Information Fusion and Response Guidance

Timothy J. Dasey and Jerome J. Braun

The uncertain and disparate information sources needed to properly assess potential threats, and the relatively untrained and inexperienced users, make development of decision support technologies critical for full realization of the value of chemical and biological (CB) defense technologies. Lincoln Laboratory is pursuing several research and development efforts in decision support for CB defense. We discuss here fusion of information sources in the context of several example algorithmic efforts, and describe applications such as decision support for mail screening and detection of biological agents in a subway station.

» **Chemical and biological (CB) detection** technologies have become more mature, and some are now available for operational use. As the sensor availability has increased, so too has the realization that the interpretation of the sensor output is nontrivial. The sensor data can be coupled with non-sensor information about the operational and environmental conditions prevailing at the time of the sensor's output. We refer to this combined information—from sources such as intelligence, adversary or agent dissemination surveillance, weather conditions, and situational vulnerabilities—as contextual information. For example, a battlefield chemical sensor alert may be interpreted differently if additional contextual information indicates that there was an artillery shell that landed upwind of the sensor position in the recent past.

What is needed is information fusion—that is, the intelligent combination of multiple information sources to enhance the decision maker's understanding of the data and its implications. The mere combination of this information through networking, communications, databases, and displays does not constitute true information fusion. One category of fusion problems requires automated pattern recognition. Pattern recognition methods produce decisions about which one out of multiple possible hypotheses is true with respect to a given object or event. One such hypothesis space is the question of whether a biological or chemical agent is present in the environment. Decision makers lack experience responding to CB contamination events, and so could use automated assistance in choosing the appropriate course of action. We define this assistance as response guidance.

By the definitions above there are few if any information-fusion or response-guidance solutions available

in either military or civilian settings for CB defense. The Joint Warning and Reporting Network (JWARN) is a joint Department of Defense acquisition program that is networking radiological, biological, and chemical sensors and the output from hazard estimation models with joint and service command and control systems [1]. The Biological Warning and Incident Characterization (BWIC) system is a Department of Homeland Security system that establishes a common view of data related to a potential biological incident [2]. These data can consist of relevant maps, sensor data from multiple sources, weather data, atmospheric dispersion models, population information, disease progression, facility maps, and public health surveillance data. Both JWARN and BWIC are the necessary first steps in improving post-attack decision processes by assembling data relevant to a decision. More work is needed, however, in the areas of automated information fusion or response guidance.

CB defense decision makers fall into four broad categories: public health and medical authorities, operations managers and planners, law enforcement, and political leadership. All of these people must be armed with information on the degree of health concern and the scale of the event. Public health and medical authorities must decide on the appropriate medical response. Operations managers such as urban facility operators or military command positions must decide on how to change operations (e.g., close a facility or have troops don protective gear) in response to the event. Law enforcement and political leadership have critical roles in attribution of the attack to individuals or groups and in public relations.

CB attacks have been rare, and none in recent history have been large scale. Thus none of the people with the responsibility for deciding how to respond to a large CB attack have actually experienced one. Public health responders have had practice in responding to naturally occurring health events, but this experience may actually adversely influence decisions following a large-scale CB event. The reasons for this are twofold. Naturally occurring health events relevant to CB attack decision making (e.g., infectious disease outbreaks, poisonings) are generally small scale; resource limits are not tested, and large-scale health care responses are not warranted. Further, exposure or illness may occur over extended time scales, and the exposure generally is of low lethality. Rapid response is therefore not required, and excessive caution is rarely penalized. In contrast, too slow a response to a large-scale biological or chemical terrorism event could be catastrophic.

Current training of CB decision makers is inadequate to test rapid and dynamic decision-making skills. Emergency preparedness exercises for homeland defense are based on scenarios decided upon well in advance. Since responders have had significant time prior to the event to think of what the responses will be, these exercises tend to test organizational roles and responsibilities, communication, and mechanisms for implementing a response, rather than the decision process for deciding on the response. The military has conducted few battlefield exercises that practice operations in a chemical or biological environment.

Well-constructed tabletop exercises are more useful than scripted exercises for testing the response decision-making process. Rarely, however, are the tabletop participants given objective feedback on whether they chose the best of the available alternative courses of action. Moreover, the process of deciding whether an attack has occurred is rarely practiced. Tabletop exercises generally do not provide realistic and complete data simulating what would be available in the aftermath of a real attack, and so decision making in these practice sessions can be highly speculative.

Current training of chemical and biological defense decision makers is inadequate to test rapid and dynamic decision-making skills.

Information fusion aims to provide the decision maker with the best possible information on the likelihood, type, extent, and spread of chemical or biological contamination. It needs to provide these answers using information sources that can have high uncertainty, that have disparate origins and information types, that may conflict with one another, and that were collected at different locations or times. The information sources can be from biological and chemical sensors of different

Assessing a Sensor's Predictive Value

How much should we trust a chemical or biological detector's alarm?

When the probability of occur-rence of an event is low, as we hope is the case for future biological and chemical attacks, the confidence that a sensor alert corresponds to a real attack is also low. This is true even for highly discriminating sensors. For example, imagine a sensor network with the following characteristics:

Probability of detection	0.9
Probability of false alarm	10^{-4}
No. independent sensors	100
Sample analysis period	3 hr

If one assumes that an attack will occur every 100 years, then that translates into the presence of an attack once per 2.92×10^7 samples (8 samples per day x 365 days per year $\times$ 100 years $\times$ 100 sensors).

For each individual alert, decision makers care about how likely it is that the alert corresponds to a real attack. For example, the medical community is interested in the probability that a positive diagnostic test indicates a patient who has the corresponding condition; they refer to this as the positive predictive value of a test. Bayes's theorem expresses the probability that a sensor alarm means an attack has actually occurred as

$$P(\text{attack}\,|\,\text{alarm}) = \frac{P_d P_a}{P_d P_a + P_{fa}\left[1 - P_a\right]},$$

where P_a is the probability of an attack occurring, P_d is the probability that an attack that does occur will be detected, and P_{fa} is the probability of a false alarm. With the assumptions specified above, the probability that any one sensor alarm indicates a true attack is only 3×10^{-4}.

In other words, the confidence that any given alarm from this sensor network is true is very small. This does not mean that no action should be taken. The decision on whether to take action must consider the cost of the action and the frequency of alerts. Actions that impose a high cost (be it economic, health, social, or political) must be reserved for when we have a predictive value near 1.0. For example, large-scale antibiotic distribution in response to a suspected biological attack will certainly have a high economic cost and may pose a health risk to the population as a result of side effects and drug interactions. And if distribution is later shown to be unnecessary—that is, the result of a false alarm—then the loss of confidence in the government and the biological defense system could be dramatic.

Developing a single sensor that has a low enough false-positive rate to result in a high positive predictive value of an alert is a significant technical challenge. Assuming the probability of attack is small and the probability of detection is high, Bayes's theorem indicates that a high predictive value requires that the probability of a false alarm be much lower than the probability of an attack. In the situation specified above, the probability of false alarm would need to be 3×10^{-10} in order to have a 99% chance that the alarm represents a true attack.

One simplifying assumption implicit in the above discussion of Bayes's theorem is that the sensor network provides an alert that indicates whether an attack has occurred. In reality, however, today's sensors indicate only whether an agent is present in the environment. For example, biological sensors that detect the presence of DNA or proteins from various agents do not provide conclusive evidence as to whether the material is unnatural or endemic to the environment or whether it was alive and pathogenic when released. These pieces of information will inevitably arise from different sources. Even when using perfect sensors, therefore, information fusion is necessary.

mechanisms and specificities. In addition, other contextual information such as meteorological data, dispersion models, operations information, intelligence, additional surveillance data, and ancillary chemical or biological sensing information can be considered as sources.

Response guidance takes as input the situation assessment computed by the information fusion subsystem and helps decision makers decide what to do as a result of that information. There should be a strong interaction between the information fusion and response guidance modules; a good information fusion subsystem can help decision makers understand what additional information would improve the certainty or accuracy of the answers. A key requirement of response guidance for either pre- or post-attack decision making is the ability to estimate the impact of various actions on life, property, and operational missions.

In this article, we emphasize the use of information fusion and response guidance following a chemical or biological attack. But similar activity is also important before such an attack occurs. In the pre-attack period, information fusion is largely an intelligence function used to discover weapon development activity or plans, or enemy doctrine and tactics. The mission planning function in the pre-attack period is a good analog for the post-attack response guidance module.

Ways to Do Fusion

There are several types of information that decision makers may want at their disposal following a CB attack. In addition to information about whether an agent is present, they will want to understand the extent and time period of contamination, whether the release was intentional, and whether it poses a health hazard. Information fusion output that informs decision makers about aspects of the current situation will demand information from a number of different sources, collected and analyzed over markedly different time scales (e.g., minutes for chemical attacks to potentially days for biological attacks).

Some approaches might hold significantly less promise than others for CB defense. Basic probabilistic inference methods, for example, normally require a construction of probability distributions representing categories of interest, such as those of the background and agent release. These methods are well established and accessible to many engineering professionals. However, it is often difficult to acquire a statistically significant amount of experimental data to construct reliable probability distributions. Simple signal processing methods work best when the sensors can be modeled reliably and when the sensing phenomenology, background clutter, and sensor response characteristics are well understood.

In biological and to some extent chemical sensing, such precision cannot be expected. Reasons include insufficient level of knowledge of the underlying biological phenomena, complexity of aerosol/fluid dynamics, or limited verifiability of theories that attempt to describe these phenomena. Approaches that rely strongly on models are brittle; constructing robust models in the CB domain, such as those of transport and dispersion, is difficult because of the lack of phenomenological knowledge and uncertainty in model input parameters.

Approaches that encode information derived from human experts are appealing because of the human comprehensibility of the rules they implement. However, such rule-based methods require an ample body of extractable human expertise, which is deficient for many tasks of biological-chemical defense. Certain forms of machine learning and automatic reasoning can cope with the uncertainty, sparseness, and incompleteness of data, and do not rely directly on the phenomenological models. The methods we developed for subway aerosol anomaly detection, microarray pattern recognition, and the FLASH (fusion, learning, adaptive super-hybrid) architecture belong in this category. The drawback to these techniques is that they are complex and thus their development requires skill and expertise.

Fusion of Disparate and Colocated Sensors

One way to improve the predictive value of a sensing system is to combine information from multiple colocated sensors. If the sensors operate on completely different measurement principles, then it is often the case that they respond differently when measuring actual chemical or biological agents from the way they do when measuring potentially interfering materials.

As an example of the use of disparate sensors, consider the measurement of chemical nerve agent A with a commercial ion mobility spectrometer (IMS). Such spectrometers detect chemical vapors by measuring the time it takes for an ionized vapor to drift through an electric field. But an innocuous chemical B produces the same drift time

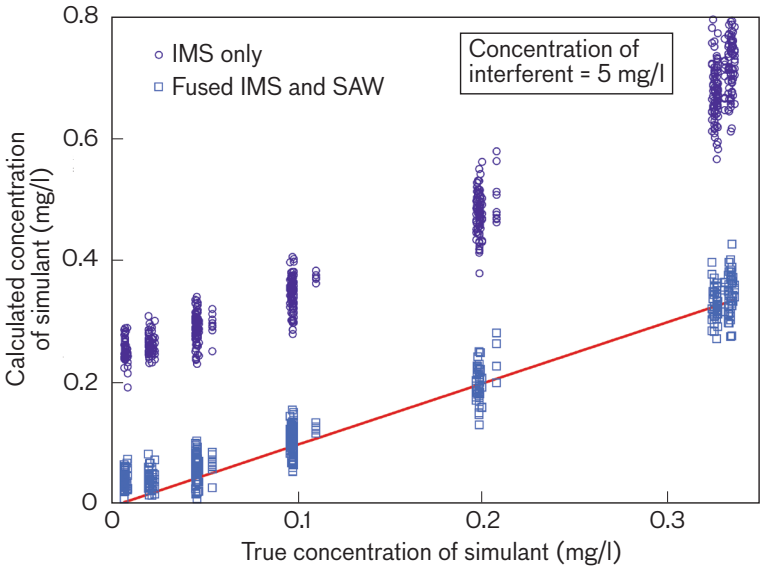


FIGURE 1. Fusing information from two sources improves accuracy. Shown here are plots of the calculated concentration of a nerve gas simulant versus the true concentration for two tests: one in which an ion mobility spectrometer (IMS) was used alone, and one in which data from the IMS data were fused with surface acoustic wave (SAW) data. The red line represents ideal performance.

as measured by an IMS as does agent A; it is therefore not possible to completely separate these two chemicals by using information only from this IMS. Chemicals can, however, be sensed by other mechanisms to improve the overall discrimination. For example, a commercial surface acoustic wave (SAW) chemical sensor produces output that is strongly correlated with the concentration of benign chemical B, but uncorrelated with harmful agent A. The SAW information on the concentration of B can be effectively subtracted from the IMS data to reveal the true amount of agent A (Figure 1).

Fusion of Many Information Sources

Multiple sensors do not necessarily need to use different detection principles or transduction mechanisms, or be housed in separate packages. In the case of biological detection, for example, multiple DNA fragments could be used to detect various DNA signatures. Such a DNA microarray, also called a gene chip, could analyze samples for hundreds or even thousands of different DNA fragments.

A DNA microarray project at Lincoln Laboratory has developed an approach that identifies pathogens by recognizing the gene expression patterns of cells that had been exposed to pathogens [3]. Cells obtained from commercially available cell lines are exposed *in vitro* to various pathogens. Messenger RNA (mRNA) is extracted from the cells after pathogen exposure. The mRNA is transformed into a stable complementary DNA, which is labeled and hybridized to DNA microarrays. This yields patterns of gene expression. The goal of the algorithms developed in this project is to enable automatic recognition of these patterns and thereby to identify the pathogen.

Such pattern recognition with DNA microarrays is challenging because of the large number of information sources; each microarray has thousands of sensing probes. The multiplicity of information sources yields a high-dimensional input space. In high-dimensionality spaces the process of deciding which patterns correspond to one category versus another—e.g., agent versus no agent—is more difficult than for low-dimensionality spaces. Much of the input data may be extraneous to the decision problem, and the algorithm must sort through which sources are important.

Most importantly, high-dimensionality spaces require a large number of training patterns—example combinations of input sources used by the learning algorithm. Unfortunately, measuring the response of a large number of microarrays to pathogens is time consuming and expensive, and a large number of training samples are not available. One alternative is to use a model of the expected DNA expression by the cells in response to a pathogen, and to use that model to predict the reaction of the DNA microarrays. However, biological science does not currently provide the knowledge and models needed for such predictions.

Under these circumstances—a large number of information sources, no available predictive model, and few training patterns—there are no existing techniques that can be fruitfully applied. For that reason, Lincoln Laboratory developed a new machine-learning-based approach, the structure of which is shown on the left side of Figure 2. The process called input space partitioning divides the input space into a number of subspaces, and different recognizers are trained for each of the subspaces. The subspaces are constructed in a manner that facilitates the

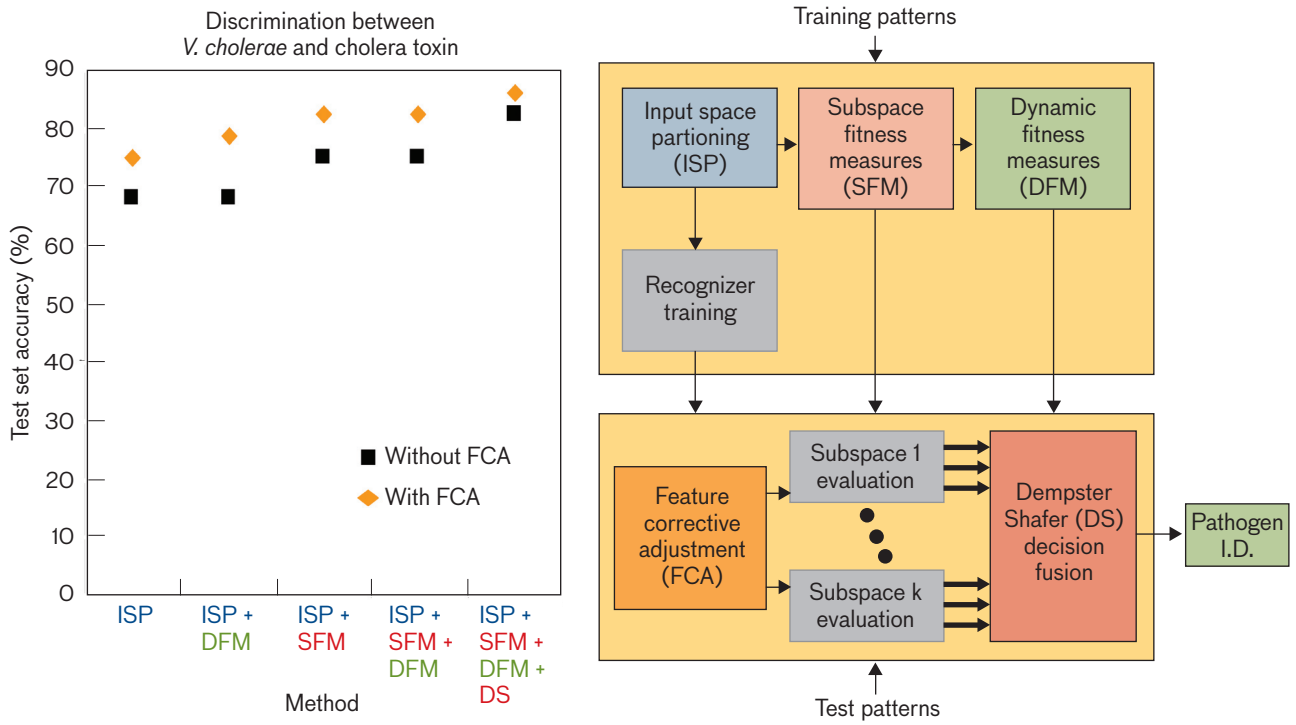


FIGURE 2. Addition of algorithm components improves discrimination between the pathogen that causes cholera and the cholera toxin (left). A block diagram of the microarray pattern analysis process is shown at right.

discrimination process, and the quality of the subspaces is characterized by fitness measures. Another process, called dynamic fitness measure, generates information about the discrimination quality of regions within the subspaces. Although input space partitioning alleviates the issue of input space dimensionality to an extent, the dimensionality of subspaces generated by it still remains significant. The subspace recognizers in our approach are based on support vector machines (SVM)—recognition methods that can cope with such challenges better than classical pattern recognition methods can.

When a test pattern is presented for recognition, the values of some of the pattern data can be modified to account for their uncertainty, using a process called feature corrective adjustment. The pattern is classified by multiple subspace classifiers, each operating in its respective subspace. The classification results and the subspace fitness measures are supplied to the final decision stage. That stage uses Dempster-Shafer theory, an approach that can be thought of as an alternative to and an extension of probability theory. Each subspace is considered a separate source of evidence and the subspace results are represented by the belief function values.

An example of the performance of the algorithm in identifying the DNA expression pattern from the organism that causes cholera from the cholera toxin is shown on the right side of Figure 2. No other algorithm technique was appropriate for this fusion task, so a comparable algorithm performance is not shown. This plot shows that additional algorithm modules improve performance. Those module additions which do not improve classification performance in this performance graph do demonstrate improvement for other pathogen classification problems. From this chart, we can see that improving discrimination performance by 10% to 20% requires making the algorithm significantly more complex. This requirement is common for other fusion problems as well. We don't know for certain what could be the best possible performance of a fusion algorithm, as such estimates cannot currently be made *a priori*.

Fusion of Qualitative and Quantitative Data

Information sources include systems that are not sensors as the term is usually defined. Nor is all information numeric. Indeed, one challenge facing comprehensive information fusion systems is the need to work with

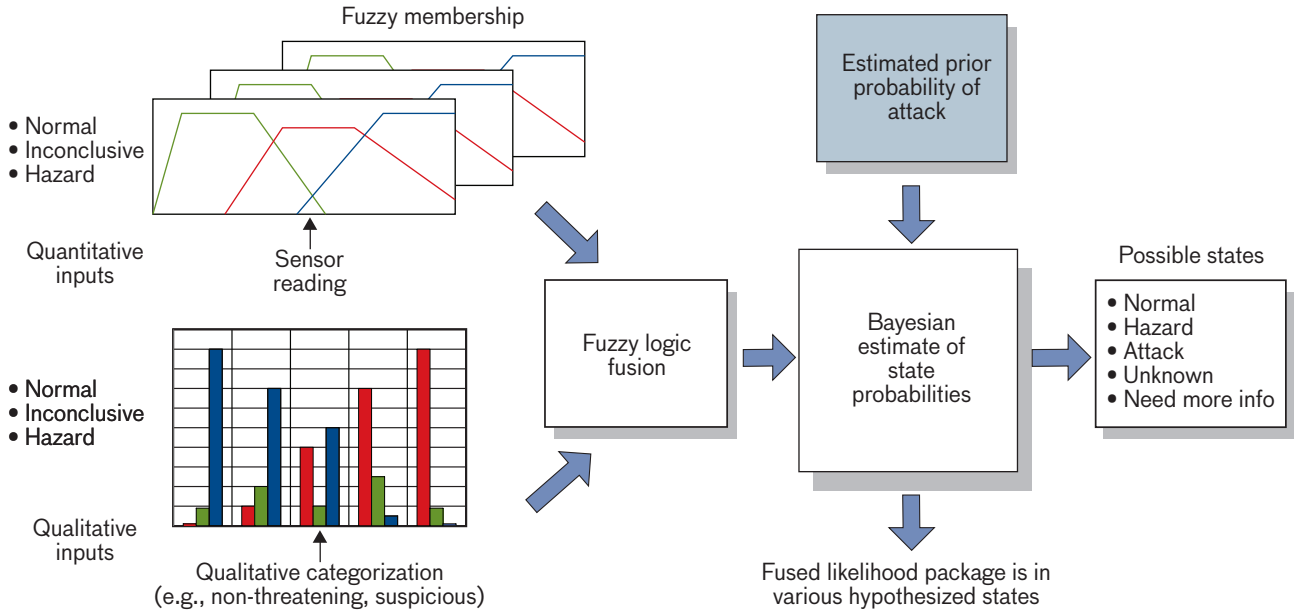


FIGURE 3. Information fusion is built into a decision support system for screening mail. This diagram shows the transformation of quantitative and qualitative information sources into fuzzy state memberships.

different fundamental forms of data representation (e.g., discrete, continuous, or qualitative).

Consider, for example, the information fusion algorithm that Lincoln Laboratory is developing for a system that may screen military mail for chemical, biological, radiological, and explosive (CBRE) threats. While the system relies in part on conventional CBRE sensors, it also takes input from human observations on the condition and attributes of packages and letters. The mail screening system incorporates the qualitative package observation information and the other quantitative data sources by using a fuzzy data fusion scheme.

Fuzzy data representations define the degree to which values from an information source belong to a set of states. For example, a numeric output from a chemical or biological sensor can be attributed as partially belonging to one or more hazardous, uncertain, or innocuous states. The functions that map the information values to these state memberships can be defined on the basis of the statistics from previous observations with that data source or they can be set heuristically. Fuzzy characterization of information sources allows the representation of inherent state ambiguity as well as probabilistic uncertainty. The package observation scores, though qualitative, can similarly be mapped to fuzzy state memberships. Thus the fusion process can use the qualitative data as well as the quantitative CBRE sensor data. The mail screening fusion process, as shown in Figure 3, uses a fuzzy version of Bayes's Rule to estimate the likelihood that a package or letter is hazardous.

Fusion of Geographically Dispersed Sensors

Biological and chemical releases in air spread material over an area large enough to expose multiple sensors to the agent. Information fusion algorithms can take advantage of the additional information from the multiple sensor measurements. For the additional sensor information to help determine whether the agent is present, however, the multisource information available to the system must contain aspects that the fusion system could exploit to distinguish a release from the background clutter. Unfortunately, we have only a weak understanding of the content, sources, and spatial-temporal patterns of biological and chemical sensor background clutter. Therefore, we usually need to determine the background patterns empirically. In contrast, observations of agent or simulant releases are sparse, and so agent releases are typically simulated with the use of various transport and dispersion models.

One such multisensor fusion application was demonstrated by a Lincoln Laboratory project that deployed a network of particle-density monitors in the Boston subway system [4]. Such instruments are not specific to biological materials, much less to a particular agent. However, such a network might be able to provide cues to anomalous

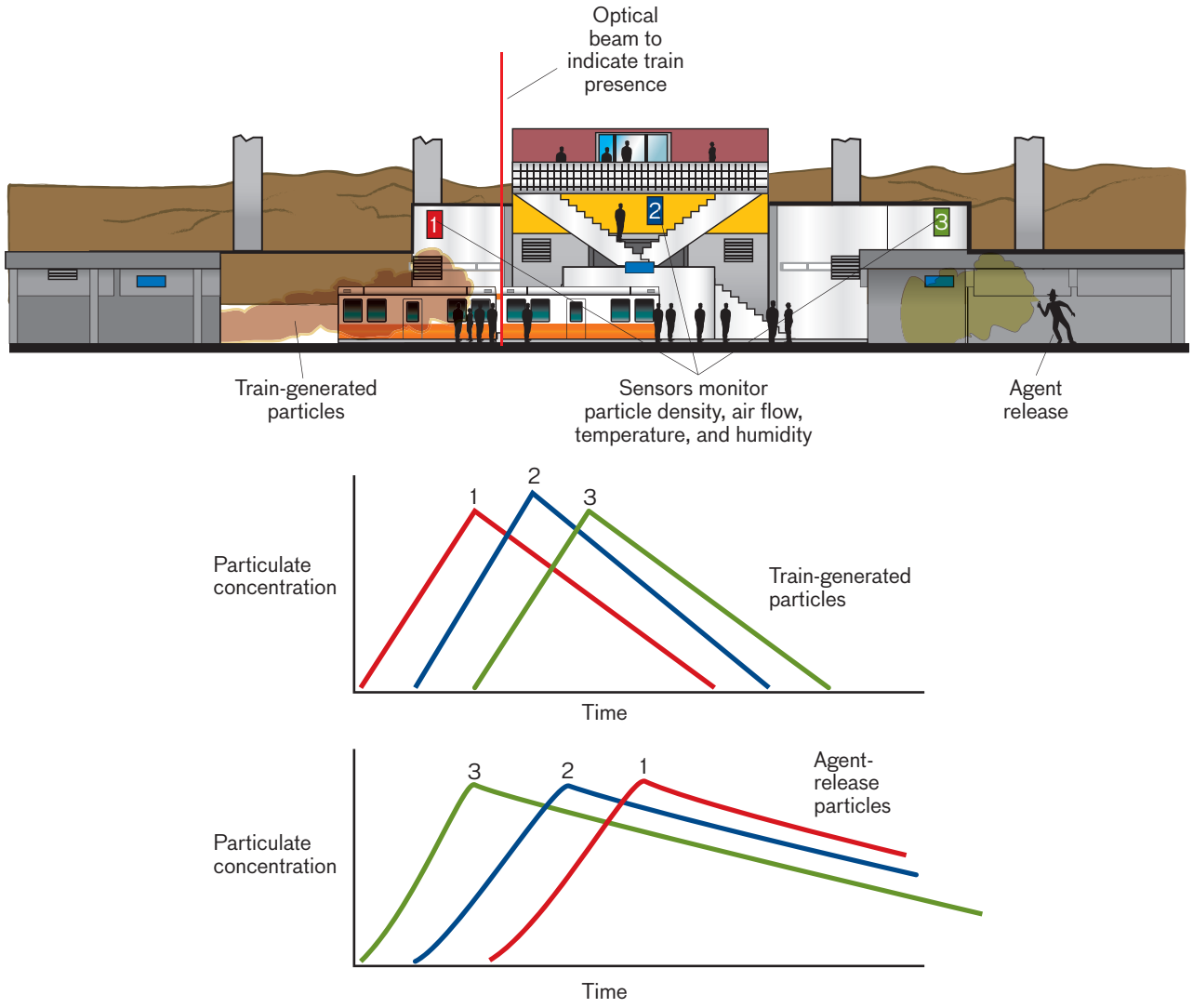


FIGURE 4. A representative illustration of the sensor deployment in the Boston subway testbed is shown in the top diagram. The patterns of activity from train-generated releases at each of the sensing nodes 1, 2, and 3 are expected to differ from the measurement patterns caused by terrorist biological releases.

aerosol activity within the subway system. This awareness could trigger additional surveillance, sample collection for more specific analysis, or other reactions such as changes to a ventilation system.

In the subway testbed, we operated three particle counters for one year at fixed points within each of two adjacent stations (Figure 4). We concurrently collected additional information on the passage of trains through the sensor network, the train speeds, and three-dimensional winds (from sonic anemometers). The subway particulate background is unusual in that it is characterized by large, rapid transients coinciding with train passage. In addition, there are daily variations in the baseline particulate density that correspond to rush hour periods and periods of relatively low train frequency.

The complex subway background presents algorithmic challenges that the use of additional contextual information sources mitigates. Train passage, particle size, and airflow information are input to the discrimination algorithm. The subway background transients are high enough that algorithms that rely on simple thresholding of particle-density measurements would result in an inability to detect all but very large biological releases. The ability to detect releases that result in concentrations well below that of the spikes that occur when a train rumbles by would be unlikely unless the algorithm were aware

of the times that trains pass. This information could be obtained in many ways; in this project, an optical beam break signaled the entrance and exit of trains from the stations. Additional information is gained by looking at sensor features other than total particulate count. The particle counters used in the Boston testbed have six particle-size channels, which present six distinct pieces of information to help distinguish intentional biological releases from the natural background. Finally, the pattern of presentation of releases to the multiple sensors can offer additional discrimination power. The airflow in subways is inconsistent and depends on external meteorological conditions. Therefore, the pattern with which a release cloud may expose a sensor network will be inconsistent as well.

The data provided by the multiple sensors were used to decide automatically whether short time periods of subway multisensor data were anomalous, and thus potentially indicative of a biological release. Learning machines, in this case artificial neural networks, were used because the release and background patterns were likely to vary from station to station and between subway systems. Artificial neural networks are pattern recognition methods that learn from examples, automatically gaining the capability to classify patterns presented to them. Two neural networks were used to perform the classification, one for periods with train activity and one for periods without train activity.

Classification results are often represented by the receiver operating curve (ROC), which illustrates the tradeoff between the probability that a system will detect a real threat and the probability that it will issue a false alarm. Figure 5 compares the ROC from a simple algorithm with the neural network. The simpler algorithm is based on whether any of the single-sensor outputs indicated an anomaly, where each sensor looks for an anomaly, determined by the integrated particle count over each analysis period. As the graph makes evident, this algorithm's results were barely better than random chance for small releases. The neural network performs significantly better.

Despite its dramatic advantage over the simple algorithm, the neural network algorithm as shown here is still not good enough for operational use in the case of small releases (red curve): its 90% probability of detection comes with a roughly 50% false-alarm rate. For larger release sizes the performance of the neural network algorithm is much better (green curve). To achieve still better performance would require either a more sophisticated algorithm, more discriminating sensors, less frequent updates of the algorithm (to allow more information to accumulate), or the inclusion of additional sensors and other information sources.

Next-Generation Information Fusion: FLASH

An algorithm developer generally aims to use as simple a method as is necessary to achieve the desired performance. Unfortunately, there is no agreed-upon theory that allows the designer to estimate the difficulty of a fusion problem ahead of time, and certainly not to estimate the achievable performance from application of a particular algorithmic technique. Designing and implementing information fusion algorithms is therefore a trial and error process. Often the developer will start by applying a simple technique; if performance turns out to be unacceptable, then that calls for either a wholesale change in approach or else application of quick-fix patches to preserve cost and schedule goals. The simple methods and the associated performance patches can produce a cobbled-together algorithm that stops working once the system has pushed even slightly beyond the original test conditions. The next-generation fusion approach described in this section avoids such compromises.

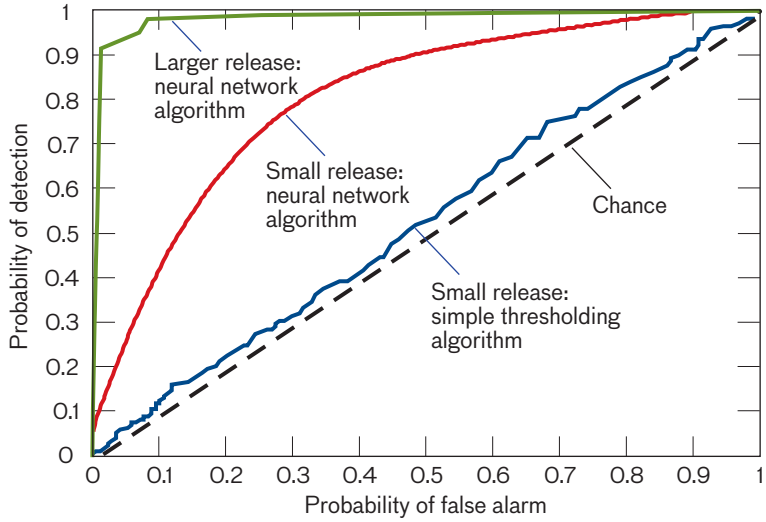


FIGURE 5. A neural network outperforms a simple thresholding algorithm in detecting a simulated agent release in the Boston subway testbed.

The FLASH architecture developed at Lincoln Laboratory represents a hybrid of several different methods [5]. This hybrid architecture, in which each individual technique compensates for shortcomings in other methods, should not only improve performance for difficult fusion problems but also provide a more general-purpose fusion engine. FLASH is designed as a hybrid of multiple heterogeneous machine-learning and approximate-reasoning methods. The first implementation, FLASH-1, was tested on the fusion of biological trigger devices and other contextual information for bio-attack detection in a building.

Figure 6 shows basically how FLASH-1 works. Information inputs are subjected to initial extraction of features—for example, the mean and variance of a time-series input signal for a particular time window. Extraction and selection of appropriate information features are critical elements to the success of any decision process. To make sure that it includes all high-information-content features, FLASH-1 extracts a large range of diverse features. Use of all features would make automated classification more difficult, however, because some features contain extraneous information. The most useful features are chosen in FLASH by a feature selection process. Techniques rooted in information theory are exploited to rank features according to the level of their usefulness for the discrimination process.

FLASH's instance recognition module contains a set of machine-learning algorithms that classify the input pattern. Instance recognition involves short-term windows of input data. This stage employs multiple support vector machines (SVM)—machine-learning constructs that tend to work better than classical methods, particularly when the data are sparse and imprecise. SVMs also generalize well; that is, their performance on data that differ drastically from the training data tends to be better than that offered by many other methods.

During training, FLASH attempts to establish whether the background training data represent different types of background. This task, performed by the background clustering module, regulates the number of background classes and influences the dynamic number of classifiers in the instance recognition module. The instance fusion module fuses the outcomes of the instance recognition classifiers and uses techniques based on the Dempster-Shafer theory of evidence to generate estimates of the degree to which the input pattern corroborates a

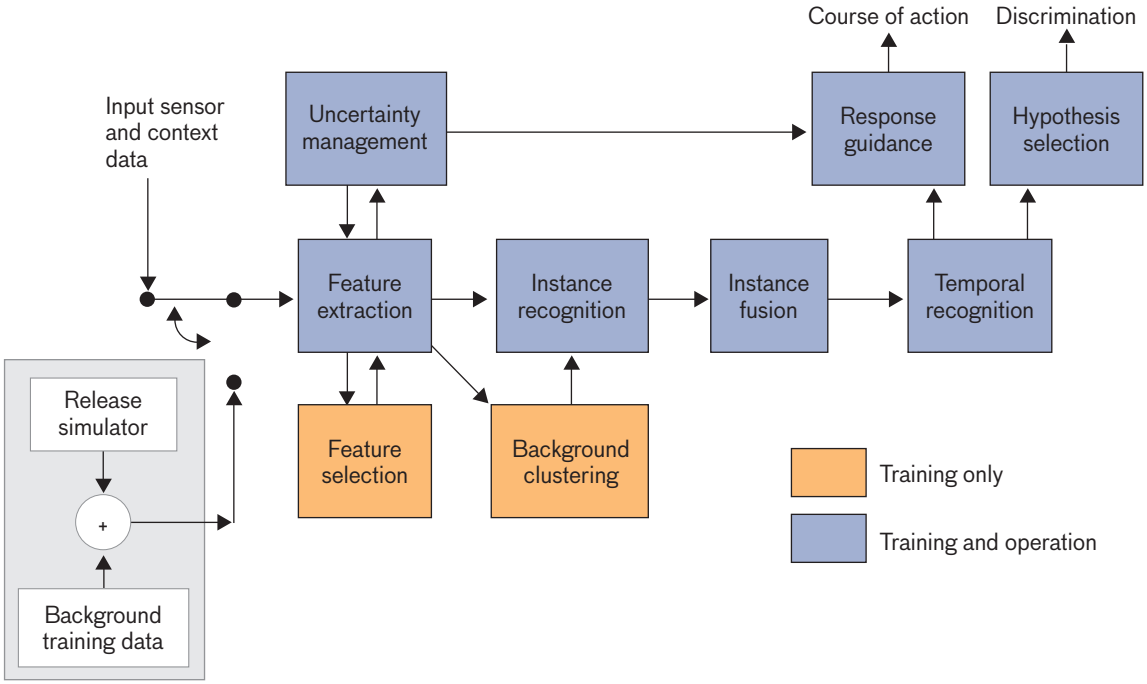
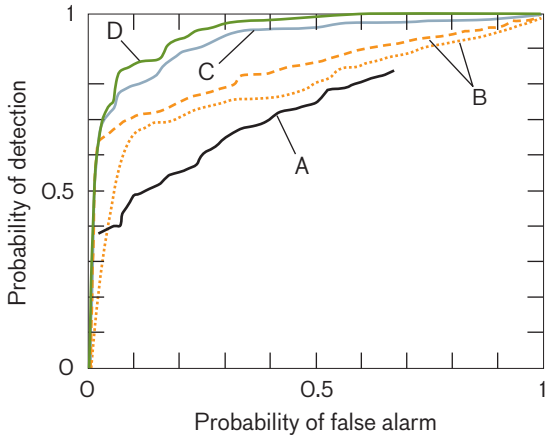


FIGURE 6. FLASH—for Fusion, Learning, Adaptive Super-Hybrid—blends multiple methods of machine learning. The implementation shown, FLASH-1, fuses data from biological trigger devices with contextual information to determine whether a building has been subjected to a biological attack.

hypothesis (e.g., a bio-attack). The temporal recognition module considers the decisions made at the instance level. The methods used at that stage include Hidden Markov Model classifiers. Hidden Markov Models are capable of learning a pattern's sequential aspects, a property that makes them particularly suitable for classification of time series. The hypothesis selection module chooses the hypothesis corresponding to the most plausible evidence from the temporal recognition module. The reasoning module has two portions. The first uses fuzzy inference to adjust the hypothesis evidence that is based on the qualitative information representing the current perceived threat level. The second uses a Bayesian network to provide response guidance that is based on the hypothesis evidence outputs from the temporal recognition module, other contextual information sources such as the threat risk level, and action utility estimates. As Figure 7 illustrates, each processing module adds discrimination value.



A. Simple multisensor fusion algorithm
B. Feature extraction, feature selection, uncertainty management, background clustering, and instance recognition (one curve for each of two instance recognition classifiers)
C. As in (B), plus instance fusion and temporal recognition
D. As in (C), plus fuzzy context modification method within hypothesis selection module (threat data simulated)

FIGURE 7. Each additional algorithmic stage adds to the FLASH-1 performance.

Response Guidance

While information fusion can bring a degree of situational awareness to CB defense, decision makers need more than that—they must understand how to interpret the information and decide what courses of action they should take. Some actions have small costs if taken on the basis of faulty information and can generally be scripted—possibly without any human oversight. Examples of such responses include adjustments to a building's ventilation system and initiation of additional surveillance or measurements. Most other response scripts, however, are usually inflexible to the contextual situation within which alerts will occur. Worse, these scripts are often based on an inadequate understanding of the risks of taking various actions and of the comparative value and reliability of different information sources.

Some response choices, moreover, cannot easily be pre-scripted, because the operational situations within which the decision must be made are too dynamic. Battlefield commanders' decisions fall into this category. For example, imagine a mobile force that is aware of chemical surface contamination at an upcoming bridge. Possible response options include going around, suiting up the troops in protective gear, and going forward without protective gear. The appropriate choice will depend on the degree of hazard of the chemical; that is one of the jobs of the information fusion process. Even with accurate information on the hazard, however, the appropriate action cannot be defined prior to the battle. The best choice depends on a balance between the delay from going around or donning protective gear; on the potential casualties from the chemical or extended use of the protective gear; and on the situation the other parts of the fighting force are in and what the enemy is currently doing.

CB defense thus needs a computational task to provide response guidance. When used in a planning phase, a response guidance algorithm can help design response scripts that appropriately balance the risks of action and inaction. Such scripts will be most useful when the number of possible operational contexts is bounded. Many existing response plans are too static and don't fully consider all of the possible situations that could occur. For example, some biological detection systems that are intended to provide information to support deployment of treatment generally have response scripts that are invariant to the nuances in the detection information and to the operational context. Instead, the appropriate responses, and the times that those responses are initiated, should be chosen by the type of agent detected (e.g., contagious versus non-contagious, treatable versus untreatable, degree of virulence), the spatial area over which the agent was detected, the amount of material detected, the weather

conditions, and the nature of the population in the area of the detection (indoor, outdoor, public gathering, VIP event), among other considerations.

Response guidance analysis can be used to help create the plans so that the response actions will be adaptive to particular threats, the incoming information content, logistical or operational constraints, and particular vulnerabilities. When used in a tactical response phase, the response guidance algorithm could adjust recommended actions appropriate for the given situation, exploiting fused information from sensors, intelligence, operations workload, and tactical constraints.

Response guidance includes response utility estimation and response selection. The utility estimation assesses the cost and benefit of various responses that have presumably been chosen ahead of time; making this estimate will inevitably involve some method for combining costs and benefits that have inherently different units (e.g., dollars, lives, mission effectiveness, social and psychological impact). It must evaluate the costs and benefits in a probabilistic sense, given uncertain situation assessments. The impact estimates can by themselves be difficult to compute. For example, the Department of Defense Joint Operational Effects Federation program [6] is developing simulation capabilities to estimate the impact of CB attacks on military missions for planning situations, and larger war-gaming simulations are being used to evaluate mission performance in situations during which the enemy responds. The response selection process determines the best combination of cost and benefit to meet the system objectives. The logic of this process will vary, depending on which system performance criterion is being optimized. Responses can be selected to maximize the expected benefit, to minimize the expected cost, to minimize the likelihood of a worst-case outcome, or to provide maximal utility, given resource constraints.

The CBRE screening of military mail, for example, represents a fairly simple response guidance problem. There are few possible courses of action (send package, examine further, destroy, forward with warning, quarantine, or communicate with addressee) and few information sources (CBRE sensors, human package-threat estimates, and intelligence). However, a mail screening system that used fixed rules for making the response decisions may not be robust. Confidence in the input information, such as sensor quality, may change with time. Tolerance for risk may dynamically change, particularly in response to new intelligence data. The costs of various actions may change with time as, for example, the costs and availability of human resources vary.

Figure 8 shows the mail screening algorithm. The response utility table, initially set to predetermined fixed values, defines the cost of various response actions. Combining the response utilities with state information about the packages allows the algorithm to compute the expected utility of each action. The best response can be selected on the basis of a number of possible criteria, such as maximum expected utility, minimum likelihood of worst case, and maximum probability of detecting a threat. FLASH-1 also contains a response guidance algorithm. It uses a Bayesian belief network, an inference process that results in the ranking of the possible courses of action, and the selection of the most desirable of those alternatives.

Extrapolation of the mail screening and FLASH response guidance algorithms to dramatically more difficult problems, such as those which arise on the battlefield, has yet to be done. Also in need of significant development are simulation-driven estimates of the utility of various courses of action. It is nevertheless clear that response guidance is aided considerably by information fusion methods.

Making Fusion Richer

Although progress lags behind more mature fields, decision support for chemical and biological defense applications is beginning to gain the attention that it deserves. Investment must be made in four important areas: information sources, algorithms and models, test and evaluation methods, and training.

There is the tendency for CB defense researchers as well as government managers to think of information fusion for CB defense as being based entirely on CB-sensor information, meteorological data, and agent dispersion models. But in fact, it is likely that there are valuable information sources that have not yet been tapped or even strongly considered for inclusion in CB fusion systems. In civilian applications, eyewitness or surveillance camera observation of unusual behavior or sensitive-area intrusion detection should be captured quickly and fed into information fusion systems. Battlefield situations, too, offer a potentially rich set of information sources

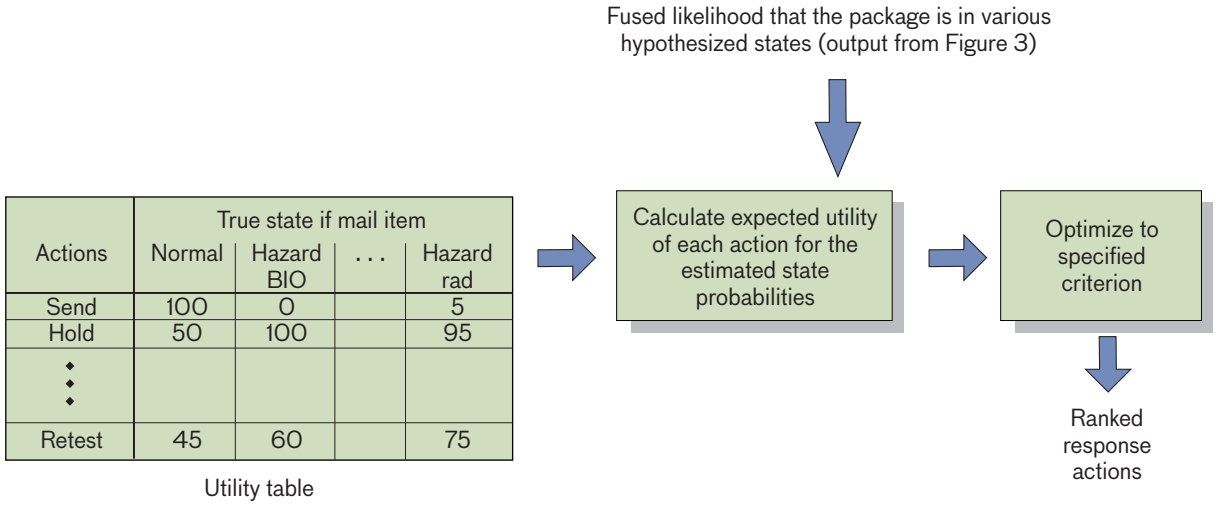


FIGURE 8. A system providing decision support for mail screening illustrates an application of information fusion and response guidance.

that could be exploited. For example, knowledge of enemy force locations, in combination with wind patterns, can permit decision makers to discount reports of remote releases. Observations of low-flying enemy aircraft of various types could change the way some sensor alerts are interpreted. Artillery observations, particular by those with radar, acoustic, seismic, or optical signatures representative of chemical or biological munitions, could substantially change the manner in which the CB sensor data are exploited and interpreted. Surveillance information related to tanker trucks or pesticide sprayers or of toxic industrial chemical sites upwind of friendly forces should be included as well. Speech recognition systems that automatically transcribe and automatically feed radio traffic related to CB events could be used to make the information rapidly available to information fusion systems. Robust information fusion demands incorporation of this rich array of disparate information sources.

At the same time, algorithms must continue to be developed that capitalize on the unique aspects of the information sources relevant to CB defense. This work should focus both on near-term fixes and on general-purpose information fusion architectures that would prove valuable in the future. Information fusion and response guidance algorithms—the latter of which has been largely ignored until now—should be considered equal in importance to information integration and presentation. There is a need for adequate long-term, multiple information-source data sets for testing and developing fusion algorithms. Furthermore, the performance characteristics of the information sources should be better characterized with performance indicators such as the receiver operating curves. Such characterization will facilitate the development of better fusion algorithms and also improve the ability to test the algorithms realistically. Finally, testing of response guidance algorithms should be done with robust response utility models. Human-in-the-loop testing of response tactics under a variety of situations will be necessary to ensure that response utility estimates properly consider all of the essential costs and benefits.

Although this article has emphasized automated post-attack decision support, we recognize that well-trained decision makers will make more effective decisions, with or without automation. A dramatic shift in our training methods for such decision makers is needed so that they can be effective in situations with sparse and uncertain information and can cope with rare and catastrophic situations. Simulation-based training, as an example, may offer some promise.

Ultimately, the various technologies for detection, protection, decontamination, and medical treatment will provide their full benefits only if decision makers can properly understand how to interpret the information and can decide consistently how to best apply the technologies in response. When implemented in operational systems, techniques such as those we have described could significantly improve the effectiveness of chemical and biological defense systems.

Acknowledgements

The authors thank Michael Walter of the Joint Program Executive Office for Chemical and Biological Defense for his prescience regarding the need for a mail screening decision support system and his guidance during the effort. Additional Lincoln Laboratory participants on that effort were Ronald Hoffeld, Gerald Larocque, Laura Brattain, Taylor Locke, Charles Yee, Sean Winkler, and Bernadette Johnson.

The National Science Foundation (NSF) funded the development of FLASH. Lincoln Laboratory participants on FLASH include Yan Glina, Laura Brattain, David Stein, Peter Skomoroch, Kevin Transue, and Raymond Uttaro. The Department of Defense Research and Engineering funded the subway, DNA microarray, and chemical sensor work. We appreciate the excellent work of other Lincoln Laboratory researchers on these projects: Yan Glina, Michael Matthews, Skip Copeland, and Jonathan Su for the subway effort; Sunil Jeswani, Yan Glina, Nicholas Judson, Rachel Herzig-Marx, Bernadette Johnson, and Kevin Transue on microarrays; and Michael Switkes and Richard Czerwinski on chemical sensing.

This material is based upon work supported by the National Science Foundation under Grant No. 0329901. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the NSF. ■

REFERENCES

1. V. Wing and S. Spadaro, "JWARN," *Military Medical Technology,* vol. 8. no. 6, 2004, www.military-medical-technology.com/article.cfm?DocID=657.
2. Department of Homeland Security Fact Sheet, "Biological Warning and Incident Characterization," no. 2006-4491P, Sandia National Laboratories, Albuquerque, July 2006.
3. J. Braun, Y. Glina, N. Judson, and K.D. Transue, "Biological Agent Detection and Identification Using Pattern Recognition," *Proc. SPIE,* vol. 5795, 2005, pp. 113–124.
4. J. Braun, Y. Glina, J.K. Su, and T.J. Dasey, "Computational Intelligence in Biological Sensing," *Proc. SPIE*, vol. 5416, 2004, pp. 111–122.
5. J. Braun and Y. Glina, "Hybrid Methods for Multisource Information Fusion and Decision Support," *Proc. SPIE*, vol. 6242, 2006, pp. 624209-1–1624209-12.
6. S. D. Kwak and E.L. Berger, "JOEF (Joint Operational Effects Federation) Architecture," MITRE Technical Paper, The Mitre Corp., Bedford, Mass., www.mitre.org/work/tech_papers/tech_papers_03/kwak_joef/, Sept. 2003.

ABOUT THE AUTHORS

Timothy Dasey is associate leader of Lincoln Laboratory's Biodefense Systems group; he manages algorithm development, system analysis, modeling, and simulation and software development for military and homeland security applications. From 1991 through 2001 he developed advanced algorithms and software for the Laboratory's Weather Sensing group; air traffic facilities nationwide use these products to enhance safety and capacity. He has a bachelor's degree in electrical and computer engineering from Clarkson University and a doctorate in biomedical engineering from Rutgers University.

Jerome J. Braun is a staff member in the Biodefense Systems group; he focuses on information fusion and intelligent decision support systems. Since 2003, he has led a National Science Foundation–sponsored program in multisensor information fusion for biodefense. He has a bachelor's degree in physics and a master's degree in computer science from the Technion – Israel Institute of Technology, and a doctorate in computer science from the University of Massachusetts Lowell.