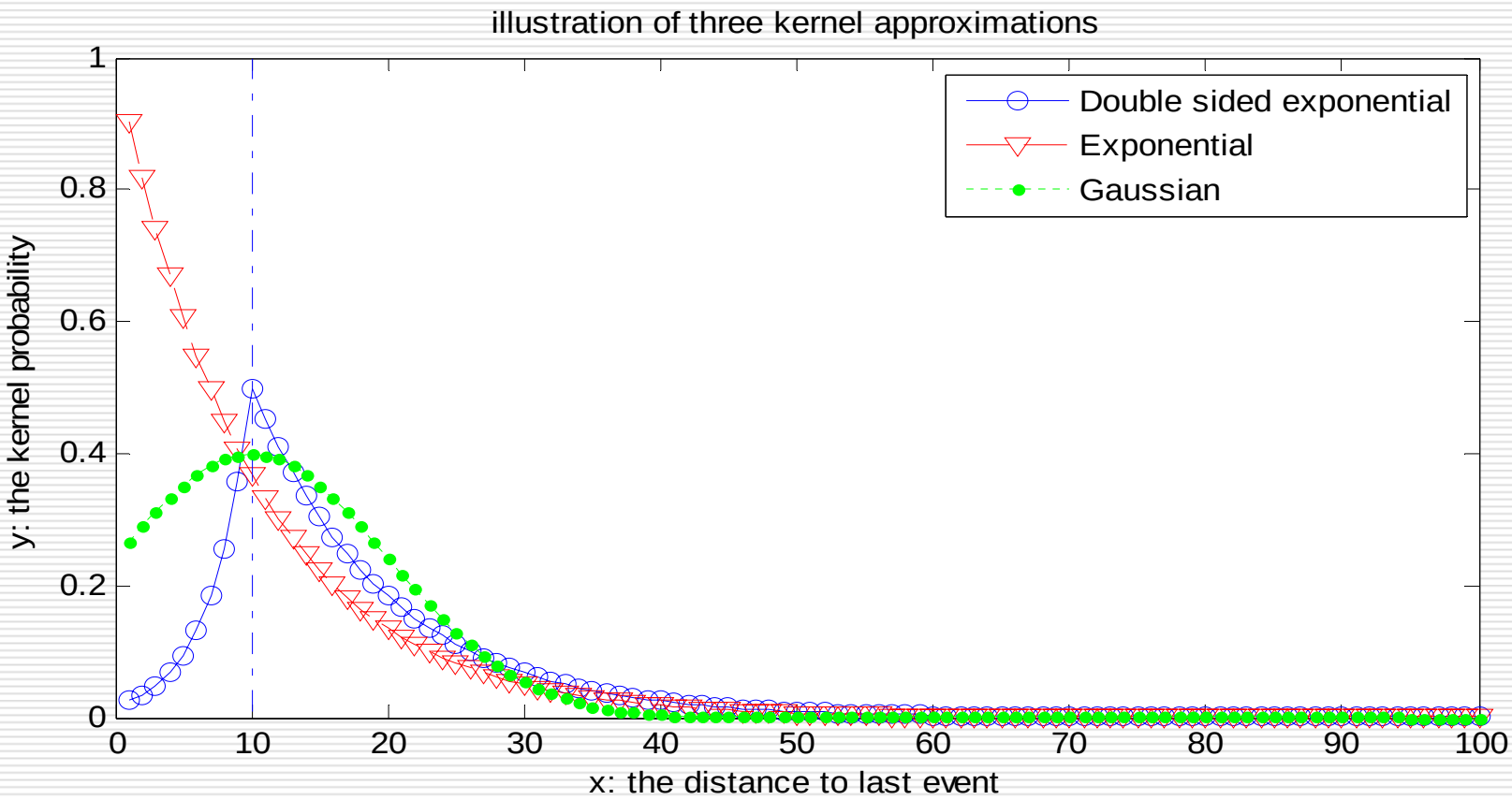
- Selected features will be placed in the inner core of the ontology which stores the features and the corresponding structures/relationships among them.



Advantages over traditional probability approaches

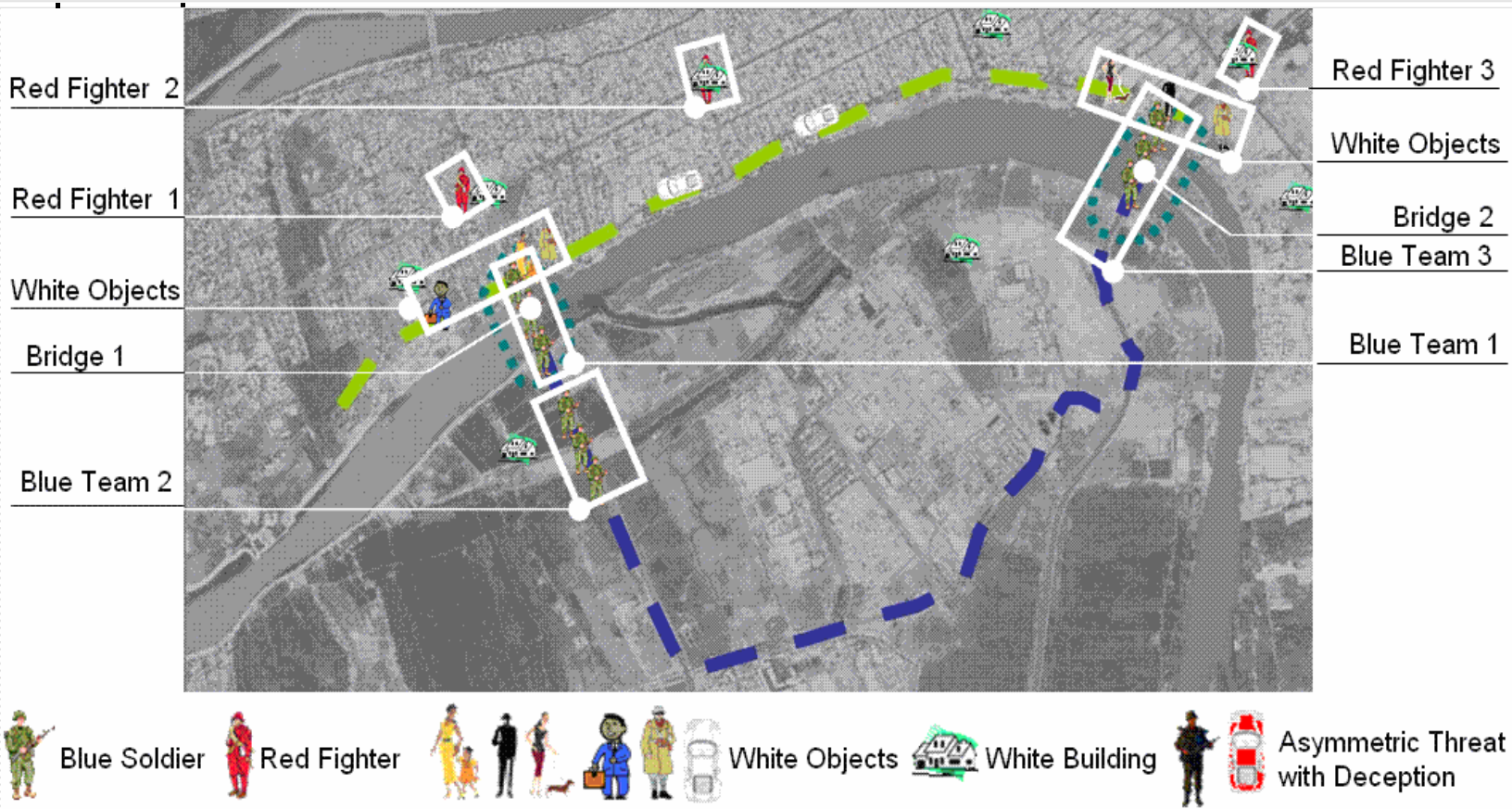
- Do not need to discard features "that do not exhibit enough variation in the event feature data set"
  - Such features are not very convenient for traditional probability estimation approach
  - However, high concentration does not necessarily mean low prediction capability
  - Our game method can make use of such features
- Refined kernel probability functions in estimation
  - Problems for Traditional Gaussian distribution approach (use last event feature value as the center-point)
    - Event distribution might be severely asymmetric
    - Many feature values are even one sided
  - Problems for traditional exponential distribution approach (use last event feature value as the starting-point, then decreasing)
    - An intelligent attacker would intuitively avoid exactly the same location/time/features
    - Thus last time's feature values do not mean the highest possibility
  - Our approach: when facing such problems, use double-sided exponential kernel distribution as the kernel probability functions.

Illustration of the approaches



Simulation: Urban Warfare Scenario

- In a typical urban warfare scenario (shown below), we intend to illustrate our dynamic adaptive hierarchical game theoretic approach for modeling and prediction of asymmetric threat



Urban Warfare Scenario

- The blue force's missions are to try their best to secure the whole area, including the urban districts, bridges, mains roads and blocks. The blue ground force consists of teams of soldiers/policemen each with small arms.
- The red force (terrorist and/or insurgent forces) includes armed fighters and some asymmetric adversaries hiding in and acting like the white objects (the civilians).
- When the battles are long-lasting and the battlefields are heavily populated by civilians, civilians sometimes play important roles in battles.
  - Civilian interest: desire/enforcement about "participation"
  - Civilian intelligence: capability about "participation"
  - Biased civilians can affect COA success probabilities: asymmetric information, asymmetric buildings, asymmetric provisions, etc.

Urban Warfare Scenario: Detailed Strategy description

- In urban scenario, we predict the changes in enemy strategies before such changes are fully implemented.
- We present a primitive prediction of ECOAs by following the pattern/feature recognition model.
- Based on such prediction, some associated best response strategies of the Blue side can be recommended.
- If the primitive prediction is almost correct, there are two possible response strategies for the blue force according to different goals.

Urban Warfare Scenario

- If the purpose of blue force is to stop the red forces' actions, the recommended COA of the blue force is to publicly send a message to the red forces, and suggest that their actions are in the control of the blue side. As a consequence, probably the red forces will change their proposed actions.
- However, if the purpose of the blue force is to set up a trap and catch them so that in the long run the total number of red attacks will go down, the blue force can only maneuver secretly.
- In such cases not only might the red use deceptions, the blue might also use some counter deceptions.
- If the first guess is incorrect (for example, the attack pattern might be new and unknown), our game theoretic data fusion module and dynamic learning module will dynamically refine the primitive prediction and update the feature/pattern records.

Urban Warfare Scenario: features

- First classify and identify different ECOAs into a small number of types of surprise attacks with associated features.
- Only after deciding which type of attacks will likely occur at some next stage with what probability, can we develop an appropriate resource allocation algorithm.
- Considering information from different resources (papers, newspapers, reports from Department of Defense: Navy, Marines, Army, Air Force), typical surprise attacks are:
  - Type 1: Gun Fighter/Mortar/Small Arms
  - Type 2: IED (Improvised Explosive Device)
  - Type 3: Kidnap/Hijack
  - Type 4: Robbery/Stealing
  - Type 5: “Dirty” bomber/Bio-attacks

Urban Warfare Scenario: features

- In a broad sense, any possible attribute (or feature) might be related to another attribute, which means any attribute can serve as a potential feature or pattern.
- However, due to real world limits such as computation requirements, usually we can only choose some measurable, available, and “probably” related attributes and put them in a pool of “raw attributes.”
- In such a raw attribute pool there might still exist hundreds or even thousands of attributes, which would greatly exceed the computation capability of existing computer systems since each attribute will serve as a dimension, and when the number of attributes increases the computation will also increase.
- As a result, before associating features into the system, a much smaller key feature set should be dynamically selected from the raw attribute pool.

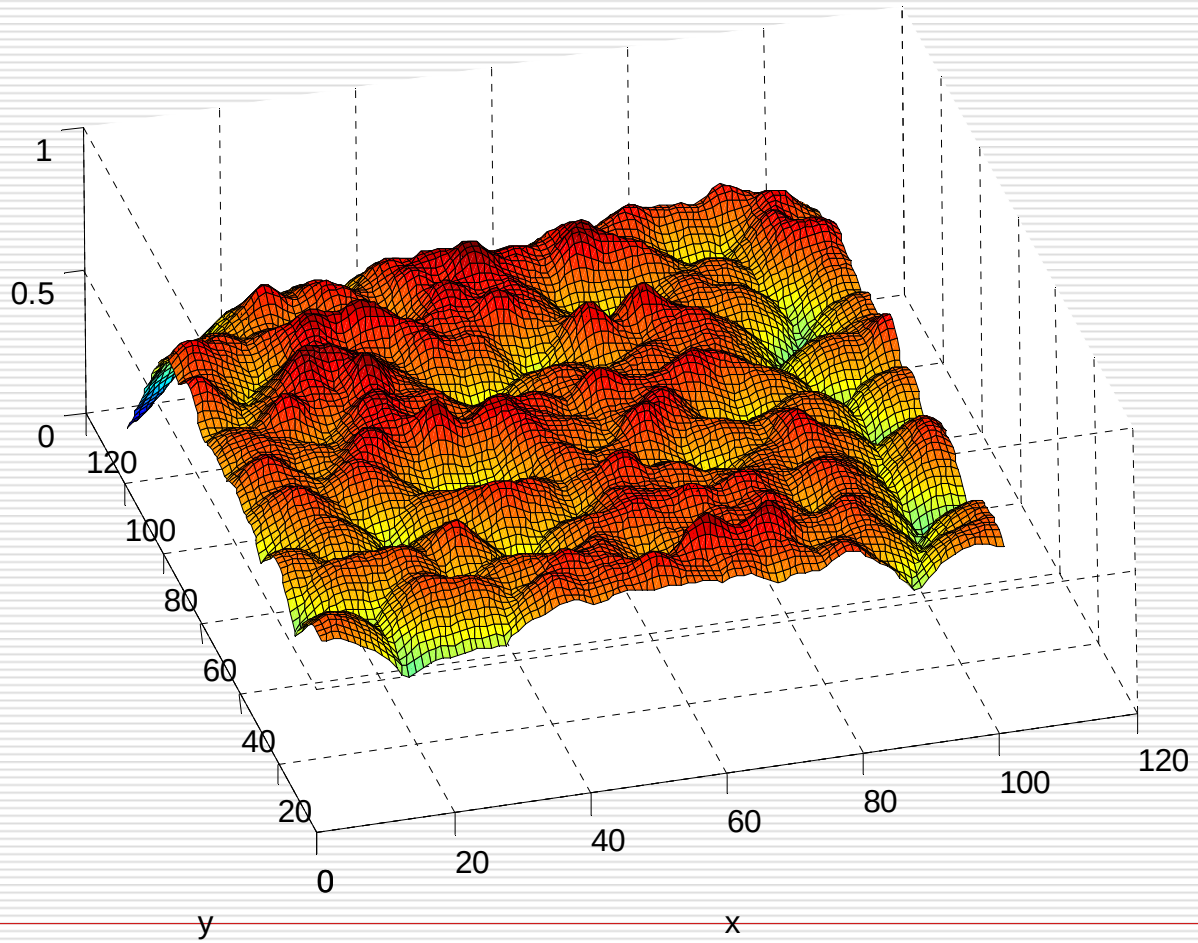
Partial List of Raw Attribute Pool: Example

- Population density per square mile
- Religious intensity
- Male people population density per square mile
- Average family size
- Young people (from 11 to 29) population density per square mile
- Average salary per year
- Average price of houses
- Ratio of children in school and out ofschool
- Percentage of people who were once involved in crimes
- Percentage of people who are in debt
- Average percentage of people who have children
- Distance to nearest soldier/policemen station
- Distance to nearest hospital
- Distance to nearest highway
- Distance to nearest church/school/library
- The time difference from the previous attack
- Distance to nearest location of previous attacks
- Morale of insurgents
- Average wellness of public utilities
- Distance to nearest desert/wood
- Average expenditure on alcohol beverages, tobacco, and smoking

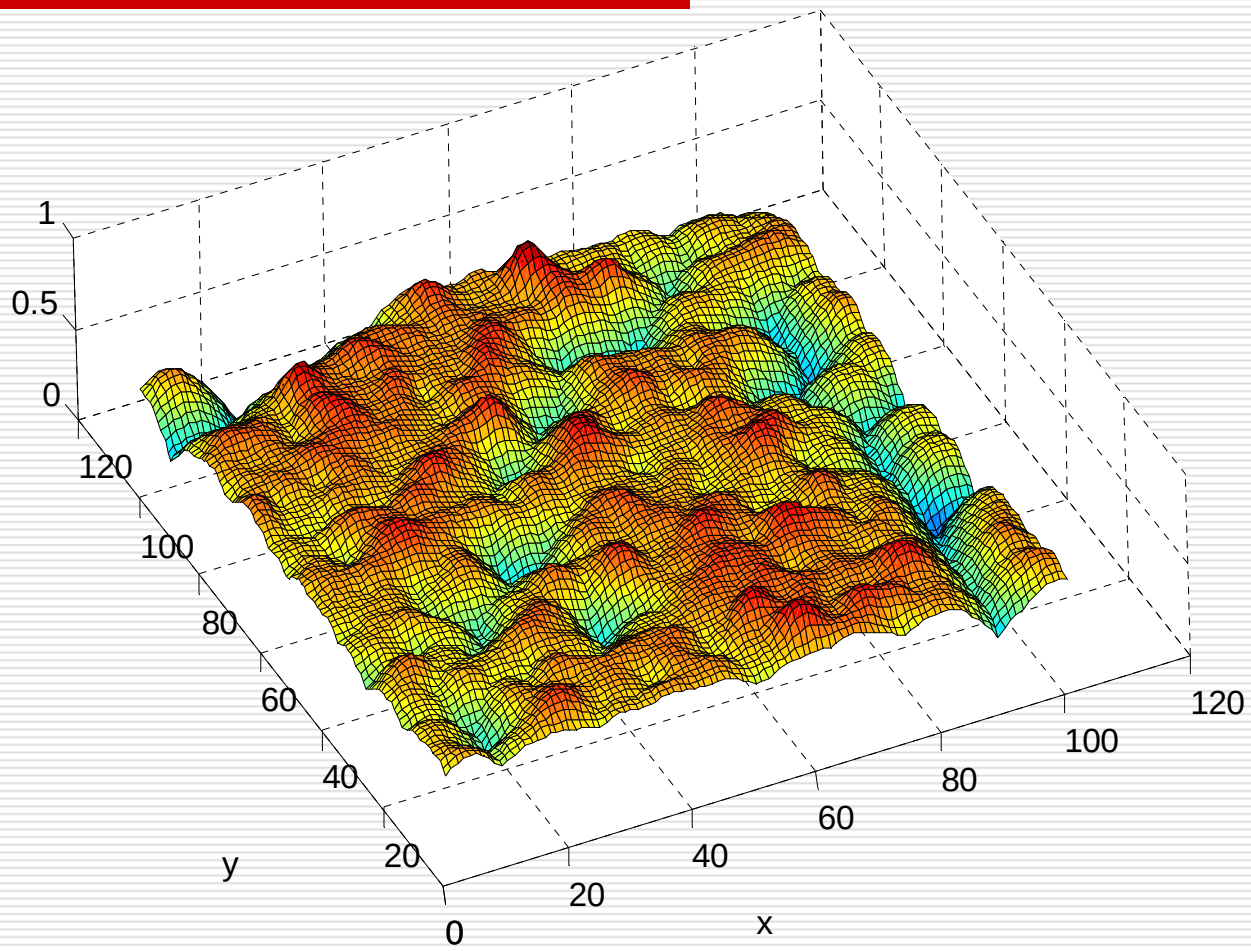
Simulation results for the scenario

- The final comprehensive probability prediction results (probability maps) in a long duration battle (which can be divided to three time-continuous stages) can be demonstrated in following figures.
- Indices of these three probability prediction maps are arranged in time sequence.
- All the strategies discussed are fused to produce the ECOA threat probabilities over city districts.
- Over the time horizon, new events are fed to the system to update the identified and/or predicted event features/patterns, and finally update the probability predictions.

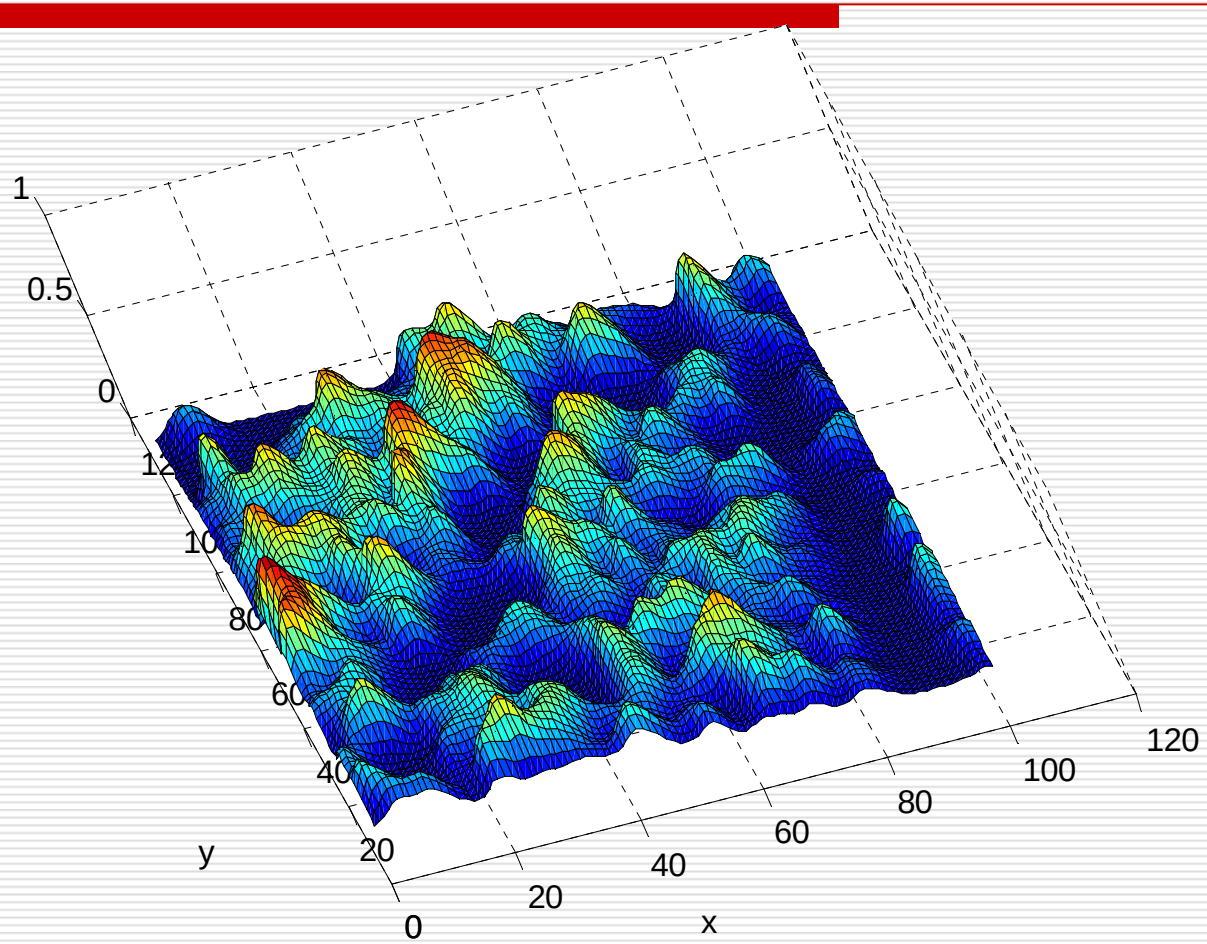
Simulations for Urban Warfare Scenario



Simulations for Urban Warfare Scenario



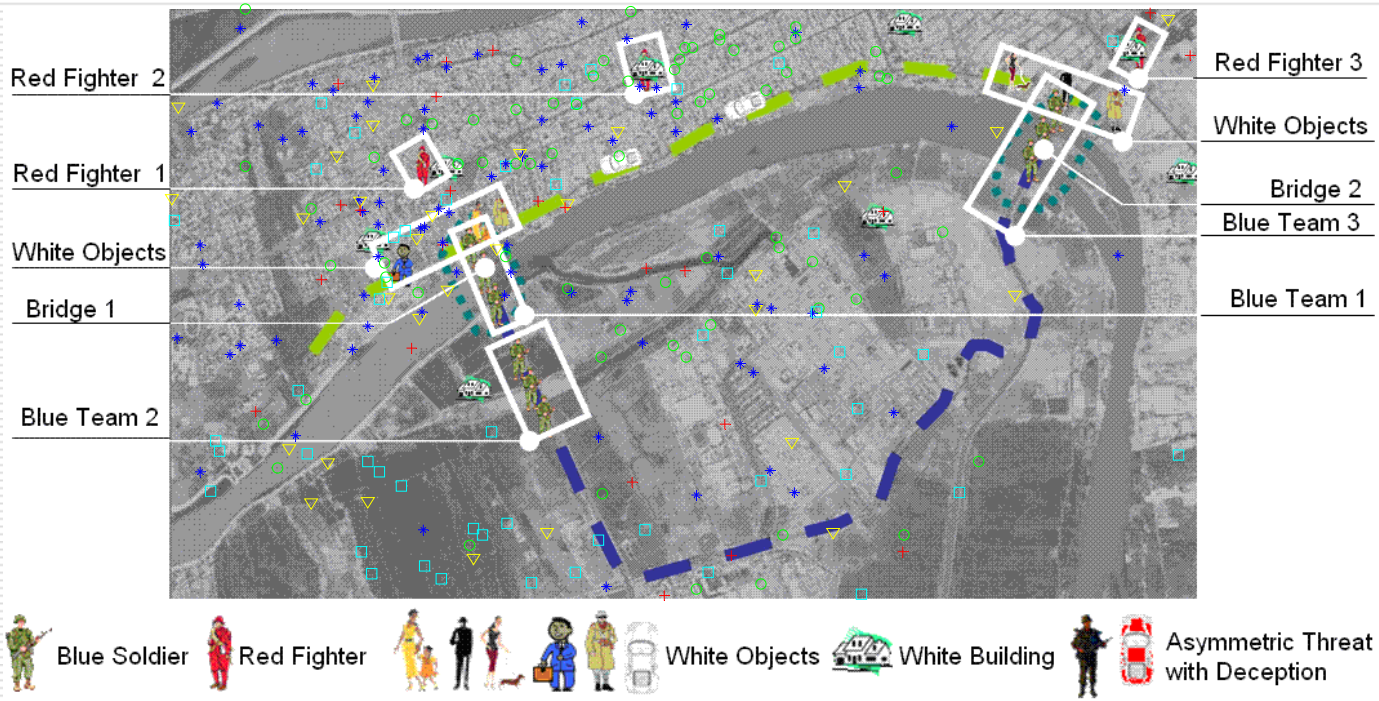
Simulations for Urban Warfare Scenario



Explanations

- The Red insurgents change their preferences during the battle.
- However, some important features such as population density and morale are always selected (in reserved feature set)
- With the help of the fusion of feature prediction and game theory, the Blue Force successfully assigns the soldier/policemen/weapon resources. In the last figure the Red insurgents have lower morale, which is reflected as a general lower probability to have a event for most location.
  - Note that the general scope of the river, which is generally not a favorite site for attacks for various reasons, is also reflected in all three maps.
  - However, it is still possible to have an attack on the river, which means it might occur on a bridge, a boat, etc.

Simulations for Urban Warfare Scenario



Conclusions

- We proposed a framework for asymmetric threat learning/adaptation detection and prediction.
- We proposed and refined advanced hybrid feature selection strategy.
- We fused Markov models with refined spatial-temporal point model prediction techniques to provide specific ECOA predictions.
- We implemented dynamic learning and adapting techniques and fused them within the ontology.
- We simulated dynamic predicting system in which enemies show learning/adapting abilities and various types of asymmetric course of actions (COA).

Acknowledgments

- This research was supported by the US Navy under contract number N00014-06-M-0237. The Program Manager is Mr. Martin Kruger.
- The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the Navy.