AWARD/CONTRACT

1 THIS CONTRACT IS A RATED ORDER UNDER DPAS (15 CFR 350)	Rating	Page 1 of Pages 25

2 CONTRACT (Proc Inst Ident) NO	3 EFFECTIVE DATE	4 REQUISITION/PURCHASE REQUEST PROJECT NO
D10PC20015	09/12/2010	SZ100653

5 ISSUED BY CODE 00063	6 ADMINISTERED BY (If other than Item 5) CODE
DOI - NBC, Ft Huachuca AZ NBC / Acquisition Services Directorate Sierra Vista Branch P O Box 12924 Fort Huachuca AZ 85670-2924	

7 NAME AND ADDRESS OF CONTRACTOR (No, street, city, county, State and ZIP Code)	8 DELIVERY FOB Origin: ✓ Other (See below)
BBN TECHNOLOGIES CORP 10 MOULTON ST CAMBRIDGE MA 02138-1119	9 DISCOUNT FOR PROMPT PAYMENT 10 days % 20 days % 30 days % 0 days %
CODE FACILITY CODE	10 SUBMIT INVOICES (4 Copies unless otherwise specified) TO THE ADDRESS SHOWN IN ITEM 12

11 SHIP TO/MARK FOR CODE DARPA	12 PAYMENT WILL BE MADE BY CODE D2730
Joseph Olive DARPA 3701 N. Fairfax Drive Arlington VA 22203-1714	Dept of the Interior - National Business Center National Business Center - Denver Fiscal Support Section - D2730-FH 7301 W Mansfield Denver, CO 80235-2230

13 AUTHORITY FOR USING OTHER THAN FULL AND OPEN COMPETITION	14 ACCOUNTING AND APPROPRIATION DATA
☐ 10 U S C. 2304(c) ( ) ☐ 41 U S C. 253(c) ( )	2010 - - WH - 6041 - - 255D - - WHDRP - - - 299400 - - 1A - - - -

15A ITEM NO.	15B SUPPLIES/SERVICES	15C QUANTITY	15D UNIT	15E UNIT PRICE	15F AMOUNT
	SEE LINE ITEM DETAIL				

15G. TOTAL AMOUNT OF CONTRACT 1,000,000 00

16 TABLE OF CONTENTS

( )	SEC	DESCRIPTION	PAGE(S)	( )	SEC	DESCRIPTION	PAGE(S)
		PART I - THE SCHEDULE				PART II - CONTRACT CLAUSES	
X	A	SOLICITATION CONTRACT FORM	1	X	I	CONTRACT CLAUSES	16-24
X	B	SUPPLIES OR SERVICES AND PRICES/COSTS	2			PART III - LIST OF DOCUMENTS, EXHIBITS AND OTHER ATTACH	
X	C	DESCRIPTION/SPECS/WORK STATEMENT	4	X	J	LIST OF ATTACHMENTS	25
X	D	PACKAGING AND MARKING	5			PART IV - REPRESENTATIONS AND INSTRUCTIONS	
X	E	INSPECTION AND ACCEPTANCE	6		K	REPRESENTATIONS, CERTIFICATIONS AND OTHER STATEMENTS OF OFFERORS	
X	F	DELIVERIES AND PERFORMANCE	7				
X	G	CONTRACT ADMINISTRATION DATA	8-10		L	INSTRS , CONDS , AND NOTICES TO OFFERORS	
X	H	SPECIAL CONTRACT REQUIREMENTS	11-15		M	EVALUATION FACTORS FOR AWARD	

CONTRACTING OFFICER WILL COMPLETE ITEM 17 OR 18 AS APPLICABLE

17 ✓ CONTRACTOR'S NEGOTIATED AGREEMENT *(Contractor is required to sign this document and return* 1 *copies to issuing office)* Contractor agrees to furnish and deliver all items or perform all the services set forth or otherwise identified above and on any continuation sheets for the consideration stated herein. The rights and obligations of the parties to this contract shall be subject to and governed by the following documents: (a) this award/contract, (b) the solicitation, if any, and (c) such provisions, representations, certifications, and specifications, as attached or incorporated by reference herein.

(Attachments are listed herein.)

18. AWARD *(Contractor is not required to sign this document.)* Your offer on Solicitation Number _____ including the additions or changes made by you which additions or changes are set forth in full above, is hereby accepted as to the items listed above and on any continuation sheets. This award consummates the contract which consists of the following documents: (a) the Government's solicitation and your offer, and (b) this award/contract. No further contractual document is necessary.

19A NAME AND TITLE OF SIGNER *(Type or print)*	20A NAME OF CONTRACTING OFFICER
William P. Quantrille Vice President, Contract Relations	Lisa A. Mattocks
19B NAME OF CONTRACTOR By William P. Quantrille *(Signature of person authorized to sign)*	19C DATE SIGNED 9/14/10
20B UNITED STATES OF AMERICA By Lisa A. Mattocks *(Signature of Contracting Officer)*	20C. DATE SIGNED 9/14/10

NSN 7540-01-152-8069
PREVIOUS EDITION UNUSABLE

STANDARD FORM 26 (REV 4-85)
Prescribed by GSA
FAR (48 CFR) 53.214(a)

Total Funding: $1,000,000.00

FYs	Fund	Budget Org	Sub	Object Class	Sub	Program	Cost Org	Sub	Proj/Job No.	Sub	Reporting Category
2010	WH	6941		255D		WHDRP			Z99400		1A

Division **Closed FYs** **Cancelled Fund**

Line Item Number	Description	CLIN Ref	Delivery Date (Start Date to End Date)	Quantity	Unit of Issue	Unit Price	Total Cost (Includes Discounts)
0001	Research and Development Services	0001		1.00	LO	$1,000,000.000	$ 1,000,000.00

(09/01/2010 to 02/28/2014)

The contractor shall furnish the necessary personnel, materials, facilities and non-personal services to perform the research and development effort in accordance with the contractor's technical and cost proposal entitled, "RATS Patrol" dated 28 March 2010 submitted in response to DARPA BAA 10-34 and incorporated to this contract by reference.

Cost $12,212,497.00
Fee $870,048.00
CPFF $13,082,545.00

Ref Req No: SZ100653

Funding Information:

2010 - - WH - 6941 - - 255D - - WHDRP - - - Z99400 - - 1A - - - -
$1,000,000.00

Total Cost: $1,000,000.00

TABLE OF CONTENTS

SECTION C -- Descriptions and Specifications 4
C.1 Descriptions And Specifications 4
C.2 Deliverables 4
SECTION D -- Packaging and Marking 5
D.1 MARKING 5
D.2 PACKAGING 5
SECTION E -- Inspection and Acceptance 6
E.1 Review and Oversight 6
E.2 Inspection of Research and Development (Short Form) 6
SECTION F -- Deliveries or Performance 7
F.1 Clauses By Reference 7
F.2 Delivery of Data 7
F.3 Period of Performance 7
SECTION G -- Contract Administration Data 8
G.1 Contracting Officer's Technical Representative 8
G.2 DARPA Program Manager 8
G.3 Contract Administration Data 8
G.4 Government Contract Administrator 9
G.5 Defense Contract Audit Agency (DCAA) Office 9
G.6 Payment Office 9
G.7 Contract Payments 10
SECTION H -- Special Contract Requirements 11
H.1 Notification Under a Cost Reimbursement Contract 11
H.2 Subcontracting Plan 11
H.3 Award Made on Behalf of A DOD Agency 11
H.4 Key Personnel Requirements 11
H.5 NOTICE OF INCORPORATION OF SECTION K 12
H.6 Equipment Clause 12
H.7 No-Cost Settlement 12
H.8 Permits, Taxes, Licenses, Ordinances and Regulations 12
H.9 Publication 12
H.10 Travel and Per Diem 13
H.11 Assertion of Data Rights 13
H.12 ACQUISITION OF COMPUTER EQUIPMENT AND RELATED ITEMS 13
H.13 SF294 AND SF295 SUBCONTRACTING REPORTING REQUIREMENTS 13
H.14 Reporting Requirements 13
H.15 Incremental Funding 15
H.16 Equipment - AGILE 15
SECTION I -- Contract Clauses 16
I.1 Clauses By Reference 16
I.2 Notification of Ownership Changes 17
I.3 Payment For Overtime Premiums 18
I.4 Authorized Deviations In Clauses 18
I.5 DEFENSE FEDERAL ACQUISITION REGULATION SUPPLEMENT 19
I.6 1452.203-70 Restriction on Endorsements JULY 1996 20
I.7 1452.204-70 Release of Claims JULY 1996 20
I.8 1452.215-70 Examination of Records by DOI APRIL 1984 20
I.9 1452.228-7 Insurance - Liability to Third Person APRIL 1984 20
I.10 1452.237-70 Information Collection - DOI JULY 1996 21
I.11 Subcontracts 21
I.12 Notification Of Changes 23
SECTION J -- List of Documents, Exhibits and Other Attachments 25
J.1 ATTACHMENTS 25

SECTION C -- DESCRIPTIONS AND SPECIFICATIONS

C.1 DESCRIPTIONS AND SPECIFICATIONS

SECTION C -- DESCRIPTIONS AND SPECIFICATIONS

The contractor shall furnish the necessary personnel, materials, facilities and non-personal services to perform the research and development effort in accordance with the contractor's technical and cost proposal entitled, "RATS Patrol" dated 28 March 2010 submitted in response to DARPA BAA 10-34 and incorporated to this contract by reference. Proposed options will not be funded.

C.2 DELIVERABLES

The deliverables for this effort are as follows:

a. Quarterly Progress Reports (QPR)
b. Papers published in international conferences and/or journals during the course of the program
c. All software and documentation developed under the project
d. Any data collected under the project
e. Integrated systems for evaluation (6 weeks before the end of each phase)
f. Trainable, military user-friendly system during Phases 2 and 3
g. Final report

(END OF SECTION C)

SECTION D -- PACKAGING AND MARKING

D.1 MARKING

Shipping documents, containers, correspondence and packages shall be marked with the following information:

Contractor:	Raytheon BBN Technologies Corp.
Contract Number:	D10PC20015
Proposal Title:	"RATS Patrol"

D.2 PACKAGING

All deliverables called for shall be packed and shipped in accordance with the Contractor's standard commercial practices in a manner that shall afford adequate protection against physical and environmental deterioration and damage during shipment.

(END OF SECTION D)

SECTION E -- INSPECTION AND ACCEPTANCE

E.1 REVIEW AND OVERSIGHT

Review and oversight of research and development services will be made by the DARPA Program Manager (PM) identified in Section G. DARPA PM will inform Government Contracting Officer's Representative identified in Section G, if research and development services do not conform to contract requirements.

(END OF SECTION E)

E.2 52.246-09 INSPECTION OF RESEARCH AND DEVELOPMENT (SHORT FORM) APRIL 1984

The Government has the right to inspect and evaluate the work performed or being performed under the contract, and the premises where the work is being performed, at all reasonable times and in a manner that will not unduly delay the work. If the Government performs inspection or evaluation on the premises of the Contractor or a subcontractor, the Contractor shall furnish and shall require subcontractors to furnish all reasonable facilities and assistance for the safe and convenient performance of these duties.

SECTION F -- DELIVERIES OR PERFORMANCE

F.1 52.252-02 CLAUSES INCORPORATED BY REFERENCE

This contract incorporates one or more clauses by reference, with the same force and effect as if they were given in full text. Upon request, the Contracting Officer will make their full text available. Also, the full text of a clause may be accessed electronically at this/these address(es):

http://farsite.hill.af.mil/

Clause	Title	Date
52.242-15 Alt I	Stop-Work Order (Aug 1989) Alternate I	April 1984
52.247-34	F.O.B. Destination	November 1991

F.2 DELIVERY OF DATA

The data to be delivered with the reports shall be delivered in accordance with the instructions provided as identified in Section J.

F.3 PERIOD OF PERFORMANCE

The contract period of performance will be a 42-month period beginning 01 September 2010 through 28 February 2014.

(END OF SECTION F)

SECTION G -- CONTRACT ADMINISTRATION DATA

G.1 CONTRACTING OFFICER'S TECHNICAL REPRESENTATIVE

Contracting Officer's Technical Representative (COTR) responsibilities will be performed by:

Department of Interior/National Business Center
Acquisition Services Directorate
Sierra Vista Division
Post Office Box 12924
Fort Huachuca, Arizona 85670-2924

POC: Lawrence H. Carter
Phone: (520) 533-1213
Fax: (520) 538-3761
E-Mail: Lawrence_H_Carter@nbc.gov

Building Location and Zip Code for Overnight Deliveries:
Augur and Adair Streets, Bldg #22208 (2nd Floor)
Zip Code: 85613-6000

G.2 DARPA PROGRAM MANAGER

The DARPA Program Manager (PM) for this contract is:

Joseph Olive
DARPA/IPTO
3701 North Fairfax Drive
Arlington, VA 22203-1714

Phone: (571) 218-4920
Fax: (703) 807-4953
E-Mail: Joseph.Olive@darpa.mil

G.3 CONTRACT ADMINISTRATION DATA

Contractor's office, which will supervise and administer the resulting contract, is as follows:

Raytheon BBN Technologies Corp.
10 Moulton Street
Cambridge, MA 02138

POC: Mr. Jamie Clark
Phone: (617) 873-8049
Fax: (617) 873-2473
E-Mail: Jclark@bbn.com

G.4 GOVERNMENT CONTRACT ADMINISTRATOR

Administration of the contract, except for COR responsibilities, shall be performed by:

POC: Thupten Tsering
Phone: (520) 538-0407
Fax: (520) 538-3761
E-Mail: Thupten_D_Tsering@nbc.gov

Mailing Address:

U.S. Department of the Interior
National Business Center
Acquisition Services Directorate
Sierra Vista Division
Post Office Box 12924
Fort Huachuca, AZ 85670-2924

Commercial Address:

U.S. Department of the Interior
National Business Center
Acquisition Services Directorate
Sierra Vista Division
Augur and Adair Streets, Bldg #22208 (2nd Floor)
Fort Huachuca, AZ 85613-6000

G.5 DEFENSE CONTRACT AUDIT AGENCY (DCAA) OFFICE

The DCAA for this contract is:

Defense Contract Audit Agency (DCAA)
Northeast Region Merrimack Valley Branch Office
77 Fawcett Street
Cambridge, MA 02138

G.6 PAYMENT OFFICE

Payment will be made by the following office:

US Dept of the Interior
Office of the Secretary
National Business Center - Denver
Fiscal Support Section - D2730-FH
7301 W. Mansfield Avenue
Denver, Co 80235-2230

E-Mail: Ft_Huachuca_Pay_NBCDENVER@NBC.GOV

G.7 CONTRACT PAYMENTS

VOUCHER SUBMISSION FOR COST:

A. The contractor has not been approved by the Government for direct submission of invoices, the contractor shall submit an original and one copy of invoices to the Government Contract Administrator, and one copy to the Contracting Officer's Representative (COR) identified in Section G.

B. Proper invoices for this contract will be submitted on at least a monthly basis in accordance with FAR Clause 52.216-7, Allowable Cost and Payment (Dec 2002).

C. For payment of fee, the contractor will invoice for 8% fee incurred not to exceed $870,048.00. In accordance with FAR 52.216-8, Fixed Fee, the withhold fee of 15% is applied in the amount of $100,000.00.

D. All requests for payment shall be submitted in accordance with FAR Clause 52.232-25, Alt 1, Prompt Payment (Feb 2002).

SECTION H -- SPECIAL CONTRACT REQUIREMENTS

H.1 NOTIFICATION UNDER A COST REIMBURSMENT CONTRACT

a. To facilitate administration of the contract and the several provisions and clauses associated with cost/performance administration, including "Limitation of Funds," etc., the Contractor agrees to give timely notification to the Contracting Officer at any time he has reason to believe the total cost to the Government for the performance of this contract will be greater or less than the estimated cost set forth in the contract.

b. As part of such notification, the Contractor shall provide a revised cost estimate for accomplishing the specified work including a detailed statement of costs incurred, a detailed estimate of cost to complete, and an analysis, accompanied with an explanation, for the projected variance from estimated cost at completion. In the case of a projected increase, such analysis and explanation must include a description of action(s) taken or projected by the contractor in attempting to eliminate or mitigate the increase while discharging the management role of assuring performance within cost.

c. It is agreed by the parties that the notification contemplated above is to be accomplished only by separate written correspondence directed to the Contracting Officer and that no other form or "notification" (e.g. mention in any type of monthly progress or status report) will effect compliance. Further, notification to any individual other than the Administrative Contracting Officer shall not constitute compliance with this requirement

H.2 SUBCONTRACTING PLAN

Raytheon BBN Technologies Corp. Master Subcontracting Plan, dated 26 October 2007 and Individual Subcontracting Plan dated 26 August 2010, as submitted to and approved by the Contracting Officer in accordance with FAR 52.219-09 is herein attached at Section J and applicable to this contract.

H.3 AWARD MADE ON BEHALF OF A DOD AGENCY

This contract is issued by the Department of Interior/National Business Center (DOI/NBC) on behalf of a Department of Defense (DoD) Agency-Defense Advanced Research Projects Agency (DARPA).

H.4 KEY PERSONNEL REQUIREMENTS

Certain skilled/experienced professional and/or technical personnel are essential for successful contractor accomplishment of the work to be performed under this contract. These are defined as "key personnel" and listed:

Prime Contractor's Name:	Dr. John Makhoul
	Richard Schwartz
	Spyros Matsoukas

The contractor agrees key personnel shall not be removed from the contract work or be replaced without compliance with the following:

a. If one or more of the key personnel for whatever reason becomes, or is expected to become, unavailable for work under this contract for a continuous period exceeding 30-calendar days, or is expected to devote substantially less effort to the work than indicated in the proposal as initially anticipated, the contractor shall promptly notify the Government Contract Administrator

identified in Section G. Upon concurrence of the Contracting Officer or authorized representative, the contractor shall promptly replace such personnel with personnel of at least equal ability and qualifications.

b. All requests for approval of substitutions hereunder must be in writing and provide a detailed explanation of the circumstances necessitating the proposed substitution(s). They must contain a complete resume for the proposed substitute, and any other information requested by the Contracting Officer to approve or disapprove the proposed substitution. The Contracting Officer or duly authorized representative will evaluate such requests and promptly notify the contractor of the approval or disapproval thereof, in writing.

H.5 NOTICE OF INCORPORATION OF SECTION K

The following section of the solicitation will not be distributed with the contract; however, this section is incorporated by reference and forms a part of the resultant contract as though furnished in full text therewith:

Section K - Representations, Certifications and Other Statements of Offerors dated 15 June 2010.
Online Representations and Certifications Application (ORCA).

H.6 EQUIPMENT CLAUSE

The Contractor shall report any purchase of property, plant or equipment whose unit cost exceeds $100,000 to the DARPA Comptroller and the Contracting Officer. DoD FMR 7000.14R, Vol. 4, Chapter 6 and Federal Financial Accounting Standards No. 6 Accounting for Property, Plant and Equipment, applies.

H.7 NO-COST SETTLEMENT

When in the best interest of the Government and Contractor, a no-cost settlement may upon mutual agreement of the parties be executed in lieu of a termination due to the immature stage and uncertainties involved in contract performance in research and development efforts. In a no-cost settlement, all costs reimbursable, not previously paid, for the performance of the contract to the date of the termination are allowable; however, the costs for anticipatory profits or consequential damages resulting from the termination of this contract (or any subcontract) to include: accounting, legal, clerical and other expenses necessary for the preparation of settlement proposals and supporting data shall be excluded from the settlement proposal submitted by the contractor. The rights and remedies of the Government in this clause are in addition to any other rights and remedies provided by law under this contract.

H.8 PERMITS, TAXES, LICENSES, ORDINANCES AND REGULATIONS

The Contractor shall, at its own expense, obtain all necessary permits, give all notices, pay all license fees and taxes, comply with all Federal, State, Municipal, County and local Board of Health ordinances, rules and regulations applicable to the business carried on under this contract, and be responsible for all applicable State Sales and Use Taxes.

H.9 PUBLICATION

PUBLIC RELEASE OR DISSEMINATION OF INFORMATION

(a) There shall be no dissemination or publication, except within and between the Contractor and any subcontractors, of information developed under this contract or contained in the reports to be furnished pursuant to this contract without prior written concurrence of the COR. All scientific and technical reports will be issued with a classification of Distribution Statement C. In consideration of the nonprofit institutions of higher education who will be subcontractors under this contract, papers resulting from unclassified contracted fundamental research (as defined by DoD Instruction Number 5230.27) are exempt from prepublication controls and the requirement for prior written concurrence of the COR.

(b) When submitting material for written approval for open publication, the Contractor/Awardee must submit a request for public release to the DARPA Technical Information Office (TIO) and include the following information: 1) Document Information: document title, document author, short plain-language description of technology discussed in the material (approx. 30 words), number of pages (or minutes of video) and document type (briefing, report, abstract, article, or paper); 2) Event Information: event type (conference, principle investigator meeting, article or paper), event date, desired date for DARPA's approval; 3) DARPA Sponsor: DARPA Program Manager, DARPA office, and contract number; and 4) Contractor/Awardee's Information: POC name, e-mail and phone. Allow four weeks for processing; due dates under four weeks require a justification. Unusual electronic file formats may require additional processing time. Requests can be sent either via e-mail to tio@darpa.mil or via 3701 North Fairfax Drive, Arlington VA 22203-1714, telephone (571) 218-4235. Refer to www.darpa.mil/tio for information about DARPA's public release process.

H.10 TRAVEL AND PER DIEM

Travel and per diem costs claimed shall be in accordance with those rates/amounts allowed by the Joint Travel Regulation (JTR) in effect at the time travel is actually performed. There may be occasions where the contractor travels to high cost areas and is unable to remain in accordance with the Joint Travel Regulation. If required, the contractor will justify those costs to the Contracting Officer prior to performing the travel. The contractor agrees to use the most economical method of travel available. All travel other than what was proposed will have prior approval of the Contracting Officer's Representative as identified in Section G. This provision also applies to all subcontractors other than educational institutions performing travel in support of this contract. Universities will be reimbursed IAW their internal travel policies pursuant to OMB Circular A-21. Foreign travel requires 90 days approval by the Contracting Officer Representative.

H.11 ASSERTION OF DATA RIGHTS

The Contractor has asserted restrictions on the Government's use, release and disclosure of technical data and computer software. The technical data and computer software assertions are attached at Section J.

H.12 ACQUISITION OF COMPUTER EQUIPMENT AND RELATED ITEMS

The contractor and subcontractors are authorized to purchase information technology (IT) equipment and services for direct support of this contract not to exceed the value indicated in this contract clause. Any claimed costs in excess of this amount shall be unallowable in accordance with FAR 52.216-7, Allowable Cost and Payment.

Contract No. D10PC20015
NTE: $801,270.00

H.13 SF294 AND SF295 SUBCONTRACTING REPORTING REQUIREMENTS

Standard Form (SF) 294, Subcontracting Report for Individual Contracts, and SF 295, Summary Subcontract Report, required in accordance with FAR 19.704(a), shall be submitted via the new electronic Subcontracting Reporting System (eSRS), which is located at the following website: https://www.esrs.gov.

H.14 REPORTING REQUIREMENTS

a. QUARTERLY RESEARCH AND DEVELOPMENT (R&D) TECHNICAL STATUS REPORT

Description: The Quarterly R&D Technical Status Report provides a means to capture a comprehensive assessment of project goals, progress, and status. The report provides a section for programmatic data, technical progress, project plans, and issues and concerns. The report shall be formatted in accordance with the sample provided at Section J.

Submission Requirements:

--Frequency: Input four (04) times yearly for the duration of the contact.
--Reporting Period: Report on performance during previous quarter.
--Due Date: Submit within fifteen (15) calendar days after the end of the previous quarter. Post-award initial submission will be submitted within 30 calendar days of award.

b. QUARTERLY FINANCIAL STATUS REPORT

Description: The Quarterly Financial Status Report provides a means to capture a comprehensive assessment of expenditures. The report provides a section for incurred expenses and invoices submitted, both for the period and for the total cumulative effort which maintains a running balance of the funds remaining. Burn rates are easy to identify as all the pertinent information is side-by-side and easy to compare. The report shall be formatted in accordance with the sample provided at Section J.

Submission Requirements:

--Frequency: Input four (04) times yearly for the duration of the contact.
--Reporting Period: Report on performance during previous quarter.
--Due Date: Submit within fifteen (15) calendar days after the end of the previous quarter. Post-award initial submission will be submitted within 30 calendar days of award.

c. YEARLY TECHNICAL REPORT

Description: The Yearly Technical Report describes and disseminates to the analytical, scientific and technical community the precise nature and results of analytical studies, research development, test and evaluation (RDT&E) on an assigned task(s). The Yearly Technical Report may be definitive for the subject presented, exploratory in nature, or an evaluation of critical subsystem or of technical problems. The document content shall be clearly written, describe accomplishments and other facts adequately with no technical errors and be acceptable for release. There is no prescribed format.

Submission Requirement:
--Frequency: Input one time at the end of the first twelve months of performance.

--Due Date: Submit within fifteen (15) calendar days after the first twelve months of the contract.

d. FINAL TECHNICAL REPORT

Description: The Final Technical Report describes and disseminates to the analytical, scientific and technical community the precise nature and results of analytical studies, research development, test and evaluation (RDT&E) on an assigned task(s). The Scientific and Technical Report may be definitive for the subject presented, exploratory in nature, or an evaluation of critical subsystem or of technical problems. The document content shall be clearly written, describe accomplishments and other facts adequately with no technical errors and be acceptable for release. There is no prescribed format. If the Final Technical Report is marked unclassified, unlimited distribution, it should be accompanied by a letter certifying that it has been cleared for public release and/or sale; to include release and/or sale to foreign nationals.

Submission Requirement:
--Frequency: Input one time at the end of the contract.
--Reporting Period: Report on performance from start of contract to the end of contract.
--Due Date: Submit within fifteen (15) calendar days after the completion of the contract.

e. DISTRIBUTION OF REPORTS

The Final Report must be electronically submitted to the Program Manager (PM) as identified in Section G. The PM is responsible for reviewing and redacting any information prior to submitting Final Report to the Defense Technical Information Center (DTIC), DTIC-BCS, 8725 John J. Kingman Road, Suite 0944, Fort Belvoir, VA 22060-6218. The PM shall send an electronic copy to the Government Contract Administrator as identified in Section G, when the final report is sent to the DTIC.

(END OF SECTION H)

H.15 INCREMENTAL FUNDING

Of the total price of CLIN 0001 the sum of $1,000,000.00 is presently available for payment and allotted to this contract. It is anticipated from time to time additional funds will be allotted to this contract until the total price of these items is allotted. It is contemplated that funds presently allotted to this contract will cover the work to be performed, as limited by the provision of (b) below, until 31 January 2011.

b. For purposes of the Limitation of Funds Clause, Section I, 52.232-22, this provision will be the "Schedule". Accordingly, the contractor is working at his own risk if funds are expended beyond those currently allotted to this contract at any given time.
(END OF SECTION G)

H.16 EQUIPMENT - AGILE

The equipment purchased under the DARPA AGILE contract (HR0011-06-C-0022) may be used on a rent free, non-interference basis.

SECTION I -- CONTRACT CLAUSES

I.1 52.252-02 CLAUSES INCORPORATED BY REFERENCE

This contract incorporates one or more clauses by reference, with the same force and effect as if they were given in full text. Upon request, the Contracting Officer will make their full text available. Also, the full text of a clause may be accessed electronically at this/these address(es):

http://farsite.hill.af.mil/

Clause	Title	Date
52.202-01	Definitions	July 2004
52.203-03	Gratuities	April 1984
52.203-05	Covenant Against Contingent Fees	April 1984
52.203-06	Restrictions On Subcontractor Sales To The Government	September 2006
52.203-07	Anti-Kickback Procedures	July 1995
52.203-08	Cancellation, Rescission, and Recovery of Funds for Illegal or Improper Activity	January 1997
52.203-10	Price Or Fee Adjustment For Illegal Or Improper Activity	January 1997
52.203-12	Limitation on Payments to Influence Certain Federal Transactions	September 2007
52.203-13	Contractor Code of Business Ethics and Conduct	April 2010
52.204-04	Printed or Copied Double-Sided on Recycled Paper.	August 2000
52.204-07	Central Contractor Registration	April 2008
52.209-06	Protecting the Government's Interest When Subcontracting with Contractors Debarred, Suspended, or Proposed for Debarment	September 2006
52.215-02	Audit and Records--Negotiation	March 2009
52.215-08	Order of Precedence--Uniform Contract Format	October 1997
52.215-10	Price Reduction for Defective Cost or Pricing Data	October 1997
52.215-11	Price Reduction for Defective Cost or Pricing Data--Modifications	October 1997
52.215-12	Subcontractor Cost or Pricing Data	October 1997
52.215-13	Subcontractor Cost or Pricing Data--Modifications	October 1997
52.215-15	Pension Adjustments and Asset Reversions	October 2004
52.215-16	Facilities Capital Cost of Money	June 2003
52.215-18	Reversion or Adjustment of Plans for Postretirement Benefits (PRB) Other than Pensions	July 2005
52.215-21	Requirements for Cost or Pricing Data or Information Other Than Cost or Pricing Data--Modifications	October 1997
52.215-23	Limitations on Pass-Through Charges.	October 2009
52.216-07	Allowable Cost and Payment	December 2002
52.216-08	Fixed Fee	March 1997
52.219-08	Utilization of Small Business Concerns	May 2004
52.219-09	Small Business Subcontracting Plan	April 2008
52.219-16	Liquidated Damages-Subcontracting Plan	January 1999
52.219-28	Post-Award Small Business Program Representation	April 2009
52.222-03	Convict Labor	June 2003
52.222-21	Prohibition of Segregated Facilities	February 1999
52.222-24	Preaward On-Site Equal Opportunity Compliance Evaluation	February 1999
52.222-26	Equal Opportunity	March 2007
52.222-35	Equal Opportunity for Special Disabled Veterans, Veterans of the Vietnam Era, and Other Eligible Veterans	September 2006

52.222-36	Affirmative Action For Workers with Disabilities	June 1998
52.222-37	Employment Reports on Special Disabled Veterans, Veterans of the Vietnam Era, and Other Eligible Veterans	September 2006
52.222-50	Combating Trafficking in Persons	February 2009
52.222-54	Employment Eligibility Verification	January 2009
52.223-06	Drug Free Workplace	May 2001
52.223-14	Toxic Chemical Release Reporting	August 2003
52.225-13	Restrictions on Certain Foreign Purchases	June 2008
52.227-01 Alt I	Authorization And Consent (Jul 1995) - Alternate I	April 1984
52.227-02	Notice And Assistance Regarding Patent And Copy Infringement	December 2007
52.227-11	Patent Rights--Ownership by the Contractor	December 2007
52.228-07	Insurance--Liability To Third Persons	March 1996
52.230-02	Cost Accounting Standards	October 2008
52.230-06	Administration of Cost Accounting Standards	March 2008
52.232-09	Limitation on Withholding of Payments.	April 1984
52.232-17	Interest	October 2008
52.232-22	Limitation of Funds	April 1984
52.232-23	Assignment Of Claims	January 1986
52.232-25 Alt I	Prompt Payment (Oct 2003) Alternate I	February 2002
52.232-33	Payment by Electronic Funds Transfer-Central Contractor Registration	October 2003
52.233-01 Alt I	Disputes (Jul 2002) - Alternate I	December 1991
52.233-03 Alt I	Protest After Award (Aug 1996) - Alternate I	June 1985
52.233-04	Applicable Law for Breach of Contract Claim	October 2004
52.242-01	Notice of Intent to Disallow Costs	April 1984
52.242-03	Penalties for Unallowable Costs	May 2001
52.242-04	Certification of Final Indirect Costs	January 1997
52.242-13	Bankruptcy	July 1995
52.243-02 Alt V	Changes--Cost-Reimbursement (Aug 1987) -Alternate V	April 1984
52.244-05	Competition In Subcontracting	December 1996
52.244-06	Subcontracts for Commercial Items	April 2010
52.245-01	Government Property	June 2007
52.245-09	Use And Charges	June 2007
52.246-23	Limitation Of Liability	February 1997
52.246-25	Limitation of Liability--Services	February 1997
52.249-06	Termination (Cost-Reimbursement)	May 2004
52.249-14	Excusable Delays	April 1984
52.253-01	Computer Generated Forms	January 1991

I.2 52.215-19 NOTIFICATION OF OWNERSHIP CHANGES OCTOBER 1997

(a) The Contractor shall make the following notifications in writing:

(1) When the Contractor becomes aware that a change in its ownership has occurred, or is certain to occur, that could result in changes in the valuation of its capitalized assets in the accounting records, the Contractor shall notify the Administrative Contracting Officer (ACO) within 30 days.

(2) The Contractor shall also notify the ACO within 30 days whenever changes to asset valuations or any other cost changes have occurred or are certain to occur as a result of a change in ownership.

(b) The Contractor shall--

(1) Maintain current, accurate, and complete inventory records of assets and their costs;

(2) Provide the ACO or designated representative ready access to the records upon request;

(3) Ensure that all individual and grouped assets, their capitalized values, accumulated depreciation or amortization, and remaining useful lives are identified accurately before and after each of the Contractor's ownership changes; and

(4) Retain and continue to maintain depreciation and amortization schedules based on the asset records maintained before each Contractor ownership change.

(c) The Contractor shall include the substance of this clause in all subcontracts under this contract that meet the applicability requirement of FAR 15.408(k).

I.3 52.222-02 PAYMENT FOR OVERTIME PREMIUMS JULY 1990

(a) The use of overtime is authorized under this contract if the overtime premium does not exceed * 0 or the overtime premium is paid for work --

(1) Necessary to cope with emergencies such as those resulting from accidents, natural disasters, breakdowns of production equipment, or occasional production bottlenecks of a sporadic nature;

(2) By indirect-labor employees such as those performing duties in connection with administration, protection, transportation, maintenance, standby plant protection, operation of utilities, or accounting;

(3) To perform tests, industrial processes, laboratory procedures, loading or unloading of transportation conveyances, and operations in flight or afloat that are continuous in nature and cannot reasonably be interrupted or completed otherwise; or

(4) That will result in lower overall costs to the Government.

(b) Any request for estimated overtime premiums that exceeds the amount specified above shall include all estimated overtime for contract completion and shall--

(1) Identify the work unit; e.g., department or section in which the requested overtime will be used, together with present workload, staffing, and other data of the affected unit sufficient to permit the Contracting Officer to evaluate the necessity for the overtime;

(2) Demonstrate the effect that denial of the request will have on the contract delivery or performance schedule;

(3) Identify the extent to which approval of overtime would affect the performance or payments in connection with other Government contracts, together with identification of each affected contract; and

(4) Provide reasons why the required work cannot be performed by using multishift operations or by employing additional personnel.

* Insert either "zero" or the dollar amount agreed to during negotiations. The inserted figure does not apply to the exceptions in subparagraph (a)(1) through (a)(4) of the clause.

I.4 52.252-06 AUTHORIZED DEVIATIONS IN CLAUSES APRIL 1984

(a) The use in this solicitation or contract of any Federal Acquisition Regulation (48 CFR Chapter 1) clause with an authorized deviation is indicated by the addition of "(DEVIATION)" after the date of the clause.

(b) The use in this solicitation or contract of any U.S. Department of the Interior Acquisition Regulation Supplement (DIARS) clause with an authorized deviation is indicated by the addition of "(DEVIATION)" after the name of the regulation.

I.5 DEFENSE FEDERAL ACQUISITION REGULATION SUPPLEMENT

DEFENSE FEDERAL ACQUISITION REGULATION SUPPLEMENT (DFARS)

Incorporate By Reference: The below referenced DFARS clauses are applicable, will be incorporated in, and will form a part of the resultant contract as though furnished in full text herewith:

252.201-7000	Contracting Officer's Representative (Dec 1991)
252.203-7000	Requirements Relating to Compensation of Former DoD Officals (Jan 2009)
252.203-7001	Prohibition on Persons Convicted of Fraud or Other Defense-Contract-Related Felonies (Dec 2008)
252.203-7002	Requirement to Inform Employees of Whistleblower Rights (Jan 2009)
252.204-7000	Disclosure of Information (Dec 1991)
252.204-7003	Control of Government Personnel Work Product (Apr 1992)
252.204-7004	Alternate A, Central Contractor Registration (Sep 2007)
252.205-7000	Provision of Information to Cooperative Agreement Holders (Dec 1991)
252.209-7004	Subcontracting with Firms That Are Owned or Controlled by the Government of a Terrorist Country (Dec 2006)
252.215-7000	Pricing Adjustments (Dec 1991)
252.215-7002	Cost Estimating System Requirements (Dec 2006)
252.219-7003	Small Business Subcontracting Plan (DoD Contracts) (Apr 2007)
252.225-7001	Buy American Act and Balance of Payments Program (Jan 2009)
252.225-7002	Qualifying Country Sources as Subcontractors (Apr 2003)
252.225-7006	Quarterly Reporting of Actual Contract Performance Outside the United States (May 2007)
252.225-7012	Preference for Certain Domestic Commodities (Dec 2008)
252.227-7013	Rights in Technical Data--Noncommercial Items (Nov 1995)
252.227-7014	Rights in Noncommercial Computer Software and Noncommercial Computer Software Documentation. (June 1995)
252.227-7016	Rights in Bid or Proposal Information (June 1995)
252.227-7025	Limitations on the Use or Disclosure of Government Furnished Information Marked with Restrictive Legends (June 1995)
252.227-7027	Deferred Ordering of Technical Data or Computer Software (Apr 1988)
252.227-7030	Technical Data - Withholding of Payments (Mar 2000)
252.227-7037	Validation of Restrictive Markings on Technical Data (Sep 1999)
252.227-7039	Patents - Reporting of Subject Inventions (Apr 1990)
252.231-7000	Supplemental Cost Principles (Dec 1991)
252.232-7010	Levies on Contract Payments (Dec 2006)
252.235-7011	Final Scientific or Technical Report (Nov 2004)
252.243-7002	Requests for Equitable Adjustment (Mar 1998)

NOTE: Full text versions may be accessed electronically at this/these address(es):

http://farsite.hill.af.mil/

Incorporate in Full Text: The below DFARS clause is applicable, will be incorporated in, and will form a part of the resultant contract:

252.235-7010 Acknowledgement of Support and Disclaimer (May 1995)

(a) The Contractor shall include an acknowledgment of the Government's support in the publication of any material based on or developed under this contract, stated in the following terms: This material is based upon work supported by the Defense Advanced Research Projects Agency (DARPA) under Contract No. D10PC20015.

(b) All material, except scientific articles or papers published in scientific journals, must, in addition to any notices or disclaimers by the Contractor, also contain the following disclaimer: Any opinions, findings and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Defense Advanced Research Projects Agency (DARPA); or its Contracting Agent, the U.S. Department of the Interior, National Business Center, Acquisition Services Directorate, Sierra Vista Division.

I.6 1452.203-70 RESTRICTION ON ENDORSEMENTS JULY 1996

1452.203-70 RESTRICTION ON ENDORSEMENTS -
DEPARTMENT OF THE INTERIOR (JUL 1996)

The Contractor shall not refer to contracts awarded by the Department of the Interior in commercial advertising, as defined in FAR 31.205-1, in a manner which states or implies that the product or service provided is approved or endorsed by the Government, or is considered by the Government to be superior to other products or services. This restriction is intended to avoid the appearance of preference by the Government toward any product or service. The Contractor may request the Contracting Officer to make a determination as to the propriety of promotional material.

(End of clause)

I.7 1452.204-70 RELEASE OF CLAIMS JULY 1996

1452.204-70 RELEASE OF CLAIMS -- DEPARTMENT OF THE INTERIOR (JUL 1996)

After completion of work and prior to final payment, the Contractor shall furnish the Contracting Officer with a release of claims against the United States relating to this contract. The Release of Claims form (DI-137) shall be used for this purpose. The form provides for exception of specified claims from operation of the release.

(End of clause)

I.8 1452.215-70 EXAMINATION OF RECORDS BY DOI APRIL 1984

1452.215-70 EXAMINATION OF RECORDS BY THE DEPARTMENT OF THE INTERIOR (APR 1984)

For purposes of the Examination of Records by the Comptroller General clause of this contract (FAR 52.215-1), the Secretary of the Interior, the Inspector General, and their duly authorized representative(s) from the Department of the Interior shall have the same access and examination rights as the Comptroller General of the United States.

(End of clause)

I.9 1452.228-7 INSURANCE - LIABILITY TO THIRD PERSON APRIL 1984

1452.228-7 Insurance - Liability to Third Persons

(a) As prescribed in 1428.311-2, the clause at FAR 52.228-7, Insurance - Liability to Third Persons, shall be modified before insertion into solicitations and contracts by (1) changing the title of the clause to read:
"INSURANCE - LIABILITY TO THIRD PERSONS (APR 1984) (DEVIATIONS)";

and (2) changing the first sentence in subparagraph (c)(2) of the clause to read:

"For certain liabilities (and expenses incidental to such liabilities) to third persons not compensated by insurance or otherwise but subject to the 'Limitation of Cost' or 'Limitation of Funds' clause of this contract."

(b) As prescribed in FAR 52.103(a) and 52.107(f), the clause at FAR 52.252-6, Authorized Deviations in Clauses, shall be inserted into solicitations and contracts containing the clause in paragraph (a) of this section.

(End of clause)

I.10 1452.237-70 INFORMATION COLLECTION - DOI JULY 1996

1452.237-70 INFORMATION COLLECTION -- DEPARTMENT OF THE INTERIOR (JUL 1996)

If performance of this contract requires the contractor to collect information on identical items from ten or more public respondents, no action shall be taken or funds expended in the solicitation or collection of such information until the contractor has received from the Contracting Officer written notification that approval has been obtained from the Office of Management and Budget (OMB) pursuant to the Paperwork Reduction Act of 1980. The Contractor agrees to provide all information requested by the Contracting Officer which is necessary to obtain approval from OMB.

(End of clause)

I.11 52.244-02 SUBCONTRACTS JUNE 2007

(a) Definitions. As used in this clause--

"Approved purchasing system" means a Contractor's purchasing system that has been reviewed and approved in accordance with Part 44 of the Federal Acquisition Regulation (FAR).

"Consent to subcontract" means the Contracting Officer's written consent for the Contractor to enter into a particular subcontract.

"Subcontract" means any contract, as defined in FAR Subpart 2.1, entered into by a subcontractor to furnish supplies or services for performance of the prime contract or a subcontract. It includes, but is not limited to, purchase orders, and changes and modifications to purchase orders.

(b) When this clause is included in a fixed-price type contract, consent to subcontract is required only on unpriced contract actions (including unpriced modifications or unpriced delivery orders), and only if required in accordance with paragraph (c) or (d) of this clause.

(c) If the Contractor does not have an approved purchasing system, consent to subcontract is required for any subcontract that--

(1) Is of the cost-reimbursement, time-and-materials, or labor-hour type; or

(2) Is fixed-price and exceeds--

(i) For a contract awarded by the Department of Defense, the Coast Guard, or the National Aeronautics and Space Administration, the greater of the simplified acquisition threshold or 5 percent of the total estimated cost of the contract; or

(ii) For a contract awarded by a civilian agency other than the Coast Guard and the National Aeronautics and Space Administration, either the simplified acquisition threshold or 5 percent of the total estimated cost of the contract.

(d) If the Contractor has an approved purchasing system, the Contractor nevertheless shall obtain the Contracting Officer's written consent before placing the following subcontracts:

-none-

(e)(1) The Contractor shall notify the Contracting Officer reasonably in advance of placing any subcontract or

modification thereof for which consent is required under paragraph (b), (c), or (d) of this clause, including the following information:

(i) A description of the supplies or services to be subcontracted.

(ii) Identification of the type of subcontract to be used.

(iii) Identification of the proposed subcontractor.

(iv) The proposed subcontract price.

(v) The subcontractor's current, complete, and accurate cost or pricing data and Certificate of Current Cost or Pricing Data, if required by other contract provisions.

(vi) The subcontractor's Disclosure Statement or Certificate relating to Cost Accounting Standards when such data are required by other provisions of this contract.

(vii) A negotiation memorandum reflecting--

(A) The principal elements of the subcontract price negotiations;

(B) The most significant considerations controlling establishment of initial or revised prices;

(C) The reason cost or pricing data were or were not required;

(D) The extent, if any, to which the Contractor did not rely on the subcontractor's cost or pricing data in determining the price objective and in negotiating the final price;

(E) The extent to which it was recognized in the negotiation that the subcontractor's cost or pricing data were not accurate, complete, or current; the action taken by the Contractor and the subcontractor; and the effect of any such defective data on the total price negotiated;

(F) The reasons for any significant difference between the Contractor's price objective and the price negotiated; and

(G) A complete explanation of the incentive fee or profit plan when incentives are used. The explanation shall identify each critical performance element, management decisions used to quantify each incentive element, reasons for the incentives, and a summary of all trade-off possibilities considered.

(2) The Contractor is not required to notify the Contracting Officer in advance of entering into any subcontract for which consent is not required under paragraph (b), (c), or (d) of this clause.

(f) Unless the consent or approval specifically provides otherwise, neither consent by the Contracting Officer to any subcontract nor approval of the Contractor's purchasing system shall constitute a determination--

(1) Of the acceptability of any subcontract terms or conditions;

(2) Of the allowability of any cost under this contract; or

(3) To relieve the Contractor of any responsibility for performing this contract.

(g) No subcontract or modification thereof placed under this contract shall provide for payment on a cost-plus-a-percentage-of-cost basis, and any fee payable under cost-reimbursement type subcontracts shall not exceed the fee limitations in FAR 15.404-4(c)(4)(i).

(h) The Contractor shall give the Contracting Officer immediate written notice of any action or suit filed and prompt notice of any claim made against the Contractor by any subcontractor or vendor that, in the opinion of the Contractor, may result in litigation related in any way to this contract, with respect to which the Contractor may be entitled to reimbursement from the Government.

(i) The Government reserves the right to review the Contractor's purchasing system as set forth in FAR Subpart

44.3.

(j) Paragraphs (c) and (e) of this clause do not apply to the following subcontracts, which were evaluated during negotiations:

University of Maryland
Johns Hopkins University
Brno University
Cambridge University
Northeastern University

(End of clause)

I.12 52.243-07 NOTIFICATION OF CHANGES APRIL 1984

(a) Definitions. "Contracting Officer," as used in this clause, does not include any representative of the Contracting Officer.

"Specifically authorized representative (SAR)," as used in this clause, means any person the Contracting Officer has so designated by written notice (a copy of which shall be provided to the Contractor) which shall refer to this subparagraph and shall be issued to the designated representative before the SAR exercises such authority.

(b) Notice. The primary purpose of this clause is to obtain prompt reporting of Government conduct that the Contractor considers to constitute a change to this contract. Except for changes identified as such in writing and signed by the Contracting Officer, the Contractor shall notify the Administrative Contracting Officer in writing, within 30 calendar days from the date that the Contractor identifies any Government conduct (including actions, inactions, and written or oral communications) that the Contractor regards as a change to the contract terms and conditions. On the basis of the most accurate information available to the Contractor, the notice shall state--

(1) The date, nature, and circumstances of the conduct regarded as a change;

(2) The name, function, and activity of each Government individual and Contractor official or employee involved in or knowledgeable about such conduct;

(3) The identification of any documents and the substance of any oral communication involved in such conduct;

(4) In the instance of alleged acceleration of scheduled performance or delivery, the basis upon which it arose;

(5) The particular elements of contract performance for which the Contractor may seek an equitable adjustment under this clause, including--

(i) What contract line items have been or may be affected by the alleged change;

(ii) What labor or materials or both have been or may be added, deleted, or wasted by the alleged change;

(iii) To the extent practicable, what delay and disruption in the manner and sequence of performance and effect on continued performance have been or may be caused by the alleged change;

(iv) What adjustments to contract price, delivery schedule, and other provisions affected by the alleged change are estimated; and

(6) The Contractor's estimate of the time by which the Government must respond to the Contractor's notice to minimize cost, delay or disruption of performance.

(c) Continued performance. Following submission of the notice required by (b) above, the Contractor shall diligently continue performance of this contract to the maximum extent possible in accordance with its terms and conditions as construed by the Contractor, unless the notice reports a direction of the Contracting Officer or a communication from a SAR of the Contracting Officer,

in either of which events the Contractor shall continue performance; provided, however, that if the Contractor regards the direction or communication as a change as described in (b) above, notice shall be given in the manner provided. All directions, communications, interpretations, orders and similar actions of the SAR shall be reduced to writing and copies furnished to the Contractor and to the Contracting Officer. The Contracting Officer shall countermand any action which exceeds the authority of the SAR.

(d) Government response. The Contracting Officer shall promptly, within 30 calendar days after receipt of notice, respond to the notice in writing. In responding, the Contracting Officer shall either--

(1) Confirm that the conduct of which the Contractor gave notice constitutes a change and when necessary direct the mode of further performance;

(2) Countermand any communication regarded as a change;

(3) Deny that the conduct of which the Contractor gave notice constitutes a change and when necessary direct the mode of further performance; or

(4) In the event the Contractor's notice information is inadequate to make a decision under subparagraphs (d)(1), (2), or (3) above, advise the Contractor what additional information is required, and establish the date by which it should be furnished and the date thereafter by which the Government will respond.

(e) Equitable adjustments. (1) If the Contracting Officer confirms that Government conduct effected a change as alleged by the Contractor, and the conduct causes an increase or decrease in the Contractor's cost of, or the time required for, performance of any part of the work under this contract, whether changed or not changed by such conduct, an equitable adjustment shall be made--

(i) In the contract price or delivery schedule or both; and

(ii) In such other provisions of the contract as may be affected.

(2) The contract shall be modified in writing accordingly. In the case of drawings, designs or specifications which are defective and for which the Government is responsible, the equitable adjustment shall include the cost and time extension for delay reasonably incurred by the Contractor in attempting to comply with the defective drawings, designs or specifications before the Contractor identified, or reasonably should have identified, such defect. When the cost of property made obsolete or excess as a result of a change confirmed by the Contracting Officer under this clause is included in the equitable adjustment, the Contracting Officer shall have the right to prescribe the manner of disposition of the property. The equitable adjustment shall not include increased costs or time extensions for delay resulting from the Contractor's failure to provide notice or to continue performance as provided, respectively, in (b) and (c) above.

NOTE: The phrases "contract price" and "cost" wherever they appear in the clause, may be appropriately modified to apply to cost-reimbursement or incentive contracts, or to combinations thereof.

SECTION J -- LIST OF DOCUMENTS, EXHIBITS AND OTHER ATTACHMENTS

J.1 ATTACHMENTS

Attachment 1 - Assertion of Data Rights
Attachment 2 - DD Form 1423-1 Contract Data Requirements List
Attachment 3 - Subcontracting Plan
Attachment 4 - Progress, Status and Management Report Format

DATE: September 3, 2010
POC/Phone no.: Lisa Mattocks, (520) 533-8944

SUBJECT: Request for Review of Subcontracting Plan for Raytheon BBN Technologies Corp.
Contract No. D10PC20015
Contract Value: $13,082,545.00

MEMORANDUM FOR:

( ) 1-Level above the CO
(X) BUDS

Misebastian 9-7-2010

1. The following review action is requested:

(X) Approval of Subcontracting Plan (FAR 19.704)
( ) Approval of Waiver to the Subcontracting Plan requirement IAW DIARS 1419.705-2(c)

2. Clauses to be included in the final contract (please check appropriate box):

( ) 52.219-3, Notice of Total HUBZone Set-Aside (Jan 1999).
( ) 52.219-6, Notice of Total Small Business Set-Aside (June 2003).
( ) 52.219-7, Notice of Partial Small Business Set-Aside (June 2003).
(X) 52.219-8, Utilization of Small Business Concerns (May 2004).
(X) 52.219-9, Small Business Subcontracting Plan (April 2008).
( ) 52.219-10, Incentive Subcontracting Program (OCT 2001).
( ) 52.219-11, Special 8(a) Contract Conditions (FEB 1990).
( ) 52.219-12, Special 8(a) Subcontract Conditions (FEB 1990).
( ) 52.219-14, Limitations on Subcontracting (DEC 1996).
(X) 52.219-16, Liquidated Damages -- Subcontracting Plan (JAN 1999).
( ) 52.219-17, Section 8(a) Award (DEC 1996).
( ) 52.219-18, Notification of Competition Limited to Eligible 8(a) Concerns (JUN 2003).
( ) 52.226-1, Utilization of Indian Organizations and Indian-Owned Economic Enterprises
(X) 52.244-2, Subcontracts, Alt I (June 2007). **(Full Text)**
( ) 52.244-4, Subcontractors and Outside Associates and Consultants (Architect-Engineer Services) (AUG 1998).
(X) 52.244-5, Competition in Subcontracting (DEC 1996).
(X) 52.244-6, Subcontractors for Commercial Items (Mar 2009)

3. Remarks:

Raytheon BBN Technologies Corp. subcontracting plan was evaluated IAW FAR 19.704, 19.705-4 and 52.219-9. Raytheon BBN Technologies Corp. did not meet the established DoD goals:

	BBNT	FY2010 (DoD) Government Goal
Small Business	16.6%	37.2%
Small Disadvantaged Business	0.0%	5.0%
Woman-owned Small Business	0.1%	5.0%
HUBZone Small Business	0.0%	3.0%
SD Veteran-Owned Small Business	0.0%	3.0%
HBCU & Minority Institutions	0.0%	5.0%

Research and Development services are inherently labor intensive and highly complex. Small Business Subcontracting possibilities did not exist for this contract. The RATS program presents a number of challenging problems that require deep expertise and new ideas in several areas of work. The goal of the RATS program is to create technology capable of accurately determining speech activity regions, detecting key words, identifying language and speaker in highly degraded, weak and/or noisy communication channels. While the current scope does not require the support of any small business subcontractors for labor or other direct costs, Raytheon BBN Technologies Corporation has proposed to subcontract with five universities in an effort to accomplish the programs stated goals. The proposed subcontractors for this effort are Johns Hopkins University, University of Maryland, Cambridge University and Brno University of Technology.

Given today's state of the art, the target goals of the program require that the error rates in four technologies be reduced by about an order of magnitude over three years. Achieving this difficult feat will require a collaborative effort of multiple sites that have established themselves as leaders in the four technologies and in dealing effectively with high-noise environments. For this program, the particular subcontractors sought by Raytheon BBN Technologies Corporation were chosen due to their unique set of skills and experiences relevant to the goals of the RATS program. The professors and their staff will be responsible for writing software, performing experiments to improve accuracy in the four RATS applications in noise and expected to deliver software that will be integrated into the system for delivery. The partner universities have excelled in their area of expertise and their knowledge will be used to assist in achieving program goals.

Lisa A. Mattocks
Contracting Officer

SUBCONTRACTING PLAN CHECKLIST
FOR AN INDIVIDUAL CONTRACT PLAN

CONTRACTOR'S NAME: Raytheon BBN Technologies Corporation

REQUIREMENT: DARPA BAA 10-34 "Robust Automatic Transcription of Speech (RATS)"

PURPOSE OF REVIEW: Approval of Individual Subcontracting Plan for Contract D10PC20015

REFERENCES: FAR 19.704, 19.705-4, 52.219-9

1. Does the subcontracting plan include:

A separate percentage goals for using small business, HUBZone small business, small disadvantaged business, women-owned small business and veteran-owned small business concerns as subcontractors?
Yes
Where? Page(s) 1 (Individual Plan)

A statement of the total dollars planned to be subcontracted and statement of the total dollars planned to be subcontracted to smal business, HUBZone small business, small disadvantaged business, women-owned small business and veteran-owned small business concerns?
Yes
Where? Page(s) 1 (Individual Plan)

A description of the principal types of supplies and services to be subcontracted and an identification of types planned for subcontracting to small business, HUBZone small business, small disadvantaged business, women-owned small business and veteran-owned small business concerns?
Yes
Where? Page(s) 2-4 (Individual Plan)

A description of the method used to develop the subcontracting goals?
Yes
Where? Page(s) 4-5 (Individual Plan)

A description of the method used to identify potential sources for solicitation purposes?
Yes
Where? Page(s) 4-5 (Individual Plan)

A statement as to whether or not the offeror included indirect costs in establishing subcontracting goals, and a description of the method used to determine the proportionate share of indirect costs to be incurred with small business, HUBZone small business, small disadvantaged business, women-owned small business and veteran-owned small business concerns?
Yes
Where? Page(s) 2-4 (Individual Plan)

The name of the individual employed by the offeror who will administer the offeror's subcontracting program, and a description of the duties of the individual?
Yes
Where? Page(s) 1 (Individual Plan)

A description of the efforts the offeror will make to ensure that small business, HUBZone small business, small disadvantaged business, women- owned small business and veteran-owned small business concerns have an equitable opportunity to compete for subcontracts?
Yes
Where? Page(s) 5 (Individual Plan)

Assurances that the offeror will include the clause at 52.219-8, Utilization of Small Business Concerns (see 19.708(a)), in all subcontracts that offer further subcontracting opportunities, and that the offeror will require all subcontractors (except small business concerns) that receive subcontracts in excess of $550,000 ($1,000,000 for construction) to adopt a plan that complies with the requirements the clause at 52.219-9, Small Business Subcontracting Plan (see19.708(b))?
Yes
Where? Page(s) 10 (Master Plan)

Assurances that the offeror will--

- Cooperate in any studies or surveys as may be required;
- Submit periodic reports so that the Government can determine the extent of compliance by the offeror with the subcontracting plan;
- Submit Standard Form (SF) 294, Subcontracting Report for Individual Contracts, and SF 295, Summary Subcontract Report, following the instructions on the forms or as provided in agency regulations; and
- Ensure that its subcontractors agree to submit SF 294 and SF 295?

Yes
Where? Page(s) 12-14 (Master Plan)

A description of the types of records that will be maintained concerning procedures adopted to comply with the requirements and goals in the plan, including establishing source lists; and a description of the offeror's efforts to locate small business, HUBZone small business,small disadvantaged business, women-owned small business and veteran-owned small business concerns and to award subcontracts to them?
Yes
Where? Page 22-26 (Master Plan)

2. Complete the following chart:

A. Total Proposed Contract Effort	$13,082,545.00	
B. Total to be Subcontracted	$ 3,660,323.00	(27.5%)
1. Large Business	$ 2,997,235.00	(83.3% of B)
2. To Small Business		
a. HUBZone SB	$100.00	
b. Non-Disadvantaged SB	$598,888.00	(16.6% of A)
c. Disadvantaged SB	$1,600.00	
d. Women-Owned SB	$2,000.00	(0.1% of A)
e. Veteran-Owned SB	$400.00	
f. SD Veteran-owned SB	$100.00	

SUBCONTRACTING PLAN GOALS SUBMITTED UNDER
BBN TECHNOLOGIES CORP.
SUBCONTRACTING MASTER PLAN
IN ACCORDANCE WITH PUBLIC LAW 95-507

1. BBN Proposal No.: P10071-BBN

2. Contractor Name: Raytheon BBN Technologies Corp.

3. Name of Individual completing plan: John Forcucci
Telephone Number: (617) 873-8110

4.	A.	Total Proposed Contract Amount	$13,082,545	over 36 months		
	B.	Total Amount to be Subcontracted	$3,600,323		27.5 % of A	
		1. Total Large Business	$2,997,235	83.3 % of B;	22.9 % of A	
		2. Total Small Business	$603,088	16.7 % of B;	4.6 % of A	
		a. Non-Disadvantaged SB	$598,888	16.6 % of B;	4.6 % of A	
		b. Disadvantaged SB	$1,600	0.0 % of B;	0.0 % of A	
		c. Women-Owned SB	$2,000	0.1 % of B;	0.0 % of A	
		d. HBCU & MI	$0	0.0 % of B;	0.0 % of A	
		e. HUBZone SB	$100	0.0 % of B;	0.0 % of A	
		f. Veteran-Owned SB	$400	0.0 % of B;	0.0 % of A	
		g. Service-Disabled Vet SB	$100	0.0 % of B;	0.0 % of A	

5. The above goals, submitted in accordance with BBN Technologies' Annual Small Business and Small Disadvantaged Business Subcontracting Master Plan, include both Direct Goals and Indirect Goals.

Raytheon BBN Technologies Corp.

By:

John Forcucci
Small Business Liaison Officer

Date: September 2, 2010

BBN Proposal No.: P10071-BBN FY09 Raytheon BBN Technologies Corp. Annual Sales: $250,000,000

1. The Total Estimated Cost of Subcontracts includes Direct Cost Subcontracts and Indirect Cost Subcontracts. The breakdown is as follows:

	Direct Cost Subcontracts	Indirect Cost Subcontracts	Total	Percent of Sub Amt	Percent of Contract Amt
Non-Disadvantaged Small Business	$568,888	$30,000	$598,888	16.6%	4.6%
Small Disadvan. Business	$0	$1,600	$1,600	0.0%	0.0%
Women Owned Small Business	$0	$2,000	$2,000	0.1%	0.0%
Hist. Black Colleges & Univ and Minority Instit.	$0	$0	$0	0.0%	0.0%
HUBZone Small Business	$0	$100	$100	0.0%	0.0%
Veteran-Owned SB	$0	$400	$400	0.0%	0.0%
Service-Disabled Vet SB	$0	$100	100	0.0%	0.0%
Large Business	$2,957,235	$40,000	$2,997,235	83.3%	22.9%
TOTAL	$3,526,123	$74,200	$3,600,323	100.0%	27.5%

2. Indirect Cost Subcontracts

The Estimated Cost of Prime Contract is $13,082,545 over 36 months. This is 1.7% of the FY09 Annual Company Sales on an annualized basis. For this subcontracting plan, 0.2 % of orders placed with Small, Small Disadvantaged, Women-Owned, HUBZone, Veteran-Owned, Service-Disabled Veteran-Owned and Large businesses charged to Overhead and G&A Accounts are allocable to this contract.

a. Non-Disadvantaged Small Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $30,000,000 in orders with Small Businesses during its fiscal year 2010. We estimate about 50% of these orders are charges to Overhead and G&A accounts, or about $15,000,000. The Small Indirect Cost Subcontracts for this contract is 0.2 % of these orders or $30,000.

b. Small Disadvantaged Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $800,000 in orders with Small Disadvantaged Businesses during its fiscal year 2010. Virtually all 100% of these orders are charges to Overhead and G&A accounts. The Small Disadvantaged Indirect Cost Subcontracts for this contract is 0.2% of these orders or $1,600.

c. Women-Owned Small Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $1,000,000 in orders with Women Owned Small Business during its fiscal year 2010. Virtually all 100% of these orders are charges to Overhead and G&A accounts. The Women Owned Indirect Cost Subcontracts for this contract is 0.2% of these orders or $2,000.

d. HUBZone Small Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $50,000 in orders with HUBZone Small Business during its fiscal year 2010. Virtually all 100% of these orders are charges to Overhead and G&A accounts. The HUBZone Indirect Cost Subcontracts for this contract is 0.2% of these orders or $100.

e. Veteran-Owned Small Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $200,000 in orders with Veteran-Owned Small Business during its fiscal year 2010. Virtually all 100% of these orders are charges to Overhead and G&A accounts. The Veteran-Owned Indirect Cost Subcontracts for this contract is 0.2% of these orders or $400.

f. Service-Disabled Veteran-Owned Small Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $50,000 in orders with Service-Disabled Veteran-Owned Small Business during its fiscal year 2010. Virtually all 100% of these orders are charges to Overhead and G&A accounts. The Service-Disabled Veteran-Owned Indirect Cost Subcontracts for this contract is 0.2% of these orders or $100.

g. Large Business

Raytheon BBN Technologies Corp. anticipates placing a total of about $40,000,000 in orders with Large Businesses during its fiscal year 2010. We estimate about 50% of these orders are charges to Overhead and G&A accounts, or about $20,000,000. The Large Indirect Cost Subcontracts for this contract is 0.2% of these orders or $40,000.

P10071-BBN

Enclosed is Raytheon BBN Technologies Corp.'s (BBN) subcontracting plan for proposal P10071-BBN. Due to the nature of this proposal, BBN is only able to set overall goals of 16.7% of anticipated subcontracted dollars to non-disadvantaged small business, disadvantaged small business, women-owned small business, HUBZone small business, veteran-owned small business and service-disabled veteran-owned small business, of which 16.6% is allocated to non-disadvantaged small business, 0.0% to disadvantaged small businesses and 0.1% to women-owned small businesses. BBN is unable to establish goals of 42% to small businesses and 5% or more to disadvantaged small businesses and women-owned small businesses because the direct subcontracting procurement is of a single source nature.

On this effort BBN intends to subcontract with University of Maryland (UMD), Cambridge University (CU), Brno University of Technology (BUT), Johns Hopkins University (JHU) and Northeastern University

The RATS program presents a number of very challenging problems that require deep expertise and new ideas in several areas of work. The program requires the performance of four tasks in very high noise environments: speech activity detection (SAD), language identification (LID), speaker identification (SID), and keyword spotting (KWS). Given today's state of the art, the target goals of the program require that the error rates in the four technologies be reduced by about an order of magnitude over three years. Achieving this difficult feat will require a collaborative effort of multiple sites that have established themselves as leaders in the four technologies and in dealing effectively with high-noise environments, as well as new ideas that will lead to achieving the goals of the program. Especially important will be the integration of different approaches to the problem of dealing with noisy speech.

University of Maryland
Professor Shihab Shamma, University of Maryland (UMD): Dr. Shamma has distinguished himself as the inventor of the cortical spectrum, a three-dimensional representation of speech as a function of time, based on the cortical processing of the human auditory system. Using this rich representation, Dr. Shamma has been able to diminish the effects of the noise, while maintaining the quality of the speech. He has been able to demonstrate the effectiveness of his method for speech activity detection (SAD). We would like to take advantage for this type of processing for all four target applications.

Cambridge University
Professors Phil Woodland and Mark Gales, Cambridge University: Woodland and Gales have an international reputation as having developed one of the best speech recognition systems in the world. They have consistently won competitive evaluations sponsored by the US Government. They have also developed mathematically sophisticated methods for adapting their speech recognition to various types of channels, speakers, and noise. These methods will be of importance to the RATS effort as we design our state-of-the-art keyword spotting system at BBN to deal effectively with changing noise environments.

Brno University of Technology
Drs. Lukas Burget and Pavel Matejka, Brno University of Technology (BUT): Burget, Matejka, and their team at BUT have built a reputation in recent years as one of the best teams in the world for language ID (LID) and speaker ID (SID). They have entered a number of evaluations sponsored by NIST in LID and SID and placed first in many of the conditions tested. They have also pioneered Joint Factor Analysis as a way to improve performance and render their systems robust to many types of variations. By combining their technologies and the noise reduction methods from JHU and UMD with the novel ideas we have for improving performance in these areas, we believe we have a very good chance of meeting the challenging goals of the RATS program.

Johns Hopkins University
Professor Hynek Hermansky, Johns Hopkins University (JHU): Dr. Hermansky has made unique contributions to the robust processing of speech by inventing the modulation spectrum, based on what is known about the human auditory perceptual mechanism. For many types of noisy speech, the noise occupies a different region of the modulation spectrum than the speech, which makes it possible to concentrate further processing on the regions that are predominantly occupied by the speech.

Northeastern University
Dr. John Makhoul is an adjunct professor of Northeastern University. As such, he collaborates with Northeaster professors to identify graduate students to work via internship under the direction of Dr.

Makhoul, who serves as tl .dvisor. The work is conducted under BBN c. .cts from DARPA, in this case Robust Automatic Transcription of Speech (RATS), on state of the art speech research and technologies. Areas available include speech recognition, speaker identification, topic detection and tracking, machine translation, information extraction and retrieval, language understanding, and human-machine interaction.

There is a current Collaborative Research Agreement in place between BBN and Northeastern University, to which the proposed effort would be incorporated. This arrangement has been on-going with Northeastern since 1997 and has proved to be a satisfactory and cost-effective augmentation of BBN's technical efforts on other efforts and BBN believe will provide the same under this effort.

Due to the on-going nature of this arrangement learning curves are essentially eliminated and the current experience of the students and their assigned roles requires little supervision. For this research effort the graduate student will assist in the development of effective algorithms and software to improve the state of the art, including collection of appropriate data for system training, and analysis of system performance evaluations..

Although BBN is unable to establish disadvantaged small business and women-owned small business goals of 5% or more for this proposal, BBN will continue to offer maximum opportunities to small business, disadvantaged small business, woman-owned small businesses, HUBZone small businesses, veteran-owned small business and service-disabled veteran-owned small business for all its procurements..

Please feel free to contact John Forcucci, our small business liaison officer at 617-873-8110 should you require additional information.

PRINCIPAL TYPES OF SUPPLIES AND SERVICES TO BE SUBCONTRACTED
BBN Proposal No.: P10071-BBN

A. DIRECT

Business Categories	Sm. Bus.	Disadvan. Sm. Bus.	WOSB	HUBZone Sm. Bus.	Veteran Sm. Bus	Service-Disabled	Large Bus.
P. .cts and Services							
Engineering Services	X						X
Hardware/Software	X						X

B. INDIRECT

Business Categories	Sm. Bus.	Disadvan. Sm. Bus.	WOSB	HUBZone Sm. Bus.	Veteran Sm. Bus	Service-Disabled	Large Bus.
Products and Services							
Consulting Services	X	X	X		X		X
Office equipment and supplies	X	X					X
Transportation and freight services	X	X					X
Desktop computers	X	X	X	X			X
Computer supplies and software	X	X	X	X	X	X	X
Printing services	X	X					X
Electronic components	X	X					X
F. .cated sheet metal	X						X
P. .d circuit boards	X						
Machined Parts	X	X					
Facilities/Maintenance/Repairs/Operations	X	X	X	X	X		X
Facilities/Construction/Renovation	X	X			X		X

BBN TECHNOLOGIES

BBN TECHNOLOGIES CORP.
10 MOULTON STREET
CAMBRIDGE, MA 02138
TELEPHONE: (617) 873-8000

MEMORANDUM OF AGREEMENT REGARDING SMALL BUSINESS, HUBZONE SMALL BUSINESS, SMALL DISADVANTAGED BUSINESS, WOMEN-OWNED SMALL BUSINESS, VETERAN-OWNED SMALL BUSINESS AND SERVICE-DISABLED VETERAN-OWNED SMALL BUSINESS SUBCONTRACTING MASTER PLAN

Prepared by:

John Forcucci*
Small Business Liaison Officer
BBN Technologies Corp.
87 Fawcett Street
Cambridge, MA 02138
Tel: 617-873-3365

Signature: John Forcucci

Approved by:

William C. Earle
Vice President, Chief Accounting Officer
BBN Technologies Corp.
10 Moulton Street
Cambridge, MA 02138
Tel: 617-873-1824

Signature: William C. Earle

Period covered: October 1, 2007 through September 30, 2010

APPROVED: Cecilia A. Shine

Cecilia A. Shine
Administrative Contracting Officer
DCMA BOSTON

October 26, 2007
Date

*Please direct all correspondence to the Small Business Liaison Officer at this address.

Section I	Company Backgrounder
Section II	Statement of Corporate Policy on Small Business, HUBZone Small Business, Small Disadvantaged Business, Women-Owned Small Business, Veteran-Owned Small Business and Service-Disabled Veteran-Owned Small Business Utilization.
Section III	Dissemination of Policy Statements
Section IV	Responsibility for Implementation
Section V	Products and Services to be Subcontracted or Purchased.
Section VI	Action-Oriented Programs
Section VII	Internal Audit and Reporting Systems

SECTION 1

COMPANY BACKGROUNDER

BBN Technologies Corp.

COMPANY BACKGROUND

Addresses:
10 Moulton Street
Cambridge, MA 02138
(617) 873-8000 (telephone)
(617) 873-5011 (fax)
Internet Address: http://www.bbn.com

Offices:
Main campus is located in Cambridge, Massachusetts, with key locations in Arlington, VA, Columbia, MD and Middletown RI.

Strategy

BBN Technologies Corp. Profile:
For over 50 years, the BBN name has been synonymous with technical innovation. Since providing the acoustical design of the UN General Assembly Hall, to implementing and operating the ARPANET, the forerunner of today's Internet, we have pioneered a number of firsts: the first packet switch, the first router, and the first network email, which established the @ sign as an icon for the digital age. We also designed, built, and operated the Defense Data Network. Today BBN scientists and engineers are pioneering new innovations to help our customers work better and smarter.

BUSINESS UNIT: NETWORK TECHNOLOGIES

Business Team: Mobile Networking Systems

Mobile Networking Systems Description
The Mobile Networking Systems (MNS) team focuses on:

- Applied research and advanced development of mobile ad-hoc networks (i.e., networks in which connectivity maps and link states change rapidly and continuously)
- Incorporation of the results of this advanced development into real-world systems through systems architecture, software development, and integration and testing of key components of mission-critical software

The MNS staff includes system architects, software developers, computer scientists, network engineers, modeling and simulation specialists, experts in RF propagation, hardware developers, test engineers, and quality assurance and infrastructure support staff.

Our primary customers in the applied research and advanced development area are DARPA and DARPA-like agencies. Our primary customers for deliverable end products are US DoD agencies and prime contractors developing systems for these agencies.

Business Team: Network Research

Network Research Description
The Network Research team performs basic and applied research in a diverse set of technical areas within the overall networking domain. Novel solutions to customer requirements are designed, analyzed, and validated via simulations, prototypes, and/or measurements as appropriate. Prototypes developed by the team range from small proof-of-concept software implementations through large-scale software systems that may integrate 3rd-party hardware or software.

The team's areas of interest include:

- Network infrastructure security
- Network-based intrusion detection and mitigation
- New mobile & wireless network system architectures, algorithms, and protocols
- Effective service provision in the presence of severe network connectivity disruptions
- Network and computer forensic technologies and systems
- Very high speed hardware/software systems, including routers & network security devices
- Novel routing solutions
- End-to-end transport protocol research
- New network management methods and architectures
- Network understanding techniques
- Efficient spectrum management
- Satellite networks, including satellite/Internet integration

BUSINESS UNIT: INFORMATION & KNOWLEDGE TECHNOLOGIES

Business Team: Intelligent Computing

Intelligent Computing Description
The Intelligent Computing team conducts research, development, and advanced proof-of-concept prototyping. The team's technology focus is in the areas of artificial intelligence and the application of inexpensive immersive simulation technology to the development of training systems.

Current areas of interest include:

- Research and development of machine learning systems.
- Development of planning and decision support capabilities based on case-based and other machine reasoning techniques.
- Development of systems for training in real-time decision-making and tactics, based on lightweight simulation environments (LWSEs).

- Application of LWSEs to investigate and improve leadership and team interaction.

Business Team: Modeling & Planning Systems

Modeling & Planning Systems Description
The Modeling and Planning Systems team designs, builds, and maintains modeling, simulation, and planning tools and systems for the end user. The team's technology focus is on modeling, simulation, and planning, with application focus areas in transportation and logistics. Our bottom-line mission is to build systems that end users will use and trust.

Current areas of interest include:

- Detailed modeling of transportation of assets from CONUS to theater
- Detailed modeling of airports and seaports
- Integration and federation of transportation models
- User-centered design techniques for modeling and decision support systems
- Logisitcs Planning, Execution and Requirements Generation

BUSINESS UNIT: SENSING & CONTROL SYSTEMS

Business Team: Distributed Sensing Solutions

Distributed Sensing Solutions Description
The Distributed Sensing Solutions team solves challenging problems of high importance to our customers in the Department of Defense, the Intelligence Community, and Industry by rapidly prototyping systems for sensing complex physical phenomena, correlating them to events of interest, and producing high-value, actionable information from the sensed data. The team specializes in delivering networked sensor system solutions by combining a strong understanding of the underlying physics, high-fidelity modeling and simulation of complex signature and propagation phenomena, deep expertise in signal processing and computer science, proven capabilities in system development, and mature tools to enable rapid design cycles from theory to demonstration.

Current areas of interest include:

- High-fidelity modeling, prediction, analysis and exploitation of acoustic, seismic, infrasonic and electromagnetic signatures and propagation in complex media
- Advanced signal processing for extracting target and event signature information from sensed data
- Networking and fusion of multiple sensors for complex signature and event characterization, and for detection, localization, classification and tracking of challenging targets against complicated backgrounds
- Development and application of sensing systems for autonomous vehicles
- Development of high performance custom hardware using advanced COTS-based platforms and tools.
- Rapid prototyping of complex real-time multi-sensor systems

Business Team: Embedded Computing & Signal Technologies

Embedded Computing & Signal Technologies Description
The Embedded Computing & Signal Technologies team performs applied research, builds proof of concept demonstrations, prototypes, simulations, and highly reliable operational systems to solve our customers' hardest problems. The team specializes in the development of technologies, applications, and systems for transmission and acquisition of signals, embedded real time signal processing, control systems, environmental modeling and simulation, and underwater acoustics.

Embedded Computing & Signal Technologies Focus areas include:

- Operational systems development and integration for maritime signal acquisition, analysis and exploitation
- Modeling of environmental phenomena and simulation / stimulation of military operational systems
- Development and implementation of low bit-rate (<300bps) vocoder to improve speech intelligibility in noisy environments
- Modeling, experimentation, and data fusion to locate and characterize UGFs
- Exploitation of EM phenomenologies to locate targets of interest, and defeat of IEDs

BUSINESS UNIT: NATIONAL INTELLIGENCE RESEARCH & APPLICATIONS

Business Team: Enterprise Solutions

Enterprise Solutions Description
The Enterprise Solutions team performs advanced research and system development for complex data environments through intelligent middleware, comprehensive semantics, knowledge modeling, and sophisticated data management techniques.

Business Team: Information Systems & Security Solutions

Information Systems & Security Solutions Description
The Information Systems and Security team is engaged in projects that address information security and the development of systems concerning the collection, storage, analysis and security of information for all levels of computer networks.

The areas of focus include:

- Information Security (Public Key Infrastructure, protocol design, secure architectures)
- Cyber Situation Awareness
- Red Teaming
- Semantic Web
- Prototype Development

BUSINESS UNIT: SPEECH AND LANGUAGE TECHNOLOGIES

Business Team: Speech & Language Research

Business Team: Multi-Media Technologies & Solutions

Multi-Media Technologies & Solutions Description
The Multi-media Technologies and Solutions team leverages our research advances to develop integrated systems and prototypes, perform application-focused research and customization, and to deploy customized and turnkey software/hardware solutions in human language and document/image analysis areas.

Current areas of interest include:

- Information Extraction from Text
- Speech Recognition and Speech-to-Speech Translation
- Question Answering
- Topic Classification and Analysis/Mining of Unstructured Data
- Document Understanding
- Video Analysis and Content Extraction
- Image Retrieval
- Broadcast Monitoring
- Speech Triage
- Multimedia Content Exploitation

Speech & Language Research Description
The Speech and Language Research team is focused on advancing the state of the art in speech and natural language processing through the development of novel algorithms and techniques that enable us to develop better solutions for Government needs.

SECTION II

STATEMENT OF CORPORATE POLICY ON SMALL BUSINESS, HUBZONE SMALL BUSINESS, SMALL DISADVANTAGED BUSINESS, WOMEN-OWNED SMALL BUSINESS, VETERAN-OWNED SMALL BUSINESS AND SERVICE-DISABLED VETERAN-OWNED SMALL BUSINESS UTILIZATION

STATEMENT OF CORPORATE POLICY

It is BBN's policy as a government prime contractor and subcontractor to actively participate in the government's subcontracting program from small businesses, HUBZone small businesses, small disadvantaged businesses, women-owned small businesses, veteran-owned small businesses and service-disabled veteran-owned small businesses and other minority institutions as promulgated in Federal Acquisition Regulation (FAR) 52.219-8 and Public Laws 95-507, Section 211; 99-661, Section 207; 100-180, Section 806; 100-656.

BBN shall afford maximum practicable opportunity to small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns to participate in the performance of contracts awarded to BBN. Participation may be achieved by subcontracting part of the work required by a contract, as well as by purchasing supplies and services for contracts, internal jobs, overhead and G & A from small businesses, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

BBN shall enable small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns to be considered fairly as subcontractors and to be given an equitable opportunity to compete for subcontract awards. BBN will aid and encourage small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns by awarding orders on a competitive basis, whenever they are capable of fulfilling requirements and price, quality and delivery.

In accordance with the requirements of the FAR, BBN will comply with the following clauses. In addition, BBN will include these clauses, in accordance with the FAR, in contracts awarded by BBN to its subcontractors.

(a) FAR clause 52.219-8, Utilization of Small Business Concerns, shall be included in all contracts exceeding $100,000, unless the contract, together with all its subcontracts, is to be performed entirely outside of any State, territory, or possession of the United States, the District of Columbia and the Commonwealth of Puerto Rico or those which are for personal services in nature.

(b) FAR clause 52.219-9, Small Business Subcontracting Plan, shall be included in all solicitations for negotiated contracts or modifications (including contracts and modifications placed on a sole source basis) which (a) offer subcontracting possibilities, (b) are expected to exceed $550,000 ($1,000,000 for the construction of any public facility) and (c) are required to include clause 52.219-8, Utilization of Small Business Concerns, unless the acquisition is a set aside or is to be accomplished under the 8(a) program.

FAR clause 52.219-8: Utilization of Small Business Concerns (May 2004)

(a) It is the policy of the United States Government that small business concerns, veteran-owned small business concerns, serviced-disabled veteran-owned small business concerns, HUBZone small business concerns, small disadvantaged business concerns, and women-owned small

business concerns shall have the maximum practicable opportunity to participate in performing contracts let by any Federal agency, including contracts and subcontracts for subsystems, assemblies, components, and related services for major systems. It is further the policy of the United States that its prime contractors establish procedures to ensure the timely payment of amounts due pursuant to the terms of their subcontracts with small business concerns, veteran-owned small business concerns, serviced-disabled veteran-owned small business concerns, HUBZone small business concerns, small disadvantaged business concerns, and women-owned small business concerns.

(b) BBN hereby agrees to carry out this policy in the awarding of subcontracts to the fullest extent consistent with efficient contract performance. BBN further agrees to cooperate in any studies or surveys as may be conducted by the United States Small Business Administration or the awarding agency of the United States as may be necessary to determine the extent of BBN's compliance with this clause.

(c) Definitions. As used in this contract -

HUBZone small business concern means a small business concern that appears on the List of Qualified HUBZone Small Business Concerns maintained by the Small Business Administration.

Service-disabled veteran-owned small business concern means a small business concern not less than 51 percent of which is owned by one or more service-disabled veterans or, in the case of any publicly owned business, not less than 51 percent of the stock of which is owned by one or more service-disabled veterans and whose management and daily business operations are controlled by one or more of such individuals.

Service-disabled veteran means a veteran, as defined in 38 U.S.C. 101(2), with a disability that is service-connected, as defined in 38 U.S.C. 101 (16).

Small business concern means a small business as defined pursuant to Section 3 of the Small Business Act and relevant regulations promulgated pursuant thereto.

Small disadvantaged business concern means a small business that represents, as part of its offer, that it has received certification as a small disadvantaged business concern consistent with 13 CFR part 124, subpart B and it is identified, on the date of its representation, as a certified small disadvantaged business in the database maintained by the Small Business Administration (Central Contractor Registration - Dynamic Small Business Search).

Veteran-owned small business concern means a small business concern not less than 51 percent of which is owned by one or more veterans (as defined at 38 U.S.C. 101(2)) or, in the case of any publicly owned business, not less than 51 percent of the stock of which is owned by one or more veterans and whose management and daily business operations are controlled by one or more of such individuals.

Women-owned small business concern means a small business concern not less than 51 percent of which is owned by one or more women or, in the case of any publicly owned business, not less than 51 percent of the stock of which is owned by one or more women and whose management and daily business operations are controlled by one or more of such individuals.

(d) BBN acting in good faith may rely on written representations by its subcontractors regarding their status as either a small business concern, a veteran-owned small business concern, a service-disabled veteran-owned small business concern or a women-owned small business concern. Small disadvantaged business concerns and HUBZone small business concerns must supply a certificate from the Small Business Administration confirming their status.

FAR clause 52.219-9: Small Business Subcontracting Plan (September 2006)

(a) This clause does not apply to small business concerns.

(b) Definitions. As used in this clause –

"Commercial Item" means a product or service that satisfies the definition of commercial item in section 2.101 of the Federal Acquisition Regulation. It also means a product that in the opinion of the Contracting Officer differs only insignificantly from BBN's commercial product.

"Subcontract" means any agreement (other than one involving an employer-employee relationship) entered into by a Federal Government prime Contractor or subcontractor calling for supplies or services required for performance of the contract or subcontract.

(c) BBN, upon request by the Contracting Officer, shall submit and negotiate a subcontracting plan, where applicable, that separately addresses subcontracting with small business concerns, veteran-owned small business concerns, serviced-disabled veteran-owned small business concerns, HUBZone small business concerns, small disadvantaged business concerns, and women-owned small business concerns veteran-owned small business concerns and service-disabled veteran-owned small business concerns which shall be included in and made a part of the resultant contract. The subcontracting plan shall be negotiated within the time specified by the Contracting Officer. Failure to submit and negotiate the subcontracting plan shall make BBN ineligible for award of a contract.

(d) The BBN's subcontracting plan shall include the following:

(1) Goals, expressed in terms of percentages of total planned subcontracting dollars, for the use of small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns as subcontractors. BBN shall include all subcontracts that contribute to contract performance, and may include a proportionate share of products and services that are normally allocated as indirect costs. Direct procurement goals are established by review of the proposal/contract bill of materials to ensure that maximum practical opportunity shall be afforded to such businesses.

When developing indirect costs for a program, the dollar value of the program, on an annual basis, is divided by BBN's annual sales. This factor is then multiplied by the budgeted indirect procurements for large businesses, small businesses, HUBZone small business, small disadvantaged businesses, women-owned small businesses, veteran-owned small businesses and service-disabled veteran-owned small businesses and the resultant dollar amount is allocated as indirect procurement dollars to the subcontracting plan for the program.

(2) A statement of (i) total dollars planned to be subcontracted, (ii) total dollars planned to be subcontracted to small business concerns, (iii) total dollars planned to be subcontracted to HUBZone small business concerns, (iv) total dollars planned to be subcontracted to small disadvantaged business concerns, (v) total dollars planned to be subcontracted to women-owned small business concerns, (vi) total dollars planned to be subcontracted to veteran-owned small business concerns and (vii) total dollars planned to be subcontracted to service-disabled veteran-owned small business concerns.

(3) A description of the principal types of supplies and services to be subcontracted, and an identification of the types planned for subcontracting to (i) small business concerns, (ii) HUBZone small business concerns, (iii) small disadvantaged business concerns, (iv) women-owned small business concerns, (v) veteran-owned small business concerns and (vi) service-disabled veteran-owned small business concerns.

(4) A description of the method used to develop the subcontracting goals in (d) (1) above.

(5) A description of the method used to identify potential sources for solicitation purposes. BBN uses existing company source lists, the Dynamic Small Business Search Section of the Central Contractor Registration (CCR) which replaced the Procurement Marketing and Access Network (PRO-Net) of the Small Business Administration (SBA), veterans service organizations, the National Minority Supplier Development Counsel Vendor Information Service, the Research and Information Division of the Minority Business Development Agency in the Department of Commerce, or small, HUBZone, small disadvantaged, and women-owned small business trade associations.

(6) A statement as to whether or not BBN included indirect costs in establishing subcontracting goals, and a description of the method used to determine the proportionate share of indirect and overhead costs incurred with (i) small business concerns, (ii) HUBZone small business concerns, (iii) small disadvantaged business concerns, (iv) women-owned small business concerns, (v) veteran-owned small business concerns and (vi) service-disabled veteran-owned small business concerns.

(7) The name of the individual employed by BBN who will administer BBN's subcontracting program, and a description of the duties of the individual.

(8) A description of the efforts BBN will make to assure that small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns will have an equitable opportunity to compete for subcontracts.

(9) Assurances that BBN will include the clause in this contract, entitled "Utilization of Small Business Concerns" in all subcontracts that offer further subcontracting opportunities, and that BBN will require all subcontractors (except small business concerns) that receive subcontracts in excess of $550,000 ($1,000,000 for the construction of any public facility) to adopt a subcontracting plan similar to the plan agreed to by BBN.

(10) Assurance that BBN will (i) cooperate in any studies or surveys as may be required, (ii) submit periodic reports in order to allow the government to determine the extent of

compliance by BBN with the subcontracting plan, (iii) submit Standard Form (SF) 294, Subcontracting Report for Individual Contracts, and/or SF 295, Summary Subcontract Report, in accordance with the instructions on the forms, and (iv) ensure that its subcontractors agree to submit Standard Forms 294 and 295.

(11) A description of the types of records BBN will maintain concerning procedures that have been adopted to comply with the requirements and goals in the plan, including establishing source lists; and a description of its efforts to locate small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns and award subcontracts to them. The records shall include at least the following (on a plant-wide or company-wide basis, unless otherwise indicated):

(i) Source lists (e.g., Central Contractor Registration - Dynamic Small Business Search of the Small Business Administration), guides, and other data that identify small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

(ii) Organizations contacted in an attempt to locate sources that are small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

(iii) Records on each subcontract solicitation resulting in an award of more than $100,0000 indicating (a) whether small business concerns were solicited, and if not, why not; (b) whether HUBZone small business concerns were solicited, and if not why, (c) whether small disadvantaged business concerns were solicited, and if not, why not; (d) whether women-owned small business concerns were solicited and if not, why not, (e) whether veteran-owned small business concerns were solicited, and if not, why not, (f) whether service-disabled veteran-owned small business concerns were solicited, and if not, why not, and (g) if applicable, the reason award was not made to a small business concern.

(iv) Records of any outreach efforts to contact: (a) trade associations, (b) business development organizations, (c) conferences and trade fairs and (d) veteran service organizations.

(v) Records of internal guidance and encouragement provided to buyers through, (a) workshops, seminars, training, etc., and (b) monitoring performance to evaluate compliance with the program's requirements.

(vi) On a contract-by-contract basis, records to support award data submitted by BBN to the Government, including the name, address, and business size of each subcontractor. Contractors having commercial plans need not comply with this requirement.

(e) In order to effectively implement this plan to the extent consistent with efficient contract performance, BBN shall perform the following functions:

(1) Assist small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns by arranging solicitations, time for the preparation of bids, quantities, specifications, and delivery schedules so as to facilitate the participation by such concerns. Where BBN's lists of potential small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business subcontractors are excessively long, reasonable effort shall be made to give all such small business concerns an opportunity to compete over a period of time.

(2) Provide adequate and timely consideration of the potentialities of small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns in all "make-or-buy" decisions.

(3) Counsel and discuss subcontracting opportunities with representatives of small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

(f) If a commercial product is offered, the subcontracting plan required by this clause may relate to BBN's production generally, for both commercial and non-commercial products, rather than solely to the Government contract. In these cases, BBN shall, with the concurrence of the Contracting Officer, submit one company-wide or division-wide plan.

The plan shall be reviewed for approval by the agency awarding BBN its first prime contract requiring a subcontracting plan during the fiscal year, or by an agency satisfactory to the Contracting Officer.

The approved plan shall remain in effect during BBN's fiscal year(s) for all of BBN's commercial products.

(g) Prior compliance of BBN with other such subcontracting plans under previous contracts will be considered by the Contracting Officer in determining the responsibility of BBN for award of the contract.

(h) The failure of BBN or subcontractor to comply in good faith with (1) the clause of this contract entitled "Utilization of Small Business Concerns"; or (2) an approved plan required by this clause, shall be a material breach of the contract.

SECTION III

DISSEMINATION OF POLICY STATEMENTS

DISSEMINATION OF POLICY STATEMENTS

In order to effectively disseminate BBN's corporate policy regarding Utilization Small Business Concerns, as promulgated in Federal Acquisition Regulation (FAR) 52.219-8, and Public Law 95-507 Section 211, a memo, including a copy of corporate policy, has been issued to all department managers and all other respective employees effected by said policy.

Additionally, BBN's corporate policy will become an integral part of the BBN Purchasing Procedures Manual, which is distributed to, and followed by, all of BBN's purchasing/buying personnel.

SECTION IV

RESPONSIBILITY FOR IMPLEMENTATION

RESPONSIBILITY FOR IMPLEMENTATION

The Small Business Liaison Officer, John Forcucci, will cooperate with purchasing management and will assume responsibility of the following:

(1) Maintain liaison on small business matters with BBN's Cognizant Federal Agency (DCMA, Boston, 495 Summer St, Boston MA 02210);

(2) Supervise compliance with the "Utilization of Small Business Concerns" clause;

(3) Administer the company's Small Business, HUBZone Small Business, Small Disadvantaged Business, Women-Owned Small Business, Veteran-Owned Small Business and Service-Disabled Veteran-Owned Small Business Concerns Subcontracting Program; and

(4) Distribute to Purchasing personnel data received from DCMA Deputy for Small Business and the Small Business Administration; and

(5) Assure that buying personnel comply with the provisions of the program;

(6) Assure that Purchasing personnel shall serve on committees rendering "Make-or-Buy" decisions and give due consideration to potentialities of small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

(7) Determine the "large" or "small" status of all suppliers.

(8) When inquiries are received from other companies that are soliciting business from BBN under the Small Business Program, they shall be referred to the Small Business Liaison Officer, who will refer them to the applicable buying personnel for consideration;

(9) Invitations to BBN to participate in exhibits, conferences, or shows sponsored by the Small Business Administration or otherwise clearly connected with advancement of small business concerns shall be referred to the company Small Business Liaison Officer for decision as to BBN's participation.

(10) Prepare and submit subcontract goals for individual proposals/contracts utilizing the form entitled "Subcontracting Plan Goals submitted under BBN Technologies Corp. Subcontracting Master Plan in accordance with Public Law 95-507" (See Exhibit A).

(11) Prepare Standard Form 294, Subcontracting Report for Individual Contracts, and Standard Form 295, Summary Subcontract Report, in accordance with the instructions on the forms.

(12) Cooperate with the DCMA Contracting Officer in any studies or surveys of our subcontracting procedures that he/she may wish to conduct.

SECTION V

PRODUCTS AND SERVICES TO BE SUBCONTRACTED

OR PURCHASED

PRODUCTS AND SERVICES THAT ARE COMMONLY
PURCHASED BY BBN TECHNOLOGIES

- Consulting/Engineering services
- Office equipment and supplies
- Transportation and freight services
- Computer Supplies
- Desktop Computers
- Computer Software
- Printing services
- Electronic components
- Fabricated sheet metal
- Printed circuit boards
- Facilities/Maintenance/Repairs/Operations
- Facilities/Construction/Renovation

SECTION VI

ACTION-ORIENTED PROGRAMS

ACTION-ORIENTED PROGRAMS

Purchasing Personnel will:

1) Assure that small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns have equitable opportunity to quote by:

   a) Reviewing all requisitions for small business potential.

   b) Interviewing all sales representatives who call upon the Purchasing organization.

   c) Using small business listings (obtainable from the regional offices of the Small Business Administration) to develop new sources of supply.

   d) Referring to indexed lists of small business suppliers (Dynamic Small Business Search Section of the Central Contractor Registration (CCR) of the Small Business Administration and National Minority Supplier Development Counsel Vendor Information Service).

   e) Exchanging information on new and potential suppliers with other company locations.

   f) Allowing sufficient time for the preparation of suppliers' quotations.

   NOTE: Where long lists of potential small business suppliers are maintained, reasonable effort shall be made to give all such small business concerns an opportunity to compete over a period of time.

2) Request quotations and place contracts and purchase orders with small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns when consistent with good business practice.

3) When buying considerations with respect to two or more potential suppliers are equal, the business shall be placed with the small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

4) If practical, and only with approval of the ordering organization, adjust quantities, specifications, and/or delivery schedules to enable placing business with small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

5) Buying personnel shall utilize source lists, (e.g., CCR), guides and other data that identify small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business concerns.

6) When placing purchasing arrangements valued in excess of $100,000 (except those to be performed entirely outside of and State, territory, or possession of the United States, the District of Columbia, and the Commonwealth of Puerto Rico or for personal services contracts), incorporate the provisions of the following clause by reference:

"FAR Clause 52.219-8 Utilization of Small Business Concerns"

In contracts that contain the clause referenced above and are valued in excess of $550,000 ($1,000,000 for construction of any public facility), insert the following additional reference:

"FAR Clause 52.219-9 Small Business Subcontracting Plan"

NOTE: The above FAR references, and the introductory and transitional language necessary to incorporate such clauses by reference, are preprinted on the appropriate contract riders in the company instructions, and are also preprinted on the Government Contract provisions of the Purchase Order form.

SECTION VII

INTERNAL AUDIT AND REPORTING SYSTEMS

INTERNAL AUDIT AND REPORTING SYSTEMS

1) Monthly reports will be prepared via an electronic file that tracks dollar commitments and purchase order volume by the business categories and geographical areas. These reports, which can be sorted by individual buyer, record both departmental performance and individual buyer performance. The electronic file also permits BBN to track purchase order commitments by contract number.

2) For all purchase orders exceeding $100,000, the purchasing files shall also indicate:

   (a) The reason for non-solicitation of small business, if such was the case; or

   (b) The reason for the failure of small business to receive the award, if such was the case when small business was solicited.

3) The BBN Purchasing department maintains an Affirmative Action Acquisition reference library that contains literature on small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business consultants and suppliers and a nationwide selection of Minority Purchasing directories from areas such as New England, New York, Los Angeles and Washington, D.C. BBN also has access to Dynamic Small Business Search Section of the Central Contractor Registration (CCR) of the Small Business Administration.

4) Individual goals will be set for each buyer and the Small Liaison Officer shall review the buyer's performance periodically. Their performance in meeting small business, HUBZone small business, small disadvantaged business, women-owned small business, veteran-owned small business and service-disabled veteran-owned small business goals will become part of the buyers' overall job performance evaluation.

Progress, Status and Management Report

Period Covered by the Report

Date of Report:

Project Title: Robust Automatic Transcription of Speech (RATS)
Contract Number: D10PC20015
Total Dollar Value: $13,082,545.00
Program Manager: Joseph Olive

Submitted by:

[Name]
[Address]

Telephone:
Fax:
Email:

Security Classification – Unclassified

Distribution List and Addresses

One report to:

Joseph Olive
Defense Advanced Research Projects Agency
3701 N. Fairfax Drive
Arlington, VA 22203-1714
E-mail: joseph.olive@darpa.mil

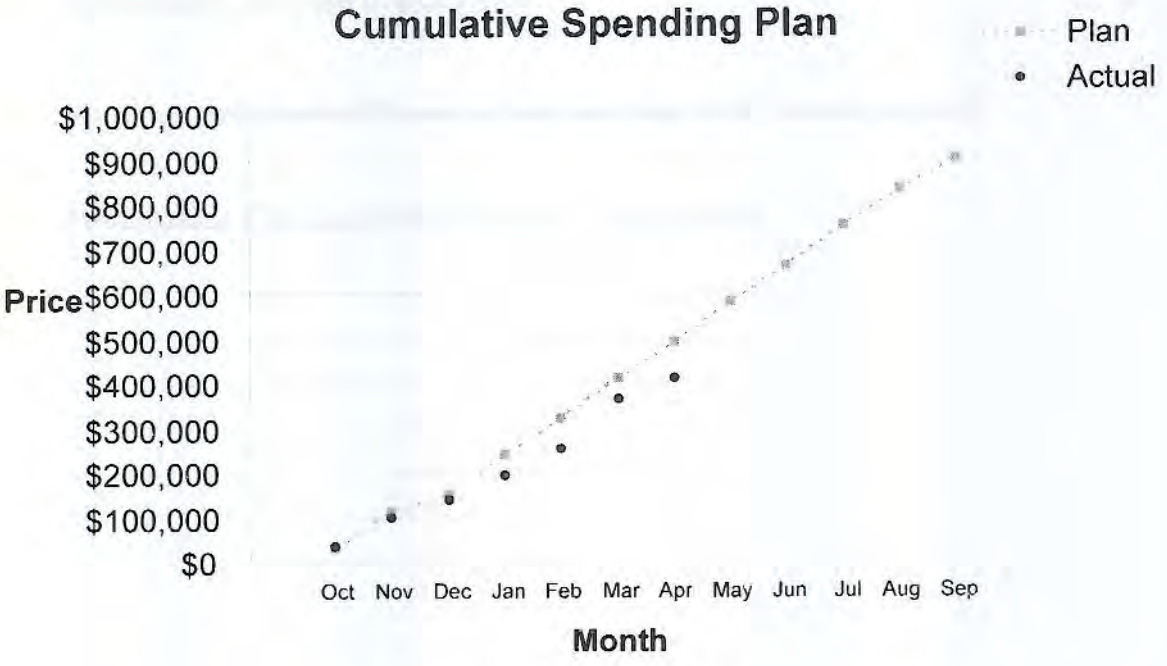
One report to:

Lawrence H Carter/COR
DOI/NBC
Acquisition Services Directorate
P.O. Box 192924
Fort Huachuca, AZ 85670-2924
E-mail: Lawrence_H_Carter@nbc.gov

Technical Information – Financial Management

1. **Technical Progress / Quarterly Expenditure Report (Please provide cumulative spending graph).**

Figure 1. Example Graph



Please provide Phase 1 schedule of tasks and events for this quarter, with financial expenditures broken down by task.

Task 1 – (Task Description): $
Task 2 – (Task Description): $
Task 3 – (Task Description): $

Total expenditures for the quarter - $

Actual Cost versus Planned Costs

	Current Quarter's Cost* ($)	Inception to Date ($)	Phase 1 Cost ($)
Plan			
Actual			
Difference			

* Current Quarter is (DATE) through (DATE).

CONTRACT DATA REQUIREMENTS LIST

(1 Data Item)

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 110 hours 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 the burden, to the Department of Defense, Executive Services and Communications Directorate (0704 0188). Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **Please do not return your form to the above organization. Send completed form to the Government Issuing Contracting Officer for the Contract/PR No. listed in Block E.**

A. CONTRACT LINE ITEM NO.	B. EXHIBIT	C. CATEGORY:
0001	A	TDP ____ TM ____ OTHER ____

D. SYSTEM/ITEM	E. CONTRACT/PR NO.	F. CONTRACTOR
	D10PC20015	Raytheon BBN Technologies Corp.

1. DATA ITEM NO.	2. TITLE OF DATA ITEM	3. SUBTITLE
A0001	Scientific and Technical Reports	

4. AUTHORITY *(Data Acquisition Document No.)*	5. CONTRACT REFERENCE	6. REQUIRING OFFICE
DI-MISC-81612	Section J	

7. DD 250 REQ	9. DIST STATEMENT REQUIRED	10. FREQUENCY	12. DATE OF FIRST SUBMISSION
8. APP CODE: A		11. AS OF DATE	13. DATE OF SUBSEQUENT SUBMISSION

16. REMARKS

17. PRICE GROUP

18. ESTIMATED TOTAL PRICE

14. DISTRIBUTION

a. ADDRESSEE	b. COPIES		
	Draft	Final	
		Reg	Repro
15. TOTAL →	0	0	0

G. PREPARED BY	H. DATE	I. APPROVED BY	J. DATE

REPRESENTATIONS AND CERTIFICATIONS OF OFFERORS

apply to technical data or computer software that will be generated under the resulting contract. Notification and identification is not required for restrictions based solely on copyright.

(c) Offers submitted in response to this solicitation shall identify, to the extent known at the time an offer is submitted to the Government, the technical data or computer software that the Offeror, its subcontractors or suppliers, or potential subcontractors or suppliers, assert should be furnished to the Government with restrictions on use, release, or disclosure.

(d) The Offeror's assertions, including the assertions of its subcontractors or suppliers or potential subcontractors or suppliers, shall be submitted as an attachment to its offer in the following format, dated and signed by an official authorized to contractually obligate the Offeror:

Identification and Assertion of Restrictions on the Government's Use, Release, or Disclosure of Technical Data or Computer Software.

The Offeror asserts for itself, or the persons identified below, that the Government's rights to use, release, or disclose the following technical data or computer software should be restricted:

Technical Data Computer Software to be Furnished With Restrictions* (LIST)*****	Basis for Assertion** (LIST)	Asserted Rights Category*** (LIST)	Name of Person Asserting Restrictions**** (LIST)
Software Suite to be developed under the RATS program including existing libraries as needed	Developed with mixed funding	Negotiated Government Purpose Rights with a 20 year term	Raytheon BBN Technologies Corp.
Byblos Software	Multiple Commercial Sales	Commercial	Raytheon BBN Technologies Corp.

*For technical data (other than computer software documentation) pertaining to items, components, or processes developed at private expense, identify both the deliverable technical data and each such item, component, or process. For computer software or computer software documentation identify the software or documentation.

**Generally, development at private expense, either exclusively or partially, is the only basis for asserting restrictions. For technical data, other than computer software documentation, development refers to development of the item, component, or process to which the data pertain. The Government's rights in computer software documentation generally may not be restricted. For computer software, development refers to the software. Indicate whether development was accomplished exclusively or partially at private expense. If development was not accomplished at private expense, or for computer software documentation, enter the specific basis for asserting restrictions.

***Enter asserted rights category (e.g., government purpose license rights from a prior contract, rights in SBIR data generated under another contract, limited, restricted, or government purpose rights under this or a prior contract, or specially negotiated licenses).

Volume 1 – Technical Proposal

Section 1 – Administrative

DARPA BAA 10-34 Confirmation Sheet	
Your Cover Sheet Has Been Accepted!	
User Id:	Diane
Technical Topic Area:	Algorithmic Development and Signal Processing
Title of Effort: **Contractors Reference Number**	RATS PATROL P10071-BBN

Tech/PI Contact	Admin Contact
Dr. John Makhoul Raytheon BBN Technologies Corp. 10 Moulton Street Cambridge, MA 02138 Country: USA Email: makhoul@bbn.com Phone: (617) 873-3332 FAX: (617) 873-2473 Type of Business: Other Large Business	Mr. Jamie Clark Raytheon BBN Technologies Corp. 10 Moulton Street Cambridge, MA 02452 Country: USA Email: jclark@bbn.com Phone: (617) 873-8049 FAX: (617) 873-2473

Comments Regarding This Effort	
Comments:	

Funding Information:

Base Funding Information			Contractual Options Information		
Total = $13,082,545 **Duration= 36 months**			**Total = $982,344** **Duration= 24 months**		
Base Funding Year	**Amount (in dollars)**	**Duration (Months)**	**Option**	**Amount (in dollars)**	**Duration (Months)**
Year 1	$5,293,372	12+	**1**	$482,785	12
Year 2	$3,895,062	12	**2**	$499,559	12
Year 3	$3,894,111	12	**3**		
Year 4			**4**		
Year 5			**5**		

Funding Comments Regarding This Effort	
Funding Comments:	

Subcontractor Information:

Company Name	Point of Contact Name
University of Maryland	Dr. Shihab Shamma
Johns Hopkins University	Dr. Hynek Hermansky
Brno University	Dr. Lukas Burget
Cambridge University	Dr. Mark Gales

Date Proposal Expires: 9/29/2010
Last Update Occurred: 3/26/2010 6:32:40 PM

RATS PATROL

Volume 1 – Technical Proposal
Section I – Administrative

DARPA BAA 10-34
BBN Proposal No.: P10071-BBN
Technical Area 1: Algorithmic Development and Signal Processing

March 28, 2010

Prepared by (Prime Contractor):
Raytheon BBN Technologies Corp.
10 Moulton Street
Cambridge, MA 02138
Type of Business: OTHER LARGE BUSINESS

Submitted to:
DARPA/IPTO (Dr. Joseph Olive)
ATTN: BAA 05-28
3701 N. Fairfax Drive
Arlington, VA 22203-1714:

Technical Point of Contact:
Dr. John Makhoul
Raytheon BBN Technologies Corp.
10 Moulton Street
Cambridge, MA 02138
Phone: 617-873-3332
Fax: 617-873-2473
Email: makhoul@bbn.com

Administrative Point of Contact:
Jamie Clark
Raytheon BBN Technologies Corp.
10 Moulton Street
Cambridge, MA 02138
Phone: 617-873-8049
Fax: 617-873-2473
Email: jclark@bbn.com

Subcontractors: Johns Hopkins University, University of Maryland, Cambridge University, Brno University of Technology

Cost Summary

	Phase 1 (18 m)	Phase 2 (12 m)	Phase 3 (12 m)	Total
Base Proposal	$5,293,372	$3,895,062	$3,894,111	$13,082,545
Option	N/A	$482,785	$499,559	$982,344

Notice: Use and Disclosure of Data

The data in this proposal shall not be disclosed outside the Government and shall not be duplicated, used or disclosed in whole or in part for any purpose other than to evaluate this proposal; provided, that if a contract is awarded to this offeror as a result of or in connection with the submission of these data, the Government shall have the right to duplicate, use or disclose the data to the extent provided in this contract. This restriction does not limit the Government's right to use information contained in these data if they are obtained from another source without restriction.

Table of Contents
Section 2 – Technical Details

1 INNOVATIVE CLAIMS 1

2 PROPOSAL ROADMAP 3

3 TECHNICAL APPROACH 6

3.1 Introduction 6
3.2 Achieving Robustness in Noise 6
3.2.1 Noise-robust Feature Extraction 6
3.2.2 Data-guided Trainable Feature Projection 7
3.2.3 Multi-resolution Tokenization 8
3.2.4 Noise Compensation 9
3.2.5 Fusion-based Classification 10
3.3 Application-Specific Methods 11
3.3.1 Speech Activity Detection (SAD) 11
3.3.2 Speaker Identification (SID) 12
3.3.3 Language Identification (LID) 12
3.3.4 Keyword Spotting (KWS) 13
3.4 Achieving the Program Goals 14
3.5 Software Integration and Deliveries 14
3.6 Speech Transcription Option 15

4 COMPARISON WITH CURRENT TECHNOLOGY 16

4.1 Noise Compensation 16
4.2 Speech Activity Detection 16
4.3 Speaker Identification 16
4.4 Language Identification 17
4.5 Keyword Spotting 17

5 STATEMENT OF WORK 18

6 INTELLECTUAL PROPERTY 28

7 SCHEDULE AND MILESTONES 29

8 PERSONNEL, QUALIFICATIONS, AND COMMITMENTS 30

8.1 BBN 30
8.2 Johns Hopkins University 32
8.3 University of Maryland 32
8.4 Cambridge University 33
8.5 Brno University of Technology 34
8.6 Time Commitments 35
8.7 Selected Bibliographies 37
8.7.1 BBN 37
8.7.2 Johns Hopkins University 39
8.7.3 University of Maryland 41
8.7.4 Cambridge University 42
8.7.5 Brno University of Technology 43

9 PROJECT MANAGEMENT AND INTERACTION PLAN 45

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

9.1 PROJECT MANAGEMENT 45
9.2 TEAM SIZE AND COMPOSITION 46
9.3 WORKING/MEETING MODELS 47
9.4 SOFTWARE/CODE MANAGEMENT 48
9.4.1 Software Integration/Management 48
9.4.2 Secure Processing 49
9.5 UNIVERSITY PARTICIPATION 49
9.6 GOVERNMENT'S ROLE 50
9.7 PROJECT MANAGEMENT AT BBN 50
10 COST SUMMARIES 51
11 ORGANIZATIONAL CONFLICT OF INTEREST AFFIRMATIONS AND DISCLOSURE 56
12 HUMAN USE 57
13 ANIMAL USE 58
14 STATEMENT OF UNIQUE CAPABILITY PROVIDED BY GOVERNMENT OR GOVERNMENT-FUNDED TEAM MEMBER 59
15 GOVERNMENT OR GOVERNMENT-FUNDED TEAM MEMBER ELIGIBILITY 60
APPENDIX A: REFERENCES 61
APPENDIX B: EFFECT OF TARGET SIZE 66

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

RATS PATROL

Volume 1 – Technical Proposal
Section II – Technical Details

Technical Area 1: Algorithmic Development and Signal Processing

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

1 Innovative Claims

The Challenge

The ability of the US Government to monitor foreign voice communication channels, in support of intelligence and military operations, is hampered by the lack of a technology that can automatically determine what is important to listen to in the presence of highly degraded and noisy voice signals. Developing such a technology is the aim of the RATS program, and BBN's response is PATROL, a system that will revolutionize US ability to pinpoint speech activity, identify the language and the speaker, and detect keywords of interest, under very noisy and degraded conditions.

Barriers to Progress

Two major barriers to dealing with noisy speech have been fragmentation of research efforts and lack of a realistic corpus on which to perform research. The RATS program promises to shatter these barriers by attacking simultaneously the full spectrum of speech detection problems and by providing the first large realistic corpus for research. BBN has responded to the challenge by forming a team whose members have distinguished themselves internationally in their areas of expertise and who collectively have the breadth and depth of experience, history of innovation, and fresh new ideas to meet the RATS challenge in all areas of this effort.

The Approach

Our approach derives from a holistic view of the problem at hand as a single overarching challenge of speech detection/recognition in noise. The image we have of noisy speech is that the speech is being held hostage by noise. Our fundamental approach takes advantage of the *predictable* properties of the speech to rescue it from the clutches of *unpredictable* noise. PATROL will accomplish the rescue through a *unified* multi-pronged attack that is common to all four RATS application areas: speech activity detection (SAD), language ID (LID), speaker ID (SID), and keyword spotting (KWS). Figure 1-1 shows a conceptual architecture of the PATROL system which combines elements of training and recognition in a single notional diagram. The double arrows indicate multiple analyses and outputs. The terms "data-guided" and "task-specific" will then be specific to the available training and the application area. We now describe each component briefly.

1. **Noise-Robust Feature Extraction:** We propose the use of auditory-based representations of sound that have been demonstrated to separate speech modulations from the noise in a higher-dimensional space, which should make it possible to focus our processing on the speech modulations. These representations include Hermansky's *modulation spectrum* and Shamma's 4-D *cortical spectrum*. (See also item 4 below.)

2. **Data-Guided Feature Projection:** To make the higher-dimensional features usable, we propose to project them onto a lower-dimensional subspace – with little loss in separability of speech from noise – by using non-linear, trainable, discriminative

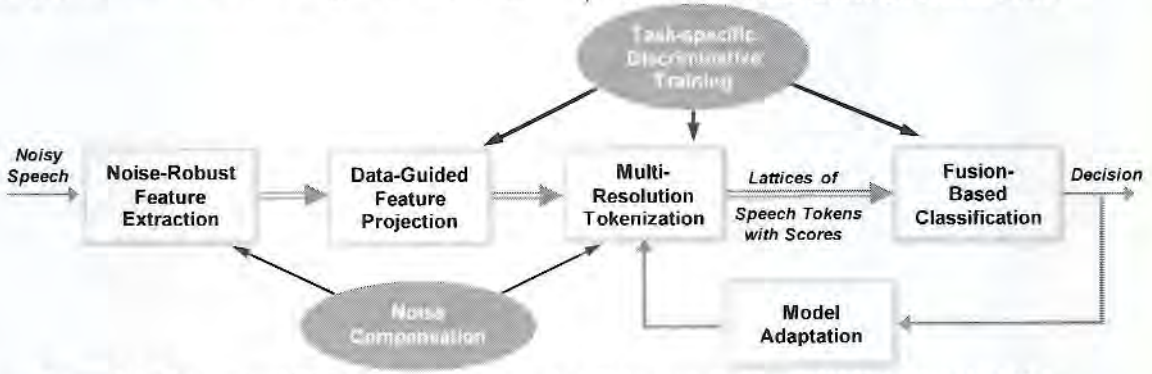


Figure 1-1 Conceptual PATROL architecture for all four applications: SAD, LID, SID, KWS.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

projections based on multi-layer perceptrons and region-dependent transforms. One particularly useful projection that has shown very good performance in noise is the estimation of phoneme posterior probabilities as features. In all cases, the training is guided by the available data and the application.

3. **Multi-resolution Tokenization:** More than anything, what characterizes speech – and differentiates it from noise – is the special manner in which speech units (tokens) vary over time. We propose to model speech as a time sequence of tokens at different time resolutions – frames, phoneme-like *self-organized units*, phonemes, and words. To model the time sequence of these tokens, we shall employ language models that are appropriate for the tokens and the application. The resulting output here is multiple lattices of tokens, with associated likelihood ratios (scores), to be used for later decision making (item 5 below). We also introduce methods that do not require extensive transcription.

4. **Noise Compensation:** Noise compensation methods improve recognition accuracy significantly. Two types of noise compensation will be employed. One, at the feature level, performs discriminative training of the features to minimize the effect of the noise. Another, model-based, is performed in conjunction with tokenization. These methods, which use discriminative adaptive training, include Vector Taylor Series compensation, Joint Uncertainty Decoding, and Predictive Constrained Maximum Likelihood Linear Regression. Joint Factor Analysis, which has been very useful in modeling speaker and channel variability, will be used for noise compensation as well.

5. **Fusion-based Classification:** The basic concept here is to fuse the outputs of multiple processes to reduce uncertainty in the presence of noise. Another is the fusion of multiple lattices of tokens and associated scores, at different resolutions, to formulate an integrated, optimal decision. Different methods of fusion will be considered. One is based on the Dempster-Shafer theory of evidence. Another novel method is a condition-adaptive fusion strategy, called *latent generalized linear model*, in which combination weights of evidence are functions of measured features of the input, such as SNR, type of channel, and speaking rate. Another aspect of fusion is a principled approach to determining an optimal threshold for acceptance/rejection.

6. **Model Adaptation:** The high-confidence fusion decisions can then be used to adapt the speech token models (item 3 above) in a feedback configuration to improve accuracy.

7. **Task-specific Discriminative Training:** Permeating the above components is the use of discriminative training, whereby the various components are trained specifically for each application to optimize performance.

Basis of Confidence

While some of the above ideas have been used individually to combat noise, what we propose is an integrated approach that combines *all* the components in a novel and unified manner for optimal performance. In Section 3, we present our novel ideas in how to optimize performance for each of the application areas (SAD, LID, SID, KWS), but the centerpiece of this proposal is the unified system conception outlined above.

The experienced, world-class team we have assembled; the large number of fresh ideas and unifying new methodology in this proposal; and BBN's track record in fielding operational systems, present a uniquely compelling and credible opportunity for achieving the ambitious performance goals of the program. We are confident that we will deliver to the Government deployable technology that will combat the endemic noise problem, rendering our monitoring operations far more effective.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

2 Proposal Roadmap

AREA	Page References
a. Main Goals of the Proposed Research: The work in this project has two main goals: (1) The first goal is to develop the technology to perform the following four speech detection tasks, for a number of foreign languages, in the presence of highly degraded and noisy voice signals, with high accuracy: Speech activity detection (SAD), language identification (LID), speaker identification (SID), and keyword spotting (KWS). The miss and false alarm target rates for the four applications are given in Table 3-1. The work will include the delivery of an evaluation system to the Evaluation Team in order to assess system performance against the targets for each phase of the project. (2) The second goal is to build and deliver to the Government a system that is military user-friendly, with an intuitive, easy-to-use graphical user interface (GUI), and that will be field-trainable for new languages, speakers, and keywords.	1 (§1) 14-15 (§3.4, 3.5)
b. Tangible Benefits to End Users: By providing the Government with the capabilities described above, with high accuracy, we would increase the ability of the Government to monitor foreign communications significantly. By making it possible for the analysts to focus only on the languages, speakers, and keywords of interest, the productivity of the analysts would increase tremendously.	1-2 (§1)
c. Critical Technical Barriers: (1) While much work has gone into the four RATS application areas, much of that work has been on relatively clean speech. As a result, the techniques that have been developed do not perform well under high noise conditions. (2) For cases where research was done with noisy signals, typically the research focused on one technique at a time – rarely combining multiple techniques. (3) The research is often done using data with artificially added noise, rather than with naturally occurring noisy speech. Any naturally occurring noisy speech is usually collected in one type of environment, e.g., in cars, and the tasks are usually very narrow, such as digit recognition. The RATS program is well poised to overcome these barriers by attacking the full spectrum of speech detection problems and by providing the first large realistic corpus for research.	6 (§3.1) 16-17 (§4)
d. Main Elements of the Proposed Technical Approach: We propose to build a system, PATROL, which will achieve the goals of the program through a *unified* approach to detection systems that are robust to noise. Instead of the piecemeal, one-technique-at-a-time approach of many past efforts, many techniques at multiple levels of processing will be developed and optimized jointly to maximize performance for each of the RATS application areas. The primary areas of work include: (1) **Noise-robust feature extraction:** It has been shown that, when noisy speech	1-2 (§1) 6-7 (§3.2.1)

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

is analyzed using multi-dimensional (multi-D), auditory-based representations, the modulation spectrum of speech typically occupies a different part of the space than the noise. We plan to use several such representations jointly, thus leading to increased robustness, including the multi-D MRASTA modulation spectrum of Hermansky and the cortical representation of Shamma (both on the PATROL team).	
(2) **Data-guided trainable feature projection:** We propose to use nonlinear discriminative projections based on multi-layer perceptrons and region-dependent transforms (RDT) to reduce the dimensionality of these high-dimensional features with minimal loss of information. The term "data-guided" means that these projections will be optimized to preserve information needed for each application (SAD, LID, SID, or KWS).	7-8 (§3.2.2)
(3) **Multi-resolution tokenization:** Speech can be analyzed as tokens (units) of different length, e.g., frames, phonemes, sub-words, or words. In contrast to current methods, we propose to use all levels of tokenization, which will increase the robustness of all applications. We also propose two methods that have been demonstrated to work well with untranscribed speech (which is abundant in an operational setting): *self-organized units* (SOUs) and semi-supervised training.	8-9 (§3.2.3)
(4) **Noise compensation:** Among *feature-based* methods, we propose to filter the multi-D features to enhance the speech and suppress the noise, and to explore discriminative RDT to map from noisy features to clean features. Among *model-based* methods, we will continue to explore sophisticated techniques pioneered by Cambridge University (e.g., Joint Uncertainty Decoding and Predictive Constrained Maximum Likelihood Linear Regression), and Joint Factor Analysis – which has been used effectively at Brno University of Technology to deal with inter-session variability.	9-10 (§3.2.4)
(5) **Fusion-based classification:** The various types of analysis and tokenization will result in a number of lattices of tokens with scores, which will then be combined using *condition adaptive fusion*. For SID and KWS, decisions will be optimized for *sets* of targets rather than for individual targets.	10 (§3.2.5)
(6) **Model adaptation:** The high-confidence fusion decisions can be used to adapt the speech token models (item 3 above) in a feedback configuration to improve detection accuracy.	10 (§3.2.5)
(7) **Task-specific discriminative training:** Permeating the above components is the use of discriminative training, whereby the various components are trained specifically for each application to optimize performance.	6-14 (§3.2, 3.3)
(8) **Speech activity detection (SAD):** In addition to using noise compensation methods, we expect SAD to improve through the use of phoneme and SOU HMMs, as well as via an integrated factorial HMM speech/noise recognizer.	11 (§3.3.1)
(9) **Language identification (LID):** LID will be more accurate with the use of more detailed SOUs, phoneme, and word models to compute acoustic likelihood ratios, and with the combination of scores across multiple levels of tokenization. Speaker Adaptive Training will reduce the speaker variation effect on LID.	12-13 (§3.3.3)
(10) **Speaker identification (SID):** SID will also improve through the use of all	12 (§3.3.2)

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

levels of tokenization for acoustic likelihood ratios and sequence probabilities. These scores will be fused with other classifiers such as SVMs that use super vector representations (e.g., MLLR parameters or mean vectors) of each speaker.	
(11) **Keyword spotting (KWS):** Our KWS system will be based on large-vocabulary speech recognition to model both keywords and other speech. We propose to use higher-level language information, such as topic-dependent language models, and unsupervised adaptation based on high-confidence hits to optimize performance.	13-14 (§3.3.4)
(12) **Software integration and deliveries:** BBN will use its vast experience in integrating software from other sites and delivering real-time, operational systems to the Government to build and deliver two systems: an *evaluation system* to be delivered to the Evaluation Team to assess system performance against program targets, and a *trainable system* as described in (a) above. The evaluation system will be delivered six weeks prior to the end of each phase, and the trainable system will be delivered in Phases 2 and 3.	14-15 (§3.5)
(13) **Speech transcription option:** We have found that KWS users prefer to have a full transcription – even if errorful – to avoid having to listen to every possible hit. In order to fulfill the promise of the 'T' in 'RATS', we have included work on full transcription as an option in Phases 2 and 3.	15 (§3.6)
e. Basis of Confidence:	
In short, two things: The approach and the team. Noise is the main problem and our whole approach is based on various principled ways of dealing with it. We have great confidence that the unified approach presented here and the specific set of novel techniques that we plan to use will, in combination, deal very effectively with the problem of noise in speech.	1-2 (§1) 6-15 (§3)
To perform this challenging work, we will employ the combined talents of an experienced team with a strong history of achievement in every aspect of the program. The existing strong relationships among members of the team will enhance our ability to perform. BBN's experience in executing multi-site speech and language projects – always with superior results – and our proven record of integrating software and deploying real-time, operational systems for the Government, gives us additional confidence that we will be able to provide the analysts with what they need to be more effective in monitoring communications.	30-35 (§8.1-8.5) 45-50 (§9)
f. Nature and Description of End Results to be Delivered to DARPA:	14-15 (§3.5)
Six weeks before the end of each Phase, BBN will deliver an *evaluation system* to the Evaluation Team. A *trainable system* will also be delivered to the Government during Phases 2 and 3.	18-19, 22-23, 25-26 (§5)
The PATROL team will also engage in the writing of technical papers and conference presentations which will detail the approaches being taken, the novel ideas that have been explored, and the results of carefully controlled experiments to assess the benefits of competing methods.	29 (§7)
g. Cost and Schedule of the Proposed Effort:	
Base Phase 1: $5,293,372; Base Phase 2: $3,895,062; Base Phase 3: $3,894,111; Option Phase 2: $482,785; Option Phase 3: $499,559.	51-55 (§10)

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

3 Technical Approach

3.1 Introduction

Much of the research in the four applications in the RATS program (SAD, LID, SID, KWS), has been done either under relatively clean speech conditions, or by artificially adding different types of stationary noise to the speech. In addition, existing research typically applies a single technique at a time to deal with the noise, rarely taking advantage of multiple methods to improve performance.

The reality for conditions of interest to the Government is quite different. The noise is typically neither simply additive nor is it stationary, and can be quite severe. Under such conditions, the current technology is inadequate for the applications of interest.

We are fortunate that the RATS program will be providing us, for the first time, the types of realistic communications noisy speech data on which to perform the research. The program has also given us the opportunity to work on the four detection problems with such data simultaneously, and to partner with other experts to solve this very challenging problem.

Our proposed PATROL system, depicted in Figure 1-1, will employ multiple noise-robust solutions across all four RATS applications in a unified way: features that exploit the distinctive spectral and temporal modulation patterns of speech and noise; acoustic and language models that capture contextual information; adaptive noise compensation in both feature and model space; multi-stream information fusion; and model adaptation. PATROL will use these solutions and the available RATS data, in combination with data-guided and discriminative training techniques, to optimize performance for each of the four RATS applications.

Our general techniques for robustness in noise are described in Section 3.2. Section 3.3 presents methods that are specific to each of the four RATS applications. A timeline for meeting the goals of the RATS program is provided in Section 3.4. Software integration and delivery to the Government are discussed in Section 3.5. Finally, in Section 3.6, we propose an option for automatic transcription of noisy speech. The numbered references below are provided in Appendix A of this proposal.

3.2 Achieving Robustness in Noise

Our approach to achieving robustness in noise is presented below under five headings: 1) noise-robust feature extraction; 2) data-guided feature projection; 3) multi-resolution tokenization; (4) noise compensation; and 5) fusion-based classification (including model adaptation). In Section 3.3, we show how these techniques apply to each of the four RATS application areas.

3.2.1 Noise-robust Feature Extraction

Our attack on noise starts at the front-end, feature extraction. MFCC [1], PLP [2], and RASTA [3] features, along with their time derivatives, the use of cepstral mean subtraction and/or spectral noise subtraction [4], form an initial line of attack against which we will compare more sophisticated noise-robust methods. Prosodic features (pitch envelope, probability of voicing) have also been found useful in LID and SID.

All successful feature extraction techniques are based on some aspect of our knowledge of human auditory perception. In this effort, we plan to explore a series of progressively more sophisticated uses of that knowledge. We start with Hermansky's Frequency Domain Perceptual Linear Prediction (FDPLP) [5], a "plug-in compatible" alternative to PLP that has been shown to provide more noise robustness by focusing attention on the high SNR regions of the signal.

Auditory experiments over the last decade suggest that humans perceive sound in a *multi-dimensional* (multi-D) space consisting of responses of neurons that operate on the

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

spectrogram of the sound and selectively extract its important spectral and temporal trajectories, or *modulations*. It appears that, in this multi-D space, speech and noise typically occupy different regions of the space, due to their different modulation patterns. One example of these multi-D representations is the MRASTA *modulation spectrum* from Hermansky [6], which extracts, at each frequency and time, vectors of information over longer time and frequency spans.

A more sophisticated multi-D processing is the *cortical representation* of Shamma and group [7], which analyzes the spectro-temporal modulations of the auditory spectrogram by applying a 2-D affine wavelet transform on the spectrogram, with spectral dilations over a range of bandwidths (or *scales*) of 0.25-8 cycles/octave and time dilation *rates* of 1-32 Hz. The output is a set of scale vs. rate matrices – the cortical spectrum (CS) – at *each* point in frequency and time. The process is applied at every 10 ms of audio, yielding a time series of 3-D (frequency/rate/scale) tensors.

To illustrate the potential of the CS in terms of robustness to noise, consider the speech-in-jet-noise example in Figure 3-1. Shown on the top left are the CS for clean speech and jet noise, *averaged* for all time and frequency points over a 1-second span of the signal. This CS clearly shows how jet noise largely occupies a different part of the CS space than does speech, while in the speech-in-jet-noise ordinary spectrogram below, the two spectra overlap. The figure then shows the CS of an enhancement or correction filter which, when applied to the speech-in-jet-noise CS, yields a cleaned-speech CS that resembles the original CS of the clean speech [8]. The resulting CS (now *without* averaging over frequency and time) will form our multi-D feature representation for further processing.

3.2.2 Data-guided Trainable Feature Projection

The above feature extraction methods produce a multi-D feature vector (~7,000 dimensions) at every 10 ms of speech! So, modeling the resulting features efficiently and reliably requires projecting them to a sub-space of significantly lower dimensionality, without losing the speech/noise separability property. Standard linear dimensionality reduction techniques, such as principal component

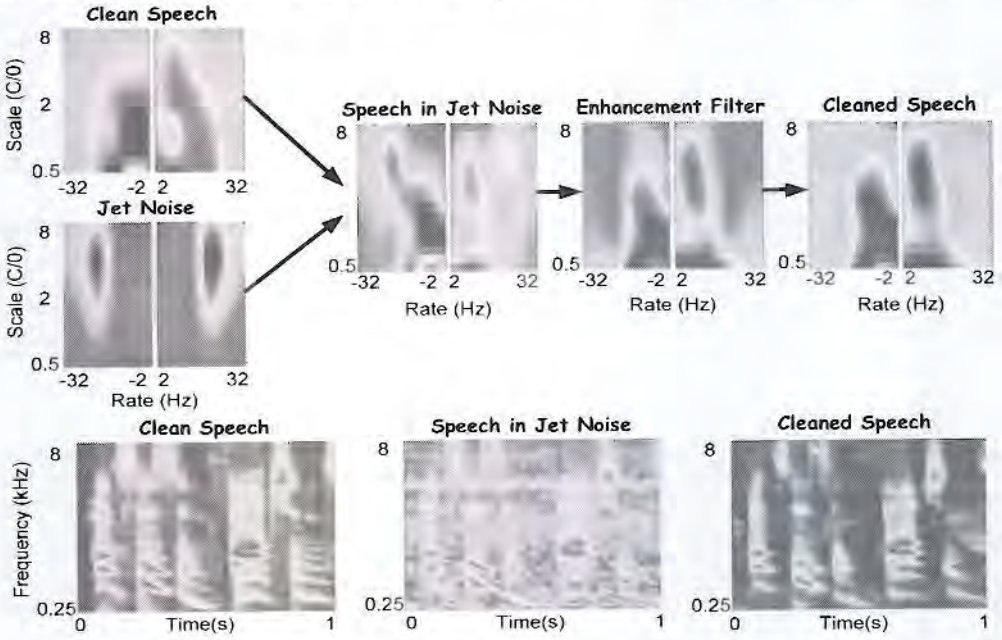


Figure 3-1: Example of noise compensation in cortical representation (top row). The bottom row shows the corresponding auditory spectrogram.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

analysis (PCA) or linear discriminant analysis (LDA), can be useful but are not powerful enough for the task at hand. We propose to investigate *non-linear discriminative* projections based on multi-layer perceptrons (MLP), and the region-dependent transforms (RDT) [9] developed at BBN.

Posterior-based MLPs map the input features to vectors of phoneme posterior probabilities. The MLP parameters are estimated on a phoneme-labeled corpus to minimize frame-level phoneme classification error. This method has been shown to yield significant increase in recognition accuracy, especially under noisy conditions [10] [11], with further improvements possible by training the MLPs in a hierarchical manner [12][13].

The RDT partitions the input feature space into regions, each with its own projection. Given an input feature vector, RDT computes the probability of each region using a partition model, which can be customized for different types of parameters. For example, a region could correspond to a single Gaussian in a Gaussian Mixture Model (GMM), or a phoneme, or even a state in a hidden Markov model (HMM). When the estimation of the transforms in RDT is performed discriminatively, the resulting projections have been shown to give significant speech recognition accuracy improvements [14].

Both MLPs and RDT are *trainable, data-guided* feature projections. Given an annotated training corpus and the associated raw features, they can learn to find the feature dimensions that make the desired distinctions that are least affected by noise. Here we refer to distinctions that are most important for the task of interest, such as speech/nonspeech in SAD, speakers in SID, and languages in LID. In KWS, MLP and RDT transforms can be trained to minimize frame phoneme classification error or expected phoneme error rate, respectively. Of the two types of discriminative non-linear projections, RDT may have the larger potential for success in this project, due to its ability to accommodate complex classifiers, such as HMMs.

3.2.3 Multi-resolution Tokenization

Tokenization is the process of converting the speech signal into a sequence of tokens, and *multi-resolution* refers to the tokens being of different time resolutions (e.g., frames, phonemes, words). We propose to model the special variations of these tokens over time to re-capture the speech from the noise. Multi-resolution models will be estimated for each of the four RATS applications.

Modeling Sequences over Time

To break the noise barrier in SAD , we must go beyond the use of time-agnostic GMMs and use the fact that noisy speech still consists of a time sequence of sounds, at a predictable rate of about 6-10 phonemes per second. We propose to discriminate speech from non-speech by modeling the sequence of speech tokens in time explicitly. For SAD to be language-independent, we would limit the resolution of the tokens up to the phone level.

LID and SID systems already model speech sequences using a tokenizer (typically a generic phoneme recognizer), and they model the likelihood of those sequences using an n-gram model. Here, we also propose to use tokenizers to model the *acoustics* for the target language or speaker, so that differences in pronunciation can be captured automatically. The acoustic likelihoods will be measured for tokens at different time resolutions. Each type of token has its advantages and disadvantages, and the choice of tokenization will be influenced by the amount of training data and the type of data annotations available. For example, we may not always have substantial amounts of transcribed training speech for training word-based recognizers. We propose two novel solutions for the case where transcriptions are not available: semi-supervised training and Self-Organized Units (SOUs).

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Semi-Supervised Training

We have shown that, given large amounts of un-transcribed audio, in conjunction with a small amount of transcribed speech (e.g., one hour), we can achieve speech recognition accuracy that approaches that of having manual transcripts for the larger corpus [15][16][17]. Starting with a model estimated from the small amount of transcribed speech, the method uses iterative re-estimation of the model based on recognizing the untranscribed audio, coupled with appropriate confidence filtering. Recently, the method has been shown to work on noisy speech data as well. We propose to use semi-supervised training with the RATS data, as well as any other untranscribed audio data we can find in the five languages of interest, to improve tokenization accuracy. Note that this method can gain special importance in an operational setting, where essentially unbounded amounts of untranscribed data are typically available.

Self-Organized Units (SOUs)

We have developed another technique that can take advantage of untranscribed audio when we have *no* transcribed speech available, or even a vocabulary for the language [18]. The method starts by dividing the audio into phoneme-like units using a simple spectral change criterion. These units are clustered to produce initial pseudo-phone models, called *self-organized units* (SOUs), which are then re-estimated using context-dependent HMMs. We can use SOUs to tokenize both training and test data. For SAD, SOUs can be used to model both speech and nonspeech segments. An acoustic model and a language model of SOU sequences for different languages can be used in LID. For SID, we can find the sequences of SOUs that are frequently produced by each speaker.

3.2.4 Noise Compensation

We group noise compensation techniques into two categories: (a) *feature-based* methods, which process acoustic features in some fashion to suppress the noise, and (b) *model-based* methods that alter the parameters of the acoustic model based on some estimate of any residual noise in the features.

Feature-based Noise Compensation

Since the modulation and cortical spectra, presented in Section 3.2.1, attempt to separate the speech from the noise in multi-D, we propose to introduce filters in those domains to compute new features that enhance the speech and suppress the noise. Shamma has already demonstrated the utility of this approach in SAD [19]. We propose to generalize the approach to other applications.

We will also explore *RDT-based enhancement*, a new method that relies on discriminatively estimated region-dependent transforms (see Section 3.2.2) to perform a mapping from noisy features to clean features. Unlike other feature mapping techniques, such as SPLICE [20], which require stereo (noise/clean) speech samples for training, our method only requires a noisy training corpus, and a target acoustic model of clean speech. The clean speech model may be obtained from any existing clean speech corpus, not necessarily of the same spoken language.

Model-based Noise Compensation

Sophisticated model-based compensation techniques, such as those pioneered by our partners at Cambridge University [21][22], have yielded excellent recognition results, even in low SNR. These approaches can also be used within an adaptive training environment allowing the estimation of a neutral, canonical, speech model that is suited for adaptation to a range of target environments. Specifically, this will include improving existing work on discriminative adaptive training approaches based on schemes such as Vector Taylor Series compensation [23], Joint Uncertainty Decoding (JUD) [24], and Predictive Constrained Maximum Likelihood Linear Regression (PCMLLR) [25]. Furthermore,

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

we will investigate methods that combine model-based noise compensation with unsupervised speaker adaptation, such as Noisy CMLLR adaptation [26].

Another technique that has been used extensively in recent SID and LID systems to deal with inter-session variability is joint factor analysis (JFA) [28]. This method models the effect of speaker and channel variability on the parameters of the acoustic model used in such applications, typically a GMM. JFA can be augmented to address variability due to noise as well, and can be applied to more complex acoustic models, such as HMMs, by generalizing the form of the transformations on the model parameters. Our partners at Brno University of Technology (BUT) will explore these research directions.

3.2.5 Fusion-based Classification

Our proposed system will use multiple feature analyses and models, as described in the previous sections, whose decisions – in the form of lattices of speech tokens with scores – will be ultimately combined, or *fused*, in order to generate the final system output for each of the four RATS applications.

A step in this direction is fusion based on the Dempster-Shafer theory of evidence [29]. This strategy suppresses streams with high entropy of the classifier output, and enhances low-entropy streams. Our partners at Johns Hopkins University (JHU) have applied this type of fusion on the phonetic posteriors from multiple feature streams [30], and are currently looking into alternative confidence measures for combining the independent stream decisions.

An alternative method for combining different classifiers is by directly optimizing the fusion model on a held-out set. We propose to extend this approach using a condition-adaptive fusion strategy, called *latent generalized linear model* (LGLM), in which the combination weights are functions of measured features of the input, such as SNR, channel type, and speaking rate. The LGLM parameters can be estimated to explicitly optimize performance in each application.

Optimal Decisions and Target Set Size

In 1986, while working on a BBN-funded speaker verification project [31], we found it best to use a different threshold for each target speaker, or conversely, normalize the scores differently for each target so that a single threshold would result in the same false alarm (FA) rate for all targets. This was done by computing the distribution of scores for a large number of impostor samples against each target model. This technique is now commonly used for speaker spotting [32]. We have applied it to topic spotting [33][34], and a modified version more recently to KWS [35]. This method is of special significance because of the importance of controlling the FA rate in RATS.

The evaluation plan for SID and KWS in the RATS BAA calls for testing on a set of N targets at a time (N=10 for SID and N=25 for KWS), instead of the usual one target at a time. For the task of detecting ***sets*** of targets, we will modify the decision criterion to use all of the targets jointly to decide whether to accept or reject a test sample. In principle, if we know this set of targets during model training, we can train all of the models discriminatively to directly estimate the posterior probability of the target set. Alternatively, we can use this information in the fusion classifier, which can estimate directly the posterior probability of the target set given all of the scores from the different models and targets.

Model Adaptation

Fusion decisions will include confidence scores. For each application, we propose to use high-confidence hits to adapt the speech models (tokenizers as well as acoustic scorers) and re-process the test data to improve detection accuracy.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

3.3 Application-Specific Methods

Here we summarize the techniques we will use for each of the four RATS applications. In all four applications, we shall assume that, at a minimum, we start with the noise-robust feature extraction techniques in Section 3.2.1, at increasing levels of complexity, up to the modulation and cortical spectra, which promise to segregate the speech and the noise in a multi-D space, followed by the feature-based noise compensation techniques in Section 3.2.4. Then, when necessary, we will use the data-guided feature projection techniques in Section 3.2.2 to reduce the feature dimensionality. This step will be performed discriminatively, based on the available training data and the application. The result will be what we will call below the *processed features*, or simply features. The types of tokenization (Section 3.2.3) to be used will depend on the application.

3.3.1 Speech Activity Detection (SAD)

The processed features will be used to discriminatively train both static frame-based classifiers (e.g., GMM, or Support Vector Machines – SVM), as well as phoneme-like HMM tokenizers. Initially, we will explore HMM tokenizers configured as we have used for SAD in broadcast news, where speech is often corrupted by background noise or music. In that work [36], we used an HMM consisting of phoneme-class tokens (e.g., nasal, fricative, vowel), as well as non-speech tokens (e.g., music, laughter, background noise). We will also investigate SOU-based HMM tokenizers, with one set of SOUs for noisy speech and another set for pure noise. A simplified version of this tokenizer, using just two SOUs (speech/non-speech), was recently applied with success to air traffic control audio data [37]. Both of these types of HMM tokenizers are expected to provide a good baseline for SAD on RATS data.

We also propose a bold novel approach that attempts to model the very process of generating noisy speech through the composition of two *hidden* stochastic processes: a clean speech process and a *noise-transform* process. The new model takes into account the fact that noise is unpredictable and differs between training and test. By factoring the noisy signal into these two processes, the proposed model will be able to better generalize to unseen combinations of speech and noise. The proposed factorial HMM [27] tokenizer is depicted in Figure 3-2.

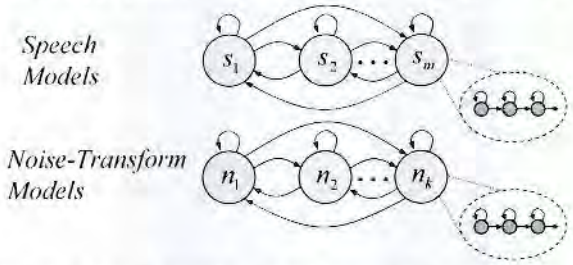


Figure 3-2: A factorial HMM recognizer proposed for SAD.

The factorial HMM recognizer consists of two parallel ergodic HMMs: an ordinary phoneme-like speech HMM (that also includes a special pause phoneme) and a novel noise-transform (or *noise-x*) HMM. These are two-level HMMs, with each state of the top level comprising a second-level HMM. The speech model is estimated from any existing corpus of clean speech data (the corpus can be in any language). The probability density functions (pdf) of the noise-x HMM are distributions over *transformations* that transform the speech model parameters into those of the noisy speech. (There will be a few noise-x HMMs, one for each expected type of noise.) Given the clean-speech HMMs, the parameters of the noise-x pdfs are estimated on RATS data. Since the noise-x transforms are state specific, the model can handle non-stationary noise. Note that the speech models are quite generic here and are for rather broad classes of sounds. The action is really in the noise-x HMMs which are trained to transform the speech models into the noisy speech models from the

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

RATS data. SAD is performed by finding the highest scoring sequence of factorial HMM states (pairs of noise-x and phoneme/pause states) given the noisy observations.

The above SAD methods will be combined or fused with that developed at the University of Maryland (UMD) using adaptive fusion (see Section 3.2.5) to optimize performance.

3.3.2 Speaker Identification (SID)

For accurate speaker identification, it is often necessary to not only determine whom the speaker sounds like (vocal tract characteristics, speaking rate, pitch contour, accent), but also whether certain pause fillers or distinguishing words or phrases are used. Relying on such *high-level information* becomes even more critical when the acoustic signal is degraded by noise or channel distortions.

Current state-of-the-art SID systems use simple models (e.g. GMMs) for modeling speaker vocal tract characteristics, and, in some cases, more complex automated speech recognition (ASR) models for deriving some of the higher-level information [38] [39]. The ASR models are trained in a typical, speaker-independent, channel-independent fashion, and are adapted to the speaker during testing using standard methods such as maximum likelihood linear regression (MLLR) [40]. However, the adaptation is estimated to optimize an ASR criterion, *not* to help in speaker identification. We believe there is a better way. Here we propose to *couple* speaker and speech recognition in a novel SID approach we call *speaker-discriminative ASR*, or SD-ASR. This approach also promises to perform especially well in noise.

Under SD-ASR, each speaker is modeled by a full ASR sub-system, with speaker-specific acoustic models derived via adaptation from a universal background acoustic model. What distinguishes SD-ASR from typical ASR is the fact that the parameters of the speaker-specific and background ASR models are estimated *jointly* so as to directly optimize SID performance on training data. The training criterion is closely related to MMIE (maximum mutual information estimation) as developed for ASR, but it aims at maximizing the posterior probability of the correct speaker over the training data, rather than maximizing the posterior of the correct word sequence. The criterion can also be easily modified to optimize parameters for the scenario in which the target is not a single speaker, but rather a set of high value individuals (HVIs).

The proposed SD-ASR system can be used even in cases where little or no transcriptions are available, by applying semi-supervised training, or using self-organizing units as explained in Section 3.2.4. By combining multiple SD-ASR systems – each trained with different tokenizations (words, sub-word units, phonemes, SOUs) –with complementary state-of-the-art SID systems from our Brno University of Technology (BUT) and JHU partners via adaptive fusion, significant improvements in SID performance will be achieved.

3.3.3 Language Identification (LID)

As in SID, accurate LID must rely on multiple types of information extracted from the acoustic signal. But, unlike SID, LID needs to factor out differences due to speakers. For that purpose, standard approaches to LID make use of speaker-independent acoustic models to recognize lattices of phonemes or sub-word units, which are then rescored by multiple language-specific acoustic models and n-gram phonotactic models [41] to find the highest scoring language on a given segment of speech. The language-specific acoustic models aim at capturing distinctive general acoustic variations, such as pronunciation, prosody, or phoneme duration that are characteristic of the language being spoken. However, the acoustic models used in state of the art LID systems are typically GMMs [42], and hence have limited modeling capabilities. To strengthen the role of the

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

acoustic model and improve performance, we propose to introduce the use of highly accurate HMMs, trained discriminatively on the noise-robust features and with the model-based noise compensation approaches described in Section 3.2.4.

To ensure that our LID acoustic models are not confused by speaker differences, we will apply speaker-adaptive training (SAT) [43], which tries to map the training speakers for the language to the same compact feature space. Then, when classifying a new sample, we adapt each of these compact acoustic models to the test speaker before computing the likelihood. This way, the acoustic models will be less sensitive to speaker differences within a language.

One issue with n-gram phonotactic models is that they rely on local context and hence might be less reliable under noisy conditions due to inaccuracies in recognizing the correct sequence of speech tokens. We propose to alleviate this problem by applying information retrieval techniques, as in [44], which attempt to model the global distribution of token n-grams.

The final LID system output will be based on adaptively fusing the decisions of multiple classifiers (GMMs, HMMs, enhanced phonotactic models) applied to multiple tokenizations of the speech input, and incorporating the LID system from BUT.

3.3.4 Keyword Spotting (KWS)

In order to perform KWS, one must have a good model of the non-keywords as well as the keywords. It is now acknowledged that the best model for the non-keywords is a set of models for those words – both acoustic and language models (LM). This is especially true in the presence of noise, where making use of context gives us a better chance of detecting the keywords. Another benefit of having good models for the non-keywords is that they can then all serve as possible keywords, since it would not be possible to know in advance what the keywords might be in the field (the case of unknown words is discussed below). BBN has developed state-of-the-art systems in both large-vocabulary speech recognition and KWS. Here, we propose extensions to that work.

One way of making better use of context, and to improve KWS performance in noise, is to use a topic-weighted LM. Since topics are typically not annotated in data, we propose to use our Unsupervised Topic Discovery (UTD) method [45] to discover a large set of topics for the language and the domain in an unsupervised manner.

Given the set of topics, we estimate from training the probability of each keyword occurring in a passage given that the passage contains each topic. Then, given a test passage, we first recognize it automatically, and then determine the likely set of topics contained in the passage. Using the topic probabilities, we compute modified prior probabilities for each keyword and then scale all the n-gram probabilities for this keyword to produce new priors. Finally, we perform a second pass of recognition – or just rescore the lattice of possible answers with this modified language model – in order to obtain an improved result.

Another method to improve KWS performance is to combine or fuse decisions from multiple independent KWS processes. Specifically, we will have a system that first runs recognition to generate a lattice of tokens (phonemes, sub-words, words), and then rescores the lattice with different acoustic models, based on HMMs or word-specific SVMs. The final adaptive fusion stage would take in the scores from the HMM and SVM classifiers, along with any additional measured features of the input in order to produce the final decision.

Handling Keywords Outside the Training

So far we have assumed that the keywords are words that are already in the system's

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

vocabulary. The case of new words, unknown to the system, must be handled properly in an operational setting, including new names. We assume here that the words are provided to the system in the form of text. For the languages under consideration in the RATS program, having the spelling of a word is sufficient for the system to deduce a pronunciation for the word and, therefore, provide a reasonable initial acoustic model, coupled with an initial back-off LM.

The new keywords will have lower recognition accuracy at first. We propose to adapt the models for each keyword based on the high-confidence hits for that keyword. This includes the parameters of the context-dependent HMMs, the pronunciation probabilities, and the LM for these words. To improve the LM, we find those other words that often occur in the same observed context as these high-confidence tokens. Then, we replace the initial back-off LM probabilities for the keyword with a weighted interpolation of all of the LM probabilities relevant to those similar words, in proportion to their similarity and the confidence of the first recognition pass. Finally, with a complete set of new models for these keywords, we perform a second pass of recognition over the data set in order to obtain more accurate estimates of the locations of the new keywords.

3.4 Achieving the Program Goals

Table 3-1 shows our proposed performance timeline for the RATS effort over the three phases of the project. The baseline numbers for SAD, LID, and KWS were obtained in a DARPA-funded joint effort between BBN and MIT Lincoln Lab (MITLL) to assess the state of the art when applied to government-supplied, noisy communication speech data in four languages: Iraqi Arabic, Farsi, Dari, and Pashto. The baseline numbers for SID were supplied by MITLL from a previous study with noisy speech data. The Phase 3 targets are those specified in the BAA and the interim targets were computed through logarithmic interpolation between the baseline and the Phase 3 targets (this results in equal percentage improvements from one phase to the next). After program start, we plan to measure a new baseline using the data to be supplied by the Government. Based on that baseline, we will work with the Program Manager to develop new interim performance targets.

	% P(Miss)	% P(FA)	% TE
SAD Base	5.0	6.0	5.5
Phase 1	4.2	3.3	3.7
Phase 2	3.6	1.8	2.4
Phase 3	3.0	1.0	1.5
LID Base	10.0	15.0	12.0
Phase 1	7.9	6.1	6.9
Phase 2	6.3	2.5	3.5
Phase 3	5.0	1.0	1.7
SID Base	13.0	5.0	7.2
Phase 1	11.9	3.3	5.2
Phase 2	10.9	2.2	3.7
Phase 3	10.0	1.5	2.6
KWS Base	40.0	6.0	10.4
Phase 1	28.8	4.2	7.3
Phase 2	20.8	2.9	5.1
Phase 3	15.0	2.0	3.5

Table 3-1 Proposed phased performance targets. TE is the harmonic mean of P(Miss) and P(FA).

The BAA specifies that, for SID, a hit is the correct identification of any of the 10 HVIs in a set. While the P(miss) for a set is somewhat less than that for a single HVI, the P(FA) for a set grows almost linearly with the number of target speakers. (A similar statement can be made for KWS where a hit is defined as the correct identification of any word in a desired set.) A detailed analysis of this effect is provided in Appendix B. We propose to work with the Program Manager to clarify the intended measure of performance.

3.5 Software Integration and Deliveries

BBN will deliver to the Government two systems. Six weeks prior to the end of each phase, BBN will deliver an *evaluation system* to the Evaluation Team. During each of Phases 2 and 3, BBN will deliver a fully functioning *trainable system* that is military user-friendly, with an intuitive, easy-to-use

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

graphical user interface (GUI), and that will be field-trainable for new languages, speakers, and keywords.

These delivery requirements present challenges that BBN is uniquely equipped to handle. We are highly experienced in the integration of technology from other sites into functioning systems, including trainable systems, with GUIs, *and* delivering real-time, end-to-end systems to Government and other customers, either for internal use or for deployment in the field for 24/7 operations. Section 9.4.1 details our experience in the building and delivery of speech-related operational systems, and our plans in this effort to integrate and test software from different sites, including the specification of appropriate APIs for the delivery of software from other sites to BBN, consistent with the delivery requirements of the Evaluation Team. (Also, Section 9.4.2 summarizes BBN's secure processing capabilities, in case they are needed in this project.)

We are fortunate that our team members have had experience delivering their software for external evaluation and integration purposes. For the evaluation system, we plan to integrate the best algorithms from the team members, as they become available. Currently, we do not expect speed to be an issue for running the system in detection-recognition mode. However, should speed become a concern for any of the software components, we will work with our software developers to find suitable speedups. Again, BBN has had extensive experience in speeding up speech-related software to real-time or faster.

For the trainable system, we will work with the Program Manager and potential users to help specify the interface and training data requirements for each of the four RATS applications. Based on extensive testing, we will then decide which components to include in the trainable system, so that we can deliver the best performing system within the constraints of third-party operation and speed. During Phase 2, we will place in the hands of users an initial trainable system. In order to ensure that the system meets the needs of the users, we will work closely with the users to get their feedback and suggestions for enhancing the system's capabilities and for improving the user interface. Improvements that can be implemented easily will be done immediately and another version will be delivered. Other suggestions, as well as further improvements in the technology, will be incorporated in the system that will be delivered during Phase 3.

3.6 Speech Transcription Option

The 'T' in the acronym RATS stands for 'transcription'. Yet, none of the tasks in the RATS program focuses explicitly on the transcription of speech into a sequence of words. Even though the short-term goals of RATS do not include transcription, since it may be viewed as too difficult a task on which high accuracy could be obtained in the short-term, we have found that KWS users prefer to have a full transcription – even if errorful – to avoid having to listen to every possible hit. Although we plan to use elements of modeling for speech transcription to perform KWS, we believe it is important to work on the full transcription task itself, not only for its own sake, but because it will also lead to better KWS performance. As such, we are proposing, as an option for Phases 2 and 3 of this program, to work explicitly on the transcription task, at least for one language. (The language will most likely be Arabic, but it would be decided in consultation with the Program Manager, should this option be exercised.) BBN and Cambridge University, both highly qualified to work on this task, will join efforts in using the various noise amelioration techniques presented in this proposal to perform research on speech transcription in noise.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

4 Comparison with Current Technology

We review current approaches to noise compensation, SAD, SID, LID, and KWS, and contrast them briefly to the proposed methods (see Section 3). We also provide state-of-the-art performance results under both clean and noisy conditions. Results with communication noise in four languages (Iraqi Arabic, Farsi, Dari, and Pashto) from the DARPA-funded RATS feasibility study by BBN and MIT Lincoln Lab are also given (see baseline numbers in Table 3-1).

4.1 Noise Compensation

Two basic approaches to dealing with noise are: enhancing the features (feature-based) and dynamically adjusting model parameters (model-based). One early feature-based approach is spectral subtraction [46][4] which tries to remove an estimate of the noise spectrum from the noisy speech spectrum. Recent extensions attempt to build statistical mappings from the noisy spectrum to the clean spectrum [47,48,20]. Model-based approaches, such as Vector Taylor Series [23-24], Parallel Model Combination [22], and predictive CMLLR [25], adapt the HMM state observation probabilities based on an estimate of the noise. These methods were mostly developed for testing automatic speech recognition (ASR) under mis-matched conditions, using standard MFCC or PLP front-ends. Researchers have also developed long-span features based on human auditory models [7][11] that are generally more robust to noise, without requiring an explicit noise estimate.

While each of the above methods has been shown to provide significant benefits to ASR in noise, there has not been a systematic effort to apply and optimize these techniques in combination. In Section 3.2 we presented a unified framework for attacking noise across all four RATS applications. This included a number of robust feature extraction and noise compensation techniques, and innovative discriminative training methods for reducing feature dimensionality and for building detailed multi-resolution models of speech, speakers, and languages. In addition, we proposed a novel condition-adaptive strategy for fusing decisions from multiple classifiers, as well as unsupervised confidence-based model adaptation, in order to reduce uncertainty in the presence of noise.

4.2 Speech Activity Detection

In the NIST Rich Transcription evaluation, the best SAD systems [49-51], using a combination of energy- and spectral-based features, and classifiers based on Gaussian Mixture Models (GMM) achieved an error rate (% of total time that was misclassified) of 3.2% when tested on the Meeting and Lecture data. Their performance degraded to 10% when tested on other noisy data. In the RATS feasibility study, the error rate for GMM-based SAD was 11% (P_{miss}=5% at P_{FA}=6%). Recently, in a speech/nonspeech classification (nondetection) task, our partners Shamma and group [19] showed that an SVM classifier trained on multi-D cortical features (Section 3.2.1) achieved an error rate of 1.2% on noisy speech – significantly better than other approaches [52-53].

In addition to using robust features, our proposed methods place greater reliance on detecting speech by virtue of its predictable time varying nature, which can be modeled by the long-term modulation spectrum and also by using longer-duration tokenizations of speech, such as phonemes or self-organized units (SOUs). In Section 3.3.1 we described a novel factorial HMM for SAD that can achieve even greater robustness by modeling the effects of noise explicitly.

4.3 Speaker Identification

In SID, broad speaker characteristics are represented by text-independent statistical models such as GMMs [54-55] that can be discriminatively trained [42]; or SVM

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

classifiers that use speaker adaptation features (such as MLLR matrices [38]) or adapted GMM supervector means [56]. Joint Factor Analysis (JFA) [28][57] and projection approaches, such as Nuisance Attribute Projection (NAP) [58], are used to reduce the effect of inter-session channel variability. Higher level features, such as word usage, pronunciation, and prosodic patterns are used when available [59]. Multiple systems are used with non-linear classifiers, such as logistic regression or neural network, to form a single decision [60-61].

In the 2008 NIST Speaker Recognition Evaluations, the best results in the telephone matched train/test condition, obtained by our partners at Brno University of Technology, were P_{miss}=5% at P_{FA}=5%. In the RATS feasibility study, the baseline result under noise was P_{miss}=13% at P_{FA}=5%.

The main weakness of current SID practices is that the acoustic likelihood ratios are based on simple GMMs, which are very susceptible to noise and cannot model distinguishing long-term speaker attributes. Our proposed approach overcomes this problem by computing acoustic likelihood ratios based on all levels of tokenization, using detailed HMMs trained via novel discriminative training, as outlined in Section 3.3.2.

4.4 Language Identification

LID is commonly performed using techniques similar to SID – such as discriminatively trained GMMs and supervector SVMs – in addition to phonotactic approaches that use n-gram scores from a phoneme recognizer [41][62]. As in SID, these systems are combined to optimize performance. In the 2009 NIST Language Recognition Evaluation, the best results were P_{miss}=1% at P_{FA}=1% for 30 second tests. In the RATS feasibility study, the baseline results under noise were P_{miss}=10% at P_{FA}=15%.

Our proposal for LID (Section 3.3.3) replaces weak GMMs with powerful multi-resolution HMM tokenizers that are trained discriminatively and in a speaker-adaptive manner in order to model and factor out inter-speaker variability. The use of semi-supervised training and SOUs will enable us to produce tokenizers for all 15 of the languages in the RATS program. We also propose more refined phonotactic models that look beyond local n-gram context for broad phonotactic patterns that are characteristic of languages.

4.5 Keyword Spotting

KWS is commonly performed with either phoneme or word recognition. In [63], we demonstrated that modeling context (words next to the keywords) significantly improved performance. In [35], we showed that using lattices in KWS with appropriately set detection thresholds is also important.

In the 2006 NIST Spoken Term Detection evaluation, for the clean conversational telephone speech domain in which all sites participated, the best performing system, submitted by BBN [35] using large vocabulary recognition, achieved P_{miss}=10% at P_{FA}=0.01% [64]. In the RATS feasibility study, the Arabic KWS system was trained with 1 hour of transcribed noisy data. Using other out-of-domain training data and applying adaptation techniques, the KWS error rates for Iraqi Arabic were P_{miss}=40% at P_{FA}=6%.

Our proposed work (Section 3.3.4) includes stronger prior models, such as topic-weighted language models, and robustness through the use of multi-resolution models, with condition-dependent fusion and model adaptation. We proposed several techniques, based on unsupervised adaptation, for improving both the acoustic and language models for keywords that are not well-represented in the training data. Semi-supervised training would allow us to dramatically improve performance when we have very little transcribed noisy audio.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

5 Statement of Work

BBN proposes to build a system, PATROL, that will revolutionize US ability to pinpoint speech activity, identify the language and the speaker, and detect keywords of interest, under very noisy and degraded conditions. The work presented in Section 3 will be accomplished in three phases: an 18-month Phase 1, a 12-month Phase 2, and a 12-month Phase 3.

Deliverables under this effort shall include:

- Quarterly Progress Reports (QPR)
- Final report
- Papers published in international conferences and/or journals during the course of the program
- All software and documentation developed under the project
- Any data collected under the project
- Integrated systems for evaluation (6 weeks before end of each phase)
- Trainable, military user-friendly system during Phases 2 and 3.

In the remainder of this section we present an itemization of the proposed tasks and subtasks, organized by phase, and grouped into one of the following areas: software integration and delivery, feature extraction and projection, noise compensation, speech activity detection (SAD), language identification (LID), speaker identification (SID), and keyword spotting (KWS). In Phases 2 and 3 we also list tasks associated with the proposed RATS Speech Transcription option described in Section 3.6.

For each subtask, we show its duration (***Dur***), objective (***Obj***), the approach taken to accomplish it (***App***), participating sites (***Sites***), completion criteria (***Criteria***), and deliverables (***Del***). We also indicate any dependencies (***Dep***) to other tasks. All acronyms are as defined in Section 3.

1. Phase 1

1.1. Software Integration and Delivery

1.1.1. Develop library for FDLP and modulation spectrum

Dur: 2 months. ***Obj***: Develop portable software library for computing FDLP and modulation spectrum. ***App***: Develop C++ library that computes FDLP and modulation spectrum; write associated regression tests; optimize code for speed. ***Sites***: JHU. ***Criteria***: Code compiling and passing regression tests at all sites. ***Del***: QPR.

1.1.2. Develop library for testing MLPs

Dur: 3 months. ***Obj***: Develop software library that provided functions for testing multi-layer perceptrons (MLPs). ***App***: Develop C++ library functions that can test MLPs on a corpus of feature vectors; write associated regression tests; optimize code for speed. ***Sites***: JHU. ***Criteria***: Library compiles and passes regression tests at all sites. ***Del***: QPR. ***Dep***: 1.1.1.

1.1.3. Develop library for cortical feature extraction

Dur: 2 months. ***Obj***: Develop portable software library for computing cortical features. ***App***: Develop C++ library that computes cortical features; write associated regression tests; optimize code for speed. ***Sites***: UMD. ***Criteria***: Code compiling and passing regression tests at all sites. ***Del***: QPR.

1.1.4. Integrate baseline UMD SAD SVM to PATROL

Dur: 2 months. ***Obj***: Integrate SAD SVM training and testing software to BBN's system. ***App***: Develop C++ library functions that can train and test SAD SVMs on a corpus of feature vectors with associated targets; check source code into BBN's software repository; write associated regression tests; optimize code for speed. ***Sites***: UMD. ***Criteria***: Code working properly at BBN. ***Del***: QPR. ***Dep***: 1.1.3.

1.1.5. Transition baseline SID and LID systems from BUT to BBN

Dur: 6 months. ***Obj***: Integrate BUT SID and LID system components to BBN's system. ***App***: Develop C++ library functions that can test BUT's SID and LID models on existing LDC/NIST corpora; check source code into BBN's software repository; write associated regression tests. ***Sites***: BUT. ***Criteria***: Obtain same results as BUT on NIST SID/LID evaluation sets. ***Del***: QPR.

1.1.6. Integrate PATROL research components into Phase 1 Deliverable system

Dur: 7 months. ***Obj***: Continuously integrate research components into PATROL system. Ensure that PATROL system runs within specified time limits in evaluation mode, on deliverable hardware. ***App***: Iteratively integrate system components to PATROL system as they become updated by research efforts. Test component connections and measure performance characteristics (accuracy, speed). Optimize components as needed in order to meet target speed. ***Sites***: BBN. ***Criteria***: PATROL system achieves expected accuracy and runs within specified time constraints. Task repeats until the start of Phase 1 evaluation. ***Del***: QPR.

1.1.7. Develop GUI for evaluation system

Dur: 5 months. ***Obj***: Develop a graphical user interface (GUI) for the PATROL system, supporting Phase 1 evaluation tasks. ***App***: Implement GUI that allows the user to easily select an evaluation task, open a test file, run the PATROL system on it, and save the output. Write source code that interprets GUI inputs and calls corresponding functions in the PATROL system library. ***Sites***: BBN. ***Criteria***: GUI is intuitive, easy to use and allows user to perform all evaluation tasks. ***Del***: QPR.

1.1.8. Test and prepare P1 PATROL system for Delivery

Dur: 2 months. ***Obj***: Ensure that Phase 1 PATROL system runs with expected accuracy and speed in evaluation mode. ***App***: Install PATROL system on hardware to be delivered to Evaluation Team. Run PATROL system in evaluation mode on all tasks. Deliver system to Evaluation Team. ***Sites***: BBN. ***Criteria***: PATROL system runs as expected on deliverable hardware. ***Del***: QPR; software and hardware delivery to Evaluation Team on 1/2/2012. ***Dep***: 1.1.6, 1.1.7.

1.2. Feature Extraction and Projection

1.2.1. Project high-dimensional features via MLP

Dur: 10 months. ***Obj***: Determine whether cortical and modulation spectrum features work well with MLPs. Also determine whether MLPs can scale to high-dimensional features. ***App***: Feed high dimensional cortical and modulation spectrum features into MLP and train MLP for each of the four RATS applications; explore different MLP architectures and training methods; optimize code for speed; support MLP integration into PATROL system. ***Sites***: JHU. ***Criteria***: MLP-derived features perform better than PCA/LDA-projected features for each of the four RATS applications. ***Del***: QPR. ***Dep***: 1.1.2.

1.2.2. Explore feature-based fusion

Dur: 8 months. ***Obj***: Determine optimal procedure for combining/fusing different types of features (e.g., PLP, FDLP, MLP). ***App***: Investigate two methods for fusing multiple feature streams: (a) feature concatenation and MLP-based projection, and (b) Dempster-Shafer fusion. Run experiments on RATS data to determine which method works best for each of the four RATS applications. Develop software library that supports feature fusion functions. Optimize code for speed. ***Sites***: JHU. ***Criteria***: Improved results across all four RATS applications from fusing multiple feature streams. ***Del***: QPR.

1.3. Noise Compensation

1.3.1. Enhance cortical features

Dur: 13 months. ***Obj***: Develop methods for noise suppression in the cortical domain. ***App***: Implement filter that suppresses regions of noise in the cortical domain. Estimate parameters of filter on RATS data. Investigate filters that adapt their parameters based on an estimate of the noise. Periodically integrate and update enhancement filters into cortical feature extraction library. Coordinate with JHU on projecting enhanced cortical features via MLPs. Optimize code for speed. ***Sites***: UMD. ***Criteria***: Improved performance of RATS applications compared to using un-enhanced cortical features. ***Del***: QPR.

1.3.2. Investigate VTS for noise adaptation

Dur: 13 months. ***Obj***: Investigate vector Taylor series (VTS) for adapting HMMs to noise on RATS data. ***App***: Consult with Cambridge University on implementation of VTS noise adaptation. Incorporate VTS in models used across all four RATS applications. Measure performance with and without VTS. Repeat measurements after incorporating MLP-projected enhanced features from UMD/JHU. Optimize implementation for speed. ***Sites***: BBN,CU. ***Criteria***: Improved performance of RATS applications compared to only using feature-based noise compensation. ***Del***: QPR. ***Dep***: 1.1.2.

1.4. SAD

1.4.1. Develop cortical domain SVM for SAD

Dur: 12 months. ***Obj***: Develop cortical domain SAD system based on Support Vector Machines (SVMs). ***App***: Investigate the use of SVMs for speech/non-speech audio classification. Run experiments on RATS data as it becomes available. Initially, make use of cortical features with PCA/LDA dimensionality reduction. Later, use MLP-based projection from JHU. Release updated PATROL software libraries for SVM testing. ***Sites***: UMD. ***Criteria***: Improved SAD performance on RATS data compared to standard speech/non-speech GMM baseline. ***Del***: QPR. ***Dep***: 1.1.4.

1.4.2. Build SOU HMM for SAD

Dur: 9 months. ***Obj***: Develop SAD system based on self-organized unit (SOU) hidden Markov models (HMMs). ***App***: Apply the process of clustering acoustic trajectories in RATS data into a number of SOUs for speech and non-speech. Design and train an HMM for each labeled SOU in the RATS training data. Contrast such system with a baseline speech/non-speech GMM SAD. Integrate MLP-projected features from UMD/JHU as they become available. Optimize SOU HMM SAD system for speed. ***Sites***: BBN. ***Criteria***: Improved SAD performance on RATS data compared to simple speech/non-speech GMM baseline. ***Del***: QPR. ***Dep***: 1.1.2.

1.5. LID

1.5.1. Establish baseline BUT LID system on RATS data

Dur: 5 months. ***Obj***: Establish a baseline BUT LID system tuned on RATS data. ***App***: Train and test all components of BUT LID system on RATS data. Optimize each component for speed and accuracy. Iteratively integrate optimized components to PATROL system. ***Sites***: BUT. ***Criteria***: BUT LID system runs properly on RATS data and achieves improved performance compared to initial un-tuned LID system. ***Del***: QPR.

1.5.2. Incorporate MLP features in BUT LID system

Dur: 6 months. ***Obj***: Test MLP-projected cortical or modulation spectrum features in LID system. ***App***: Periodically integrate UMD/JHU's MLP-projected cortical or modulation spectrum features into BUT LID system and re-optimize LID on RATS data. Compare against current best baseline on RATS development set. ***Sites***: BUT. ***Criteria***: Perform periodically until best front-end processing is integrated to Phase 1 BUT LID system components. ***Del***: QPR. ***Dep***: 1.1.2, 1.5.1.

1.5.3. Develop SOU-based LID system

Dur: 13 months. ***Obj***: Develop SOU-based BBN LID system. ***App***: Tokenize RATS training data based on self-organized units (SOUs). Estimate language-specific SOU GMM and HMM acoustic models. Train language-specific SOU n-gram models. Optimize LID system components for speed and accuracy. Incorporate MLP features as they become available. ***Sites***: BBN. ***Criteria***: BBN LID system runs properly on RATS data and achieves performance comparable to or better than BUT's LID system on RATS devset. ***Del***: QPR. ***Dep***: 1.1.2.

1.6. SID

1.6.1. Develop JHU SID system

Dur: 12 months. ***Obj***: Develop JHU SID system tuned on RATS data. ***App***: Train and test all components of JHU SID system on RATS data. Optimize each component for speed and accuracy. Incorporate MLP features as they become available. Iteratively integrate optimized components to PATROL system. ***Sites***: JHU. ***Criteria***: JHU SID system runs properly on RATS data and achieves improved performance compared to initial un-tuned SID system. ***Del***: QPR. ***Dep***: 1.1.2.

1.6.2. Establish baseline BUT SID system on RATS data

Dur: 5 months. ***Obj***: Establish a baseline BUT SID system tuned on RATS data. ***App***: Train and test all components of BUT SID system on RATS data. Optimize each component for speed and accuracy. Iteratively integrate optimized components to PATROL system. ***Sites***: BUT. ***Criteria***: BUT SID system runs properly on RATS data and achieves improved performance compared to initial un-tuned SID system. ***Del***: QPR.

1.6.3. Incorporate MLP features in BUT SID system

Dur: 6 months. ***Obj***: Test MLP-projected cortical or modulation spectrum features in LID system. ***App***: Periodically integrate UMD/JHU's MLP-projected cortical or modulation spectrum features into BUT LID system and re-optimize LID on RATS data. Compare against current best baseline on RATS development set. ***Sites***: BUT. ***Criteria***: Perform periodically until best front-end processing is integrated to Phase 1 BUT LID system components. ***Del***: QPR. ***Dep***: 1.1.2, 1.6.2.

1.6.4. Try BBN's ASR model adaptation parameters as features for SID

Dur: 4 months. ***Obj***: Investigate the use of BBN's ASR speaker adaptation parameters as features in SID. ***App***: Train SID classifiers that rely on ASR speaker adaptation parameters (MLLR, CMLLR matrices) provided by BBN. Test on RATS development data. Repeat investigation periodically as BBN's ASR system evolves during Phase 1. ***Sites***: BUT. ***Criteria***: Perform periodically until best SID classification results are obtained in time for integration to final Phase 1 PATROL system. ***Del***: QPR.

1.6.5. Develop SOU-based SID system

Dur: 13 months. ***Obj***: Develop SOU-based BBN SID system. ***App***: Tokenize RATS training data based on self-organized units (SOUs). Estimate universal background SOU HMM (UBM) . Train speaker-dependent SOU HMM acoustic models. Train speaker-specific SOU n-gram models. Optimize SID system components for speed and accuracy. Incorporate MLP features as they become available. ***Sites***: BBN. ***Criteria***: BBN SID system runs properly on RATS data and achieves performance comparable to or better than BUT's SID system on RATS devset. ***Del***: QPR.

1.7. KWS

1.7.1. Establish baseline KWS using LVCSR

Dur: 9 months. ***Obj***: Develop Arabic and Farsi KWS systems using large vocabulary recognizers (LVCSR). ***App***: Demonstrate that the performance is better than using phonemes or sub-words, and better than using a small vocabulary that includes only the

keywords plus sub-words. Initially, run experiments on existing Arabic LDC corpora. When RATS data becomes available, retrain KWS system and measure performance on RATS Arabic and Farsi devsets. ***Sites***: BBN, CU. ***Criteria***: A KWS system using LVCSR. Comparative results with phoneme based KWS. ***Del***: QPR.

1.7.2. Integrate MLP features from UMD/JHU

Dur: 6 months. ***Obj***: Integrate improved MLP-projected features from UMD and JHU. ***App***: Periodically evaluate the impact of UMD/JHU MLP-projected features on KWS. After each feature update, retrain and test KWS systems on RATS data. Optimize system components as needed on the new features. ***Sites***: BBN, CU. ***Criteria***: Improved KWS performance due to the use of more advanced front-ends. ***Del***: QPR. ***Dep***: 1.1.2, 1.7.1.

1.7.3. Develop confidence measures for keywords

Dur: 5 months. ***Obj***: Develop good confidence measures for keywords on noisy data. ***App***: Associate each keyword with features extracted from phonetic lattices. Train classifier (e.g., generalized linear model) on RATS annotated data to predict probability of the recognized keyword being correct. ***Sites***: BBN, CU. ***Criteria***: Obtain confidence measures that can be used to decide whether to return a recognized keyword. ***Del***: QPR. ***Dep***: 1.7.1.

1.7.4. Develop adaptation procedures for KWS

Dur: 6 months. ***Obj***: Investigate adaptation of KWS tokenizers and classifiers for improved accuracy. ***App***: Use high-confidence detected keywords on test set in order to adapt KWS models. Re-decode test data with adapted models. Optimize adaptation and recognition stages for speed. ***Sites***: BBN. ***Criteria***: Improved KWS performance due to confidence-based adaptation. ***Del***: QPR. ***Dep***: 1.7.3.

1.7.5. Generate pronunciations for OOV keywords

Dur: 4 months. ***Obj***: Generate set of alternative pronunciations for keywords that are not in the training lexicon. ***App***: Explore two approaches to handling unknown keywords. First, use graphemic ASR models, that map each letter in a word to a phoneme. Second, develop algorithms for generating a set of plausible pronunciations from the surface word form, and use these pronunciations within phonetic ASR models for KWS. ***Sites***: BBN. ***Criteria***: Ability to spot words that are outside of the training lexicon. ***Del***: QPR. ***Dep***: 1.7.1.

2. Phase 2

2.1. Software Integration and Delivery

2.1.1. Develop system for training in the field

Dur: 10 months. ***Obj***: Develop infrastructure to support training in the field for all four RATS applications. ***App***: Start with snapshot of Phase 1 PATROL system; integrate all BBN training modules into deliverable pipeline; provide GUI-accessible functions that perform training on given data; test training pipeline extensively on deliverable hardware; optimize components for speed; write supporting documentation. ***Sites***: BBN. ***Criteria***: Deliverable system performs all training tasks reliably and efficiently. ***Del***: QPR; software and hardware delivery to Government on 1/2/2013.

2.1.2. Develop GUI for trainable system

Dur: 5 months. ***Obj***: Develop a graphical user interface (GUI) for the trainable version of the PATROL system. ***App***: Implement military user friendly GUI that allows the user to train and test the PATROL system on a given set of data. Write source code that interprets GUI inputs and calls corresponding functions in the PATROL system library. ***Sites***: BBN. ***Criteria***: GUI is intuitive, easy to use and allows user to perform all training/testing tasks. ***Del***: QPR; software. ***Dep***: 2.1.1.

2.1.3. Integrate PATROL research components into Phase 2 deliverable system

Dur: 7 months. ***Obj***: Continuously integrate research components into PATROL system. Ensure that PATROL system runs within specified time limits in training/evaluation mode, on deliverable hardware. ***App***: Iteratively integrate system components to PATROL system as they become updated by research efforts. Test component connections and measure performance characteristics (accuracy, speed). Optimize components as needed in order to meet target speed. ***Sites***: BBN. ***Criteria***: PATROL system achieves expected accuracy and runs within specified time constraints. Task repeats until the start of Phase 2 evaluation. ***Del***: QPR.

2.1.4. Test and prepare P2 PATROL system for delivery

Dur: 2 months. ***Obj***: Ensure that Phase 2 PATROL system runs with expected accuracy and speed in training/evaluation mode. ***App***: Install PATROL system on hardware to be delivered to Government and Evaluation Team. Run PATROL system in training/evaluation mode on all tasks. Deliver system to Government and Evaluation Team. ***Sites***: BBN. ***Criteria***: PATROL system runs as expected on deliverable hardware. ***Del***: QPR; software and hardware delivery to Government and Evaluation Team on 1/2/2013. ***Dep***: 2.1.1, 2.1.2, 2.1.3.

2.2. Feature Extraction and Projection

2.2.1. Explore hierarchical MLP training

Dur: 7 months. ***Obj***: Determine whether MLP-based projection can be improved via hierarchical training. ***App***: Use output of trained MLP as input to second level MLP; optimize code for speed; support MLP integration into PATROL system. ***Sites***: JHU. ***Criteria***: Improved classification performance in RATS applications. ***Del***: QPR.

2.2.2. Project high-dimensional features via RDT

Dur: 12 months. ***Obj***: Determine whether RDT-based feature projection is better than MLPs. ***App***: Train discriminative RDT to project UMD/JHU features to lower dimensional sub-space for each of the four RATS applications; optimize code for speed. ***Sites***: BBN. ***Criteria***: Improved classification performance in RATS applications. ***Del***: QPR.

2.3. Noise Compensation

2.3.1. Investigate RDT-based enhancement

Dur: 12 months. ***Obj***: Apply RDT to cortical spectrum in order to suppress the noise. ***App***: Estimate clean speech model from clean corpus; Initialize RDT mapping based on UMD's noise filter; Train RDT discriminatively on RATS data to improve classification with clean speech model. ***Sites***: BBN,UMD. ***Criteria***: Improved classification performance in RATS applications. ***Del***: QPR.

2.3.2. Investigate JUD and PCMLLR for noise adaptation

Dur: 10 months. ***Obj***: Investigate joint uncertainty decoding (JUD) and predictive CMLLR (PCMLLR) for adapting HMMs to noise on RATS data. ***App***: Consult with Cambridge University on implementation of JUD and PCMLLR noise adaptation. Incorporate compensated models used across all four RATS applications. Contrast performance with VTS noise compensation. Repeat measurements after incorporating MLP/RDT-projected enhanced features from UMD/JHU. Optimize implementation for speed. ***Sites***: BBN,CU. ***Criteria***: Improved performance of RATS applications compared to only using feature-based noise compensation. ***Del***: QPR.

2.3.3. Investigate condition-adaptive fusion

Dur: 8 months. ***Obj***: Determine optimal procedure for combining/fusing scores and features from multiple classifiers (e.g., HMMs, GMMs). ***App***: Use measured features of input (e.g., SNR) as additional features in final classification stage, along

with scores from other classifiers. Develop software library that supports score fusion functions. Contrast with Dempster-Shafer methods. Optimize code for speed. ***Sites***: JHU. ***Criteria***: Improved results across all four RATS applications from fusing multiple classifier scores. ***Del***: QPR.

2.4. SAD

2.4.1. Build phoneme-based HMM for SAD

Dur: 6 months. ***Obj***: Develop SAD system based on phoneme-class HMMs. ***App***: Use semi-supervised training methods to label RATS SAD training data at phoneme level. Design and train an HMM for each class of phonemes. Contrast such system with the SOU HMM SAD. Integrate MLP/RDT-projected features from UMD/JHU as they become available. Optimize HMM SAD system for speed. ***Sites***: BBN. ***Criteria***: Improved SAD performance on RATS data compared to simple speech/non-speech GMM baseline. ***Del***: QPR.

2.4.2. Investigate adaptation for SAD

Dur: 5 months. ***Obj***: Improve SAD performance through unsupervised adaptation. ***App***: Use high-confidence speech/non-speech boundaries detected on test data to adapt acoustic models and then retest with adapted models. ***Sites***: BBN. ***Criteria***: Improved SAD performance on RATS data compared to un-adapted SAD baseline. ***Del***: QPR.

2.5. LID

2.5.1. Develop phoneme-based LID system

Dur: 12 months. ***Obj***: Develop phoneme-based BBN LID system. ***App***: Use semi-supervised training methods to label RATS training data based on phonemes. Estimate language-specific phoneme HMM acoustic models. Train language-specific phoneme n-gram models. Optimize LID system components for speed and accuracy. Incorporate MLP/RDT-projected features as they become available. ***Sites***: BBN. ***Criteria***: Comparable performance to SOU-based LID. ***Del***: QPR.

2.5.2. Discriminative training of HMMs for LID

Dur: 12 months. ***Obj***: Explore discriminative training of HMMs for LID. ***App***: Estimate HMM parameters so as to maximize the posterior probability of the correct language on RATS training data. Incorporate MLP/RDT-projected features as they become available. ***Sites***: BBN. ***Criteria***: Improved performance compared to maximum likelihood HMM LID. ***Del***: QPR.

2.5.3. Improve JFA in LID

Dur: 12 months. ***Obj***: Investigate the use of joint factor analysis (JFA) for modeling noise variability in LID. ***App***: Apply JFA on RATS data. Extend JFA framework to account for noise variability. ***Sites***: BUT. ***Criteria***: Improved performance compared to baseline BUT system. ***Del***: QPR.

2.6. SID

2.6.1. Develop phoneme-based SID system

Dur: 12 months. ***Obj***: Develop phoneme-based BBN SID system. ***App***: Tokenize RATS training data based on phonemes. Estimate universal background phoneme HMM (UBM) . Train speaker-dependent phoneme HMM acoustic models. Train speaker-specific phoneme n-gram models. Optimize SID system components for speed and accuracy. Incorporate MLP features as they become available. ***Sites***: BBN. ***Criteria***: BBN SID system runs properly on RATS data and achieves performance comparable to or better than BUT's SID system on RATS devset. ***Del***: QPR.

2.6.2. Discriminative training of HMMs for SID

Dur: 9 months. ***Obj***: Explore discriminative training of HMMs for SID. ***App***: Estimate HMM parameters so as to maximize the posterior probability of the correct speaker on RATS training data. Incorporate MLP/RDT-projected features as they become available. ***Sites***: BBN,JHU. ***Criteria***: Improved performance compared to maximum likelihood HMM SID. ***Del***: QPR.

2.6.3. Improve JFA in SID

Dur: 12 months. ***Obj***: Investigate the use of joint factor analysis (JFA) for modeling noise variability in LID. ***App***: Apply JFA on RATS data. Extend JFA framework to account for noise variability. ***Sites***: BUT. ***Criteria***: Improved performance compared to baseline BUT system. ***Del***: QPR.

2.7. KWS

2.7.1. Investigate semi-supervised training methods

Dur: 10 months. ***Obj***: Explore semi-supervised training methods for KWS acoustic model training. ***App***: Build initial acoustic models on transcribed RATS corpus, then decode large un-transcribed audio and train on select automatic transcripts. ***Sites***: BBN. ***Criteria***: Improved KWS performance compared to training on manually transcribed data. ***Del***: QPR.

2.7.2. Use topic-weighted LM

Dur: 8 months. ***Obj***: Investigate topic-weighted language model (LM) for KWS. ***App***: **Run UTD on transcribed text for the language to derive a set of topics.** Run recognition on test passage; detect topic automatically; derive topic-weighted LM; rerun recognition. ***Sites***: BBN. ***Criteria***: Improved KWS performance compared to using a generic LM. ***Del***: QPR.

2.8. Speech Transcription Option

2.8.1. Establish ASR baseline on RATS devset

Dur: 5 months. ***Obj***: Establish training/test sets and proper ASR baseline using RATS data. ***App***: Coordinate with Data Team to form a proper RATS ASR development set; build initial RATS ASR system on designated training corpus. ***Sites***: BBN,CU. ***Criteria***: Reasonable ASR accuracy on RATS data. ***Del***: QPR.

2.8.2. Develop ASR RATS system

Dur: 7 months. ***Obj***: Further improve accuracy of PATROL ASR system. ***App***: Incorporate features and noise compensation methods developed for KWS as they become available. Explore discriminative training and adaptation methods. ***Sites***: BBN,CU. ***Criteria***: Improved ASR accuracy compared to baseline ASR system on RATS data. ***Del***: QPR. ***Dep***: 2.8.1.

3. Phase 3

3.1. Software Integration and Delivery

3.1.1. Develop system for training in the field

Dur: 10 months. ***Obj***: Develop infrastructure to support training in the field for all four RATS applications. ***App***: Integrate training modules from UMD/JHU/BUT into deliverable pipeline; test training pipeline extensively on deliverable hardware; optimize components for speed; write supporting documentation. ***Sites***: BBN. ***Criteria***: Deliverable system performs all training tasks reliably and efficiently. ***Del***: QPR; software and hardware.

3.1.2. Integrate PATROL research components into Phase 3 deliverable system

Dur: 10 months. ***Obj***: Continuously integrate research components into PATROL system. Ensure that PATROL system runs within specified time limits in training/evaluation mode, on deliverable hardware. ***App***: Iteratively integrate system components to PATROL system as they become updated by research efforts. Test component connections and measure performance characteristics (accuracy, speed). Optimize components as needed in order to meet target speed. ***Sites***: BBN. ***Criteria***: PATROL system achieves expected accuracy and runs within specified time constraints. Task repeats until the start of Phase 3 evaluation. ***Del***: QPR. ***Dep***: 3.1.1.

3.1.3. Test and prepare P3 PATROL system for Delivery

Dur: 2 months. ***Obj***: Ensure that Phase 3 PATROL system runs with expected accuracy and speed in training/evaluation mode. ***App***: Install PATROL system on hardware to be delivered to Government and Evaluation

Team. Run PATROL system in evaluation mode on all tasks. Deliver system to Government and Evaluation Team. ***Sites***: BBN. ***Criteria***: PATROL system runs as expected on deliverable hardware. ***Del***: QPR. Software and hardware delivery to Government and to Evaluation Team on 1/2/2014. ***Dep***: 3.1.1, 3.1.2.

3.2. Feature Extraction and Projection

3.2.1. Improve feature extraction methods

Dur: 12 months. ***Obj***: Improve existing front-end representations. ***App***: Analyze results obtained in all four RATS applications and adjust existing or develop new feature extraction modules in order to achieve further improvements to noise robustness; update software libraries; optimize code for speed. ***Sites***: UMD,JHU. ***Criteria***: Improved robustness to noise. ***Del***: QPR.

3.3. Noise Compensation

3.3.1. Investigate condition-adaptive fusion

Dur: 8 months. ***Obj***: Determine optimal procedure for combining/fusing scores and features from multiple classifiers (e.g., HMMs, GMMs). ***App***: Continue development of condition-adaptive fusion, investigating new features such as prosody, articulation, phoneme rate. ***Sites***: JHU. ***Criteria***: Improved results across all four RATS applications from fusing multiple classifier scores. ***Del***: QPR.

3.3.2. Investigate RDT-based enhancement

Dur: 12 months. ***Obj***: Apply RDT to cortical spectrum in order to suppress the noise. ***App***: Continue developing RDT based enhancement by investigating different methods for partitioning the cortical feature space into regions. ***Sites***: BBN,UMD. ***Criteria***: Improved classification performance in RATS applications. ***Del***: QPR.

3.3.3. Investigate Noisy CMLLR for noise adaptation

Dur: 10 months. ***Obj***: Investigate Noisy CMLLR (NCMLLR) for adapting HMMs to noise on RATS data. ***App***: Consult with Cambridge University on implementation of NCMLLR adaptation. Incorporate compensated models used across all four RATS applications. Contrast performance with Phase 2 noise compensation methods. Repeat measurements after incorporating MLP/RDT-projected enhanced features from UMD/JHU. Optimize implementation for speed. ***Sites***: BBN,CU. ***Criteria***: Improved performance of RATS applications compared to only using feature-based noise compensation. ***Del***: QPR.

3.3.4. Develop noise adaptive training methods

Dur: 9 months. ***Obj***: Derive acoustic models that adapt better to noise. ***App***: Apply noisy CMLLR adaptation methods to training in order to obtain canonical, noise-neutral acoustic models; also investigate discriminative methods for estimating the adaptation parameters. ***Sites***: BBN,CU. ***Criteria***: Improved adaptation performance on test. ***Del***: QPR.

3.4. SAD

3.4.1. Investigate factorial HMM for SAD

Dur: 12 months. ***Obj***: Explore the use of factorial HMMs for SAD. ***App***: Estimate clean speech phoneme-class models from clean speech corpus; initialize factorial HMM noise states based on identity transforms; train factorial model on RATS data, allowing noise transforms to model the effect of noise on clean speech; investigate both maximum likelihood and discriminative training criteria. ***Sites***: BBN. ***Criteria***: Improved SAD performance compared to Phase 2 SAD models. ***Del***: QPR.

3.4.2. Explore sequence scores for SAD

Dur: 8 months. ***Obj***: Investigate the use of higher-level features in SAD. ***App***: Apply phonotactic-type modeling to learn patterns about the sequences of phonemes or SOUs that are characteristic of speech; fuse such scores with acoustic HMM SAD scores. ***Sites***: BBN. ***Criteria***: Improved SAD performance compared to Phase 2 SAD models. ***Del***: QPR.

3.5. LID

3.5.1. Use SAT for LID

Dur: 10 months. ***Obj***: Use speaker adaptive training (SAT) in LID. ***App***: Apply speaker change detection and clustering methods on RATS LID training data; Estimate adaptation transforms for each speaker cluster; incorporate transforms during LID HMM training in order to obtain speaker-neutral models; investigate discriminative training extensions. ***Sites***: BBN. ***Criteria***: Improved performance compared to standard discriminative HMM LID training. ***Del***: QPR.

3.5.2. Use information retrieval methods for LID

Dur: 10 months. ***Obj***: Explore techniques for global phonotactic modeling in LID. ***App***: Collect statistics about phoneme or SOU co-occurrences in speech tokenization output; build models that look beyond local n-gram context; include scores from such models in LID fusion. ***Sites***: BUT. ***Criteria***: Improved performance compared to baseline LID system. ***Del***: QPR.

3.5.3. JFA for HMM-based LID

Dur: 12 months. ***Obj***: Extend JFA to work with HMM-based SID systems. ***App***: Estimate basis of MLLR transformations to model variability in HMM Gaussian parameters; adapt HMMs using JFA-inspired methods. ***Sites***: BUT,BBN. ***Criteria***: Improved performance compared to baseline HMM SID system. ***Del***: QPR.

3.6. SID

3.6.1. JFA for HMM-based SID

Dur: 12 months. ***Obj***: Extend JFA to work with HMM-based SID systems. ***App***: Estimate basis of MLLR transformations to model variability in HMM Gaussian parameters; adapt HMMs using JFA-inspired methods. ***Sites***: BUT,BBN. ***Criteria***: Improved performance compared to baseline HMM SID system. ***Del***: QPR.

3.7. KWS

3.7.1. Apply SVM phonetic rescoring

Dur: 10 months. ***Obj***: Use phoneme or word-based SVM models to complement HMMs. ***App***: Estimate an SVM for each context-dependent phoneme or word in training; rescore lattice of tokens in test using SVMs; include SVM scores in KWS fusion. ***Sites***: BBN. ***Criteria***: Improved performance compared to baseline HMM KWS system. ***Del***: QPR.

3.7.2. Adapt LM to new keywords

Dur: 10 months. ***Obj***: Improve estimate of n-gram LM probability for new keywords. ***App***: Run KWS system on test data; find high-confidence detections of the new keyword; look at n-gram context; find words that have been observed in similar contexts in LM training; re-estimate LM probability of new keyword based on these similar words; re-run recognition on test, using adapted LM. ***Sites***: BBN. ***Criteria***: Improved performance compared to baseline generic LM KWS system. ***Del***: QPR.

3.8. Speech Transcription Option

3.8.1. Develop ASR RATS system

Dur: 12 months. ***Obj***: Further improve accuracy of PATROL ASR system. ***App***: Incorporate features and noise compensation methods developed for KWS as they become available. Explore semi-supervised training methods. ***Sites***: BBN,CU. ***Criteria***: Improved ASR accuracy compared to baseline ASR system on RATS data. ***Del***: QPR.

6 Intellectual Property

Any data that BBN collects or acquires under this effort will be made available to all sites in the RATS program.

Noncommercial Items (Technical Data and Computer Software)

NONCOMMERCIAL ITEMS			
Technical Data and Computer Software To be Furnished With Restrictions	**Basis for Assertion**	**Asserted Rights Category**	**Name of Person Asserting Restrictions**
Software Suite to be developed under the RATS Program including existing libraries as needed	Developed with mixed funding	Negotiated Government Purpose Rights	Raytheon BBN Technologies Corp.

Commercial Items (Technical Data and Computer Software)

COMMERCIAL ITEMS			
Technical Data and Computer Software To be Furnished With Restrictions	**Basis for Assertion**	**Asserted Rights Category**	**Name of Person Asserting Restrictions**
Byblos Software*	Multiple commercial sales	Commercial	Raytheon BBN Technologies Corp.

*Byblos Software will be made available to the Government for the duration of the program effort pursuant to a no-cost license.

Patents

BBN hereby represents that it is the sole or joint owner of the inventions that will be utilized in the proposed program. BBN asserts that it has all of the necessary rights to utilize them to execute the program. BBN will obtain for the Government any necessary licenses in its subcontractor's inventions.

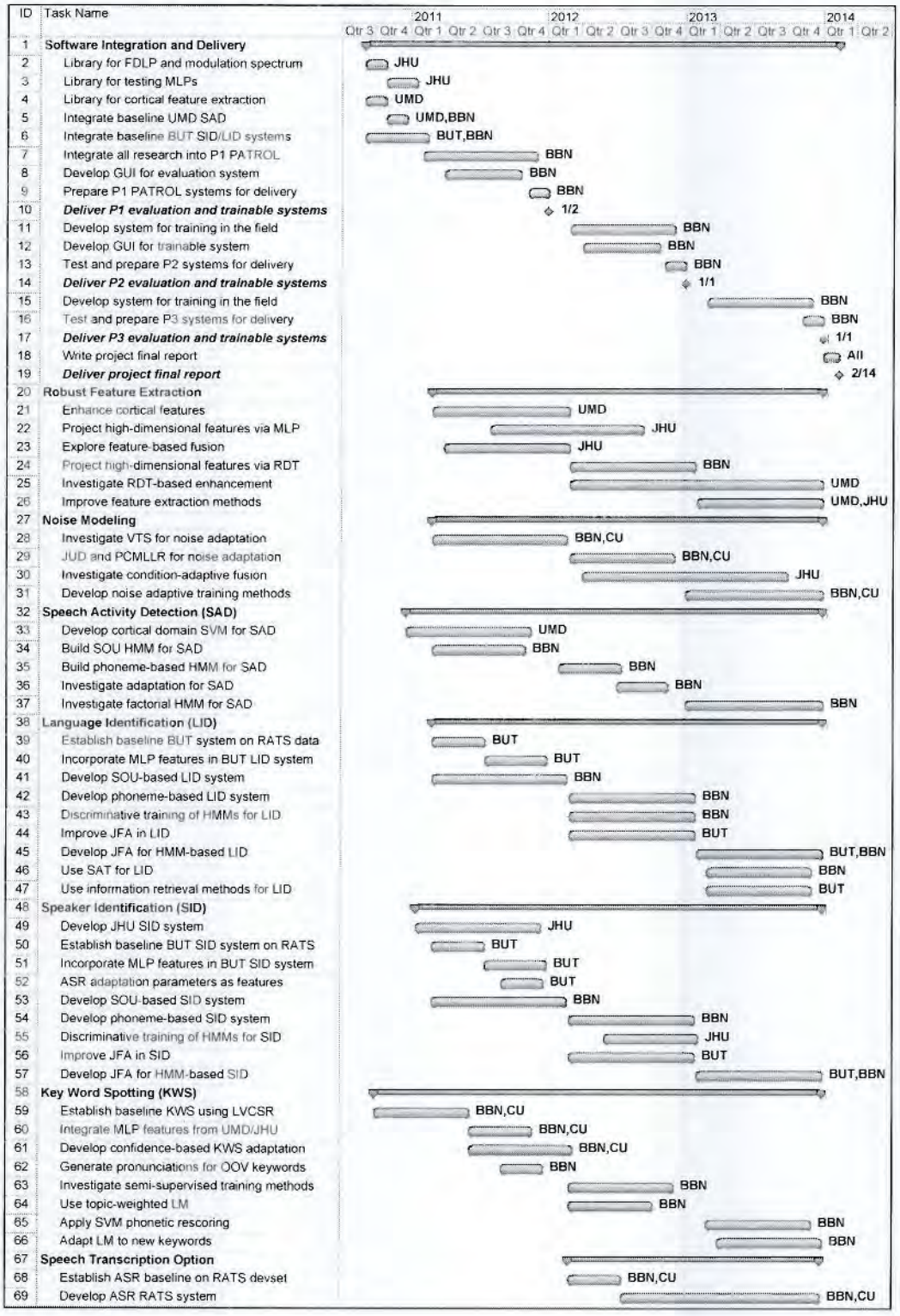
Intellectual Property Representations

BBN provides its good faith representation that BBN either owns or possesses appropriate licensing rights to all other intellectual property that will be utilized under BBN's proposal for the RATS Program. BBN will obtain from its subcontractors appropriate licensing right to all subcontractor intellectual property that will be utilized under BBN's proposal for the RATS Program.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

7 Schedule and Milestones

The proposed schedule is given below. Milestones are indicated in ***bold italic*** font.



Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

8 Personnel, Qualifications, and Commitments

For the very challenging problem of performing automatic recognition of language, speaker, and speech in real-world high-noise environments, we have assembled a team ("PATROL") that is equal to the task. Starting with the overall leadership of an experienced staff at BBN, we have added four sites that have unique world-class capabilities in all areas of this project. Dr. John Makhoul at BBN will serve as PI for this effort. He will be assisted at BBN by Rich Schwartz as Technical Lead and Spyros Matsoukas as co-PI and Technical Manager, both of whom played major roles in the DARPA EARS and GALE programs, which required leading multi-site teams working across geographically diverse sites. Makhoul, Schwartz, and Matsoukas will work with staff at BBN with wide and deep experience in all areas of recognition problems in speech, as well as the building and deployment of multi-media operational systems.

Because of the importance of dealing with high noise in this project, we have added three sites that have made significant contributions to the problem of speech in noise. Dr. Hynek Hermansky of Johns Hopkins University brings years of experience in developing robust methods of speech processing, many of which have been adopted by researchers worldwide. Dr. Shihab Shamma of the University of Maryland has developed a unique multi-dimensional cortical representation of sound that he has used successfully to combat many types of unwanted signals and noise in speech. Professors Mark Gales and Phil Woodland of Cambridge University, who have developed one of the top-performing speech recognition systems in the world, have also made novel contributions to the modeling of noise in speech, to the use of discriminative training and adaptation (which will be important to dealing with noise), and in performing speech recognition and keyword spotting under noisy conditions. To round out the team, we have added Drs. Lukas Burget and Pavel Matejka of the Brno University of Technology (BUT), who have developed world-class language ID and speaker ID systems that have competed successfully in NIST evaluations, year after year, scoring top marks in many of the conditions that were competed.

Below is a listing of key personnel in each organization and the unique capabilities they bring to this effort. A table of the time commitments of these key personnel is given in Section 8.6.

8.1 BBN

Key personnel at BBN include John Makhoul, Richard Schwartz, and Spyros Matsoukas. (As the need arises, we will draw upon the extensive reservoir of expert advice from other senior staff at BBN, including Herb Gish, Owen Kimball, Man-hung Siu, Walter Andrews, Long Nguyen, Rohit Prasad, Amit Srivastava, and Prem Natarajan. These staff's expertise spans the full range of the RATS applications, as well as the integration and delivery of operational speech systems.)

Dr. John Makhoul, a Chief Scientist at BBN, is also an Adjunct Professor at Northeastern University, where he supervises students who do their graduate thesis work at BBN. He has led projects in various aspects of speech and language processing, including speech coding, speech synthesis, speech recognition, speech enhancement, speaker recognition, artificial neural networks, optical character recognition, language understanding, information retrieval and extraction, human-machine interaction using voice, and machine translation. He has served as PI on numerous DoD speech and language projects, including the DARPA EARS speech recognition project and the ongoing DARPA GALE machine translation project, which includes work in speech recognition from Arabic and Chinese. Makhoul's involvement in speech

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

enhancement goes back to a DoD project in 1979 to enhance the quality of noisy speech. In 1989, he led a National Council panel that produced a study about the Removal of Noise from Noise-Degraded Speech Signals. Makhoul is a Life Fellow of IEEE and a Fellow of the Acoustical Society of America. He has received a number of IEEE awards for his technical accomplishments, including the 2009 IEEE James L. Flanagan Speech & Audio Processing Award and Medal "for pioneering contributions to speech modeling."

Richard Schwartz, a Principal Scientist at BBN, is a recognized expert in various aspects of pattern recognition, including speech recognition, HMM modeling, probabilistic language modeling, speaker verification and identification, speech recognition in noise, fast search, neural networks, robust multilingual OCR, and, more recently, machine translation. He has been technical lead or senior technical advisor on numerous DoD projects in those areas, including the EARS and GALE projects mentioned above. Schwartz was the chief designer of the BBN Byblos speech recognition system, which continues to form the backbone of all our state-of-the-art speech recognition and keyword spotting work. The Byblos system has been one of the consistent top performers in official government speech recognition evaluations. Among his many original contributions, Schwartz was the first to place speaker identification on a firm statistical basis in 1982; in 1990 he invented the N-best algorithm, an efficient method for computing the N-best sentence hypothesis in speech recognition; and in 1993 he was the first to demonstrate large-vocabulary, continuous speech recognition in real-time on COTS hardware, using a new search algorithm he had invented. He was a primary contributor to the BBN work on speech enhancement in 1979 – work that continues to be referenced to this day. He designed and supervised the implementation of the first speaker verification system that used hidden Markov models; the system was deployed at BBN in 1986. He has made numerous contributions to speaker change detection, speaker clustering, and search algorithms for 50x faster than real-time keyword spotting.

Spyros Matsoukas, a Lead Scientist at BBN, has been working on various aspects of statistical modeling in speech recognition since 1996. He was the top technical contributor in the DARPA EARS project and has taken over the leadership in the continued improvement of the BBN Byblos speech recognition system. Among his original technical contributions is the development of speaker-adaptive training techniques for speech recognition – which is now used in a number of research organizations worldwide – and segmental trajectory modeling for speech recognition. He has also done significant work in the detection of speech in the presence of noise and music, the detection of changes in the acoustic channel and the speaker, and speaker clustering. Matsoukas was the direct supervisor of the PhD thesis of one of our Northeastern University research assistants whose paper with Matsoukas on region-dependent feature transforms for speech recognition won a Spoken Language Processing Grant best paper award at the 2006 IEEE ICASSP conference. More recently, Matsoukas has been in charge of the machine translation work under the DARPA GALE project, which has borrowed a number of statistical modeling techniques from the speech area. A paper resulting from work that he led on system combination in machine translation won the best paper award at the 2007 Association for Computation Linguistics conference on Human Language Technology. In addition to overseeing the technical work of the machine translation work at BBN, he also manages the technical interaction with our subcontractors and is responsible for integrating the work of all sites to produce the highest possible quality machine translation from Arabic and Chinese (speech or text) to English. Matsoukas' excellent cross-site technical leadership skills will serve us well in this proposed work.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

8.2 Johns Hopkins University

Dr. Hynek Hermansky is a Full Professor of Electrical and Computer Engineering at the Johns Hopkins University in Baltimore, Maryland. He is also a Professor (on leave of absence) at the Brno University of Technology, Czech Republic; an Adjunct Professor at the Oregon Health and Sciences University, Portland, Oregon; and an External fellow at the International Computer Science Institute at Berkeley, California. He is a Fellow of IEEE for "invention and development of perceptually-based speech processing methods." His main area of research is robust acoustic processing of speech. Among the techniques that he has pioneered are: Perceptual Linear Prediction (PLP), RASTA modulation frequency filtering, multi-stream recognition of speech, data-guided feature extraction, frequency-domain PLP, and posterior-based features for speech recognition. He has applied these methods to speech recognition, speaker recognition, and speech and audio coding. A number of the robust processing techniques that he pioneered are now in wide usage worldwide.

Dr. Hermansky has been working in speech processing for over 30 years and has held the following positions: Director of Research at the IDIAP Research Institute, Martigny, Switzerland; Research Fellow at the University of Tokyo; Research Engineer at Panasonic Technologies in Santa Barbara, California; Senior Member of Research Staff at US WEST Advanced Technologies; and Professor and Director of the Center for Information Processing at OHSU, Portland, Oregon.

In his professional activities, Dr. Hermansky served as Technical Chair at the 1998 ICASSP in Seattle and an Associate Editor for IEEE Transaction on Speech and Audio, and is a member of the Organizing Committee for ICASSP 2011 in Prague. He is member of the Editorial Board of Speech Communication, holds 7 US patents, and authored or co-authored over 200 papers in reviewed journals and conference proceedings.

8.3 University of Maryland

Dr. Shihab Shamma is a Full Professor, Department of Electrical Engineering, University of Maryland (UMD), College Park, with a joint appointment at the Institute for System Research, UMD. His research deals with questions in computational neuroscience, speech and audio processing, neuromorphic engineering, and the development of microsensor systems for experimental research and neural prostheses. Primary focus has been on studying the computational principles underlying the processing and recognition of complex sounds in the auditory system, and the relationship between auditory and visual processing. Signal processing algorithms inspired by data from neurophysiological and psychoacoustical experiments are being developed and applied in a variety of systems such as speech and voice recognition, diagnostics in industrial manufacturing, and underwater and battlefield acoustics.

Dr. Shamma has developed novel models of auditory cortical processing, inspired by physiological and psychoacoustic experiments over the last decade that have revealed a rich and flexible representation of the perceptually significant features of sound. The resultant higher-dimensional cortical representation (a frequency/rate/scale vector as a function of time) is able to separate speech sounds from other unwanted signals, such as various types of noise or music. He has been able to use this unique separation property to design multi-dimensional filters that enhance the quality of noisy speech and suppress the noise. He has also used the cortical representation to design front ends to a speech activity detector (SAD) for many types of noise. He has tested his SAD with animal vocalizations, music, and environmental sounds, including jet

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

engine noise, factory noise, destroyer engine sounds, military vehicles, cars, and speech babble recorded in different environments like restaurants, airports, and exhibition halls, and, more recently, reverberant speech. His speech activity detector has been found to be superior to other SAD methods in independent tests by several government agencies.

Dr. Shamma is a Fellow of the Acoustical Society of America, Senior Member of IEEE, member of the Association for Research in Otolaryngology, and member of the Society for Neuroscience. He is also Action Editor for the Journal of Computational Neuroscience, Academic Editor for PLoS, and Academic Board Member for Trends in Cognitive Sciences. He has been Co-organizer and Director of numerous Workshops and Symposia including most recently the Annual Telluride Workshops on Neuromorphic Engineering (1997- present, partially funded by NSF, ONR, DARPA, and the Whitaker and Gatsby Foundations), and the NIPS and COSYNE workshops on Neural Mechanisms of Music Perception (1999), Thalamo-cortical Processing (2002), and Attention and Streaming (2006). He is the holder of a patent on "A Cochlear Filter Bank with Switched Capacitor Circuits", and has patents pending on "Intelligibility Assessment Using Spectro-temporal Modulations" and "Speech Discrimination Based on Multiscale Spectro-Temporal Modulations."

8.4 Cambridge University

Dr. Mark Gales will serve as the primary person in charge of the RATS work at Cambridge University. Professor Phil Woodland will serve as senior technical advisor.

Dr. Mark Gales is Reader in Information Engineering at Cambridge University where he has been a member of faculty staff since 1999. Prior to this he was a Research Staff Member at IBM. He has worked on speech processing for over 18 years, with a particular interest in acoustic modeling, acoustic environment robustness, and speaker adaptation. He has participated in a range of DARPA/NIST automatic speech recognition evaluations since 1994, at both Cambridge and IBM. These include recent evaluations under the DARPA EARS and GALE programs. Dr. Gales has been awarded multiple IEEE paper awards for his work on speech recognition. From 2001 to 2004 he was a member of the IEEE Speech Technical Committee. He is currently an Associate Editor of the IEEE Signal Processing Letters, IEEE Transactions on Audio Speech and Language, and on the editorial board of Computer Speech and Language journal. He has published over 100 papers, primarily in the area of acoustic modeling for speech recognition.

Dr. Gales has extensive experience in noise robustness, speaker adaptation, and speech recognition. He developed one of the first approaches to model-based compensation for acoustic models to handle noise conditions – parallel model combination. This model-based approach to handling high levels of background noise has been refined to handle a range of important techniques to make it applicable to state-of-the-art speech processing applications, including discriminative adaptive training , and improving efficiency to allow its application to large systems. He developed one of the standard approaches for speech recognition and adaptive training, Constrained Maximum Likelihood Linear Regression (CMLLR), as well as refining this to be more suited for speaker adaptation in the presence of high levels of background noise. Under the DARPA EARS and GALE programs he has been involved in developing large vocabulary transcription systems for English, Mandarin and Arabic. In addition, he has worked on sequence kernels that form an important part of speaker identification and verification systems and has recently been applied to handling high-levels of background noise.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Phil Woodland is Professor of Information Engineering at the University of Cambridge where he leads the Speech Group. He has been a member of faculty staff at Cambridge since 1989. He has published almost 200 academic papers in the area of speech and language technology. Many of these papers are highly cited, including the most highly cited paper to appear in the journal *Computer Speech and Language*. He has won best paper awards for work in the areas of speaker adaptation and discriminative training of large vocabulary speech recognition systems. He has also worked on statistical language modeling, auditory modeling, statistical speech synthesis, and spoken document retrieval. His team has developed large-vocabulary speech recognition systems that have frequently given the lowest error rates in international research evaluations organized by the US Government. He was one of the original co-authors of the widely-used HTK toolkit, which now has about 100,000 registered users worldwide, and has continued to play a major role in its development. He was a member of the IEEE SPS Speech Technical Committee from 1999 to 2003, a member of the Editorial Board of Computer Speech and Language from 1994 until 2009, and is a current member of the Editorial Board of Speech Communication.

Prof. Woodland has extensive experience in many areas of speech technology, particularly in speech recognition systems, and has worked in the field for the last 25 years. Of special relevance to RATS is his work on linear transform-based adaptation approaches, in which he developed the widely-used MLLR technique, as well as techniques for unsupervised adaptation in low recognition rate situations, via lattice-based adaptation, and discriminative adaptation based methods. He has developed widely used discriminative training techniques and shown that such techniques are useful for situations where test data is poorly matched to the training. He has worked on developing speech transcription technology for many years, including work on the DARPA EARS and GALE programs, for both of these he was the PI at Cambridge University. During this time his team has developed systems in a number of languages including English, Mandarin and Arabic.

8.5 Brno University of Technology

Dr. Lukas Burget will serve as the primary person in charge of the RATS work at BUT. Working closely with Dr. Burget will be Dr. Pavel Matejka.

Dr. Lukas Burget has been an Assistant Professor at the Faculty of Information Technology (FIT), University of Technology, Brno, Czech Republic, since 2004. He serves as Scientific Director of the Speech@FIT research group. From 2000 to 2002, he was a visiting researcher at OGI, Portland, OR, under supervision of Professor Hynek Hermansky. In 2008, he was invited to lead the "Robust Speaker Recognition over Varying Channels" team at the Johns Hopkins University CLSP summer workshop, and will lead the team of BOSARIS workshop in 2010.

Dr. Burget's scientific interests are in the field of speech processing, namely acoustic modeling for speech, speaker and language recognition, including their software implementations. He has authored or co-authored more than 40 papers in journals and conferences. He was leader of the top-performing teams in the NIST LID 2005 and 2007 evaluations and the NIST SID 2006 and 2008 evaluations, scoring first in a remarkable number of conditions. He contributed significantly to the team developing AMI LVCSR systems successful in the NIST RT 2005, 2006 and 2007 evaluations.

Dr. Burget has participated in EU-sponsored projects "Multimodal Meeting Manager" (M4), "Augmented MultiParty Interaction" (AMI), and "Augmented MultiParty interaction with

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Distant Access" (AMIDA), as well as in several projects sponsored at the local Czech level. Currently, he is participating in the EU project "Mobile Biometry". He is principal investigator of US-Air Force EOARD sponsored project "Improving the capacity of language recognition systems to handle rare languages using radio broadcast data".

Dr. Pavel Matejka is senior researcher in the Speech@FIT research group, Department of Computer Graphics and Multimedia, Faculty of Information Technology, Brno University of Technology (BUT). He was a visiting scientist with the Anthropic speech processing group at the Oregon Graduate Institute of Science and Technology in 2002-2003. He has participated in the EU projects M4, AMI, and AMIDA mentioned above, and in language identification projects sponsored by US Air Force EOARD and the Czech Ministry of Defense. He took part in NIST 2005, 2007 and 2009 language recognition and NIST 2006 and 2008 speaker recognition evaluations, where BUT had excellent results. He is currently leading BUT's speaker identification activities in the Mobile Biometry EU project.

Dr. Matejka is author or co-author of more than 15 papers in journals and at reviewed international conferences. His research interests include speaker recognition, language identification, speech recognition (namely, phone recognition based on novel feature extraction using temporal patterns and neural networks). He is active in keyword spotting and on-line implementation of speech processing algorithms. He was finalist in the Student paper contest at ICASSP 2006 in Toulouse and participated in the JHU workshop group "Recovery from Model Inconsistency in Multilingual Speech Recognition" in 2007.

8.6 Time Commitments

Table 8-1 shows the planned time commitments for the Key Personnel for calendar years 2010, 2011, and 2012. The numbers for 2010 assume that the effort will start 15 August 2010, so 2010 is only a partial year. It is anticipated that a relatively small effort will be devoted for Technical Area 1 during the first six months of the project while data is being collected by the Data Collection team.

For the BBN Key Personnel, we have shown the time commitments both in terms of hours and percentage time. At certain universities, personnel time is accounted for by percentage of time rather than by hours, while at others, time is specified in months.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Key Individual	Project	Pending/Current	2010	2011	2012
John Makhoul	RATS	Proposed	90 hrs (15%)	528 hrs (30%)	616 hrs (35%)
	GALE	Current	350 hrs (60%)	403 hrs (23%)	n/a
Richard Schwartz	RATS	Proposed	100 hrs (17%)	792 hrs (45%)	880 hrs (50%)
	GALE	Current	410 hrs (70%)	440 hrs (25%)	n/a
Spyros Matsoukas	RATS	Proposed	153 hrs (25%)	1196 hrs (65%)	1472 hrs (80%)
	GALE	Current	429 hrs (70%)	460 hrs (25%)	n/a
Hynek Hermansky	RATS	Proposed	12%	25%	25%
	IARPA/ARL Speaker ID	Current	12.5%	12.5%	n/a
Shihab Shamma	RATS	Proposed	0.25 month	1 month	1 month
	IARPA/ARL Speaker ID	Current	1 month	2 months	n/a
	ONR	Pending	0.25 month	1 month	1 month
Mark Gales	RATS	Proposed	50 hrs	200 hrs	200 hrs
	GALE	Current	70	70	n/a
Phil Woodland	RATS	Proposed	20 hrs	50 hrs	50 hrs
	GALE	Current	70 hrs	70 hrs	n/a
Lukas Burget	RATS	Proposed	10%	50%	40%
	IARPA Speaker ID	Current	41%	40%	40%
Pavel Matejka	RATS	Proposed	10%	55%	40%
	IARPA Speaker ID	Current	61%	35%	40%

Table 8-1 Key personnel and their time commitments.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

8.7 Selected Bibliographies

Below is a selected bibliography from each of the PATROL sites with references pertaining to the RATS program.

8.7.1 BBN

R. Prasad, S. Tsakalidis, I. Bulyko, C. Kao, and P. Natarajan, "Pashto Speech Recognition with Limited Pronunciation Lexicon," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Dallas, TX, March 2010.

S. Tsakalidis, R. Prasad, and P. Natarajan, "Context-dependent Pronunciation Modeling for Iraqi ASR," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Taiwan, April 2009.

M. Siu, H. Gish and X. Yang, "Discriminatively Trained GMMs for Language Classification Using Boosting Methods", *IEEE Trans. on Audio, Speech and Language Processing*, Vol. 17, No. 1, 2009.

L. Nguyen, T. Ng, K. Nguyen, Rabih Zbib, and John Makhoul, "Lexical and phonetic modeling for Arabic automatic speech recognition," *InterSpeech*, Brighton, UK, pp. 712-715, 2009.

S. Novotney, R. Schwartz, and J. Ma, "Unsupervised Acoustic and Language Model Training with Small Amounts of Labeled Data," *ICASSP*, Taipei, China, pp. 4297-4300, 2009.

J. Ma and R. Schwartz, "Unsupervised versus Supervised Training of Acoustic Models," *Interspeech*, Brisbane, Australia, pp. 2378-2381, 2008.

J. Au-Yeung and M. Siu, "Evaluation of the robustness of the polynomial segment models to noisy environments with unsupervised adaptation", *Speech Communications*, Vol. 50, 2008.

K. Subramanian, R. Prasad, and P. Natarajan, "Optimal Estimation of Rejection Thresholds for Topic Spotting," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Honolulu, HI, April 2007.

D. R. H. Miller, M. Kleber, C. Kao, O. Kimball, T. Colthurst, S. A. Lowe, R. Schwartz, and H. Gish, "Rapid and Accurate Spoken Term Detection", *Interspeech*, pp. 314-371, Antwerp, Belgium, 2007.

J. Ma, S. Matsoukas, O. Kimball, and R. Schwartz, "Unsupervised Training on Large Amounts of Broadcast News Data," *ICASSP*, Toulouse, France, vol.3, pp. 1056-1059, 2006.

S. Saleem, R. Prasad, and P. Natarajan, "Colloquial Iraqi ASR for Speech-to-Speech Translation," *Intl. Conf. on Spoken Language Proce*ssing, Pittsburg, PA, September, 2006.

B. Zhang, S. Matsoukas, R. Schwartz: "Recent progress on the discriminative region-dependent transform for speech feature extraction", *Interspeech*, Pittsburgh, PA, September, 2006.

H. Jin, R. Schwartz, F. G. Walls, and S. P. Sista, "Method and Apparatus for Score Normalization for Information Retrieval Applications", US Patent No. 6,651,057, issued on November 18, 2003, and US Patent No. 7,062,485, issued on June 13, 2006 (continuation).

L. Zhai, M. Siu, X. Yang and H. Gish, "Discriminatively trained Language Models Using Support Vector Machines for Language Identification," *Odyssey*, June 2006.

B. Zhang, S. Matsoukas, and R. Schwartz, "Discriminatively Trained Region Dependent Feature Transforms for Speech Recognition," *ICASSP*, Toulouse, France, vol.1, pp.313-316, 2006.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

B. Zhang, S. Matsoukas, and R. Schwartz, "Recent Progress on the Discriminative Region-dependent Transform for Speech Feature Extraction," *Interspeech*, Pittsburgh, PA, pp. 1495-1498, 2006.

K.Y. Leung, M.W. Mak, M. Siu, and S.Y. Kung, "Speaker Verification Using Adapted Articulatory Feature-based Conditional Pronunciation Modeling," *IEEE Int. Conference on Acoustics, Speech and Signal Processing*, pp. 181-184, Philadelphia, PA, 2005.

B. Xiang, L. Nguyen, S. Matsoukas, and R. Schwartz, "Cluster-dependent acoustic modeling," *ICASSP 2005*, Philadelphia, Pennsylvania, 2005.

K.K. Wong, M. Siu, "Automatic language identification using discrete hidden Markov model," *International Conference on Spoken Language Processing*, pp. 1633-1636, Jeju Island, Korea, 2004.

W. Andrews and J. Hernandez-Cordero, "Using Speech Compression for Speaker Detection", *Speaker Odyssey Workshop*, Toledo, Spain, 2004.

L. Nguyen and B. Xiang, "Light supervision in acoustic model training," *ICASSP*, Montreal, Canada, 2004.

D. Reynolds, W. Andrews, J. Campbell, J. Navrtil, B. Peskin, A. Adami, Q. Jin, D. Kluscek, J. Abramson, R. Mihaescu, J. Godfrey, D. Jones, and B. Xiang, "SuperSID Project: Exploiting High-Level Information for High Accuracy Speaker Recognition", *ICASSP*, Hong Kong, China, 2003.

J. Navrtil, Q. Jin, W. Andrews and J. Campbell, "Phonetic Speaker Recognition Using Maximum-Likelihood Binary-Decision Tree Models", *ICASSP*, Hong Kong, China, 2003.

L. Nguyen, X. Guo, R. Schwartz, and J. Makhoul, "Japanese broadcast news transcription system," *ICSLP*, Denver, Colorado, 2002.

M. Siu, A. Chan, "Robust Speech Recognition Against Short-Time noise," *International Conference on Spoken Language Processing*, Vol. 2, pp. 1049-1052, 2002.

R. Prasad, L. Nguyen, R. Schwartz, and J. Makhoul, "Automatic Transcription of Courtroom Speech," *Int. Conf. Spoken Language Processing*, Denver, CO, pp.1745-1748, 2002.

R. Prasad, L. Nguyen, R. Schwartz, and J. Makhoul, "MeetingLogger: Rich Transcription of Courtroom Speech," *Int. Conf. Human Language Technology*, San Diego, CA, Morgan Kaufmann Publishers, pp. 303-306, March 24-27, 2002.

S. Sista, R. Schwartz, T. Leek, and J. Makhoul, "An Algorithm for Unsupervised Topic Discovery from Broadcast News Stories," *Int. Conf. Human Language Technology*, San Diego, CA, Morgan Kaufmann Publishers, pp. 110-114, 2002.

W. Andrews, M. Kohler, and J. Campbell, "Phonetic Speaker Recognition," *EuroSpeech*, Aalborg, Denmark, 2001.

H. Jin, R. Schwartz, S. Sista, F. Walls,"Topic Tracking for Radio, TV Broadcast, and Newswire," *Eurospeech*, Budapest, Hungary, pp. 2439-2442, 1999.

H. Jin, S. Matsoukas, R. Schwartz, and F Kubala, "Fast Robust Inverse Transform SAT and Multi-stage Adaptation", *DARPA Broadcast News Transcription and Understanding Workshop*, Leesburg VA, 1998.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

H. Jin, F. Kubala, and R. Schwartz, "Automatic speaker clustering," *DARPA Speech Recognition Workshop*, Chantilly VA, 1997.

S. Matsoukas, R. Schwartz, H. Jin, and L. Nguyen, "Practical Implementations of Speaker Adaptive Training", *DARPA Speech Recognition Workshop*, Chantilly VA, 1997.

T. Anastasakos, J. McDonough, and J. Makhoul, "Speaker Adaptive Training: A Maximum Likelihood Approach to Speaker Normalization", *ICASSP*, Washington DC, 1997.

S. Matsoukas, R. Schwartz, H. Jin, and L. Nguyen, "Practical Implementations of Speaker-Adaptive Training", *IEEE Recognition Workshop*, Chantilly, Virginia, pp 133-138, 1997.

A. Anastasakos, J. Mcdonough, R. Schwartz, and J. Makhoul, "A Compact Model for Speaker-Adaptive Training," *ICSLP*, 1996

R. Schwartz, A. Derr, A. Wilgus, and J. Makhoul, "Speaker Verification IR&D Final Report," BBN Technical Report No. 6661, Nov. 1987.

H. Gish, M. Krasner, W. Russell, and J. Wolf, "Methods and experiments for text-independent speaker recognition over telephone channels," *IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 865-868, 1986.

V. Viswanathan, C.M. Henry, R. Schwartz, and S. Roucos, "Evaluation of Multisensor Speech Input for Speech Recognition in High Ambient Noise," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tokyo, Japan, pp. 85-88, 1986.

S. Roucos, V. Viswanathan, C. Henry, and R. Schwartz, "Word Recognition Using Multisensor Speech Input in High Ambient Noise," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tokyo, Japan, pp. 737-740, 1986.

H. Gish, K. Karnofsky, M. Krasner, S. Roucos, R. Schwartz, and J. Wolf, "Investigation of Text-Independent Speaker Identification Over Telephone Channels," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Tampa, FL, pp. 379-382, 1985.

M. Krasner, J. Wolf, K. Karnofsky, R. Schwartz, S. Roucos, and H. Gish, "Investigation of text-independent speaker identification techniques under conditions of variable data," *IEEE Int. Conf. Acoust., Speech, Signal Processing*, pp. 18B.5.1-18B.5.4, 1984.

J. Wolf, M. Krasner, K. Karnofsky, R. Schwartz, and S. Roucos, "Further Investigations of Probabilistic Methods for Text-Independent Speaker Identification", *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Boston, MA, pp. 551-554, 1983.

R. Schwartz, S. Roucos, and M. Berouti, "The Application of Probability Density Estimation to Text-Independent Speaker Identification," *ICASSP*, Paris, France, pp. 1649-1652, May 1982.

M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of Speech Corrupted by Acoustic Noise," *ICASSP*, Washington, DC, pp. 208-211, 1979. Also in J.Lim (Ed.), *Speech Enhancement*, New Jersey; Prentice-Hall, 1983.

8.7.2 Johns Hopkins University

M. Athineos, H. Hermansky, and D. P. W. Ellis, "LP-TRAP: Linear Predictive Temporal Patterns," *Proc. ICSLP*, Jeju, S. Korea, pp. 1154-1157, 2004.

S. Ganapathy,S. Thomas and H. Hermansky, "Temporal Envelope Subtraction for Robust Speech Recognition Using Modulation Spectrum", *Proc. ASRU*, Merano, Italy,2009.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

S. Ganapathy, S. Thomas and H. Hermansky, "Robust Spectro-temporal Features based on Autoregressive Models of Hilbert Envelopes", *Proc. ICASSP 2010*, Dallas, TX, 2010.

S. Thomas, S. Ganapathy and H. Hermansky, "Tandem Representations of Spectral Envelope and Modulation Frequency Features for ASR", *Proc. Interspeech*, Brighton, UK, 2009.

S. Thomas, S. Ganapathy and H. Hermansky, "Recognition of Reverberant Speech Using Frequency Domain Linear Prediction", *IEEE Signal Processing Letters*, 2008.

H. Hermansky, "Speculations on Knowledge Versus Data in ASR," *Proc. ESCA-NATO Tutorial and Research Workshop on Robust Speech Recognition for Unknown Communication Channels*, Pont-a-Mousson, France, 1997.

F. Valente and H. Hermansky, "Discriminant Linear Processing of Time-Frequency Plane," *Proc. International Conference on Spoken Language Processing*, Pittsburgh, PA, 2006.

H. Hermansky and D.P.W. Ellis and S. Sharma, "Connectionist Feature Extraction for Conventional HMM Systems", *Proc. ICASSP '00*, 2000.

H. Hermansky and N. Morgan, "RASTA Processing of Speech", *IEEE Transactions on Speech and Audio Processing*, Vol. 2, No. 4, pp. 587-589, 1994.

H. Hermansky, "Robust Speech Recognition," in Cole, Hirshman et al., "The Challenge of Spoken Language Systems, Research Directions for the Nineties," *IEEE Transactions on Speech and Audio Processing*, Vol. 3, No. 1, pp. 1-21, 1995.

H. Hermansky and S. Sharma, "TRAPS - Classifiers of Temporal Patterns," *Proc. International Conference on Spoken Language Processing*, Sydney, Australia, 1998.

H. Hermansky and S. Sharma, "Temporal Patterns (TRAPS) in ASR of Noisy Speech," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Phoenix, AZ 1999.

H. Hermansky and P. Fousek, "Multi-resolution RASTA filtering for TANDEM-based ASR," *Proc. European Conference on Speech Communication and Technology*, Lisbon, Portugal, 2005.

N. Mesgarani, G.S.V.S. Sivaram, S. Nemala, H. Hermansky, "Discriminant Spectrotemporal Features for Phoneme Recognition", *Proc. Interspeech 2009*, Brighton, UK, 2009.

G.S.V.S. Sivaram, S. Nemala, M.Elhilali, T. Tran, and H. Hermansky, "Sparse Coding for Speech Recognition", *ICASSP 2010*, Dallas, TX, 2010.

J. Pinto, Y. Bayya, H. Hermansky, M. Magimai-Doss, "Exploiting Contextual Information For Improved Phoneme Recognition", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.

F. Valente and H. Hermansky, "Hierarchical and Parallel Processing of Modulation Spectrum for ASR applications", *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP)*, 2008.

H. Fletcher and R.H. Galt, "The perception of speech and its relation to telephony", *The Journal of the Acoustical Society of America*, 1950.

A. Boothroyd and S. Nittrouer, "Mathematical treatment of context effects in phoneme and word recognition", *Journal of the Acoustical Society of America*, Vol 84(1), pp. 101-114, July 2008.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

H. McGurk and J. MacDonald, "Hearing lips and seeing voices", *Nature 264*, pp. 746 – 748, December 1976.

F. Valente and H. Hermansky, "Combination of Acoustic Classifiers based on Dempster-Shafer Theory of Evidence," *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Honolulu, HI, 2007.

8.7.3 University of Maryland

S. Shamma, "Encoding sound timbre in the auditory system" *IETE J. Res.* 49(2), 193-205, 2003.

L. Atlas and S. Shamma, "Joint acoustic and modulation frequency", *EURASIP J. Appl. Sig. Proc. 7*, Jeju, S. Korea, pp. 668-675, 2003.

P. Julián, A. Andreou, L. Riddle, S. Shamma, D. Goldberg, and G. Cauwenberghs. "A Comparative Study of Sound Localization Algorithms for Energy Aware Sensor Network Nodes," *IEEE Trans. Cir. Sys.*, 51(4), 2004

A. Schai, S Shamma "A Neuromorphic Sound Localizer for a Smart MEMS System," *Analog Integrated Circuits and Signal Processing*, 39, 267–273, 2004

N. Massgarani, M. Slaney, and S. Shamma, "Content-Based Audio Classification Based On Multiscale Spectro-Temporal Features", *IEEE Trans. Audio and Speech*, 2005.

N. Massgarani, M. Slaney, and S. Shamma, "Content-Based Audio Classification Based On Multiscale Spectro-Temporal Features", *IEEE Trans. Audio and Speech,* 2005.

D. Zotkin, S. Shamma, P. Ru, and R. Duraiswami, L. Davis "Neuromimetic sound representation for percept detection and manipulation," *EURASIP J. Applied Signal Processing*, 2005.

D. Zotkin, S. Shamma, P. Ru, and R. Duraiswami, L. Davis "Neuromimetic sound representation for percept detection and manipulation," *EURASIP J. Applied Signal Processing*, 2005.

J. Fritz, S. Shamma, M. Elhiliali, "Active Listening: Task-dependent plasticity of receptive fields in primary auditory cortex", *Hearing Res*, 206, 159-176, 2005.

T. Chi, P. Ru, and S. Shamma, "Multiresolution spectrotemporal analysis of complex sounds", *J. Acoust. Soc. Am*, 118, pp.887, 2006.

T. Chi and S. Shamma, "Spectrum restoration from multiscale auditory phase singularities by generalized projections", *IEEE Trans. Audio and Speech*, 14(4), 118, 887 2006.

J. Fritz, M. Elhilali, S. David, S Shamma "Auditory Attention – focusing the search light on sound", *Curr. Opin. Neurobiol*, 17(4):437-55, 2007

M. Elhilali and S. Shamma "The cocktail party problem-with a cortical twist", *J. Acoust. Soc. Am.*, 124(6), 2008

S. Shamma, "On the emergence and awareness of auditory objects," *PLoS biology*, 6(6) 2008

S. Shamma, "Characterizing Auditory Receptive Fields", *Neuron*, 58(6), p.829-831, 2008

M. Elhilali, J Xiang, S Shamma , and J Simon "Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene," *PLoS* 7(6) 2009.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

N. Mesgarani, J Fritz, and S Shamma, "A Computational Model of Rapid Task-Related Plasticity of Auditory Cortical Receptive Fields," *J. Comp. Neuroscience*, 2009.

8.7.4 Cambridge University

F. Flego and M.J.F. Gales, "Discriminative Adaptive Training with VTS and JUD," *Proc. IEEE Automatic Speech Recognition and Understanding Workshop*, December 2009.

H. Xu, M.J.F. Gales, and K.K. Chin, "Improving Joint Uncertainty Decoding Performance by Predictive Methods for Noise Robust Speech Recognition," *Proc. IEEE Automatic Speech Recognition and Understanding Workshop*, December 2009.

M.J.F. Gales, "Maximum Likelihood Linear Transformations for HMM-based Speech Recognition," *Computer Speech and Language*, Volume 12, 1998.

D.K. Kim and M.J.F. Gales, "Noisy Constrained Maximum Likelihood Linear Regression for Noise-Robust Speech Recognition," submitted to *IEEE Transactions on Audio Speech and Language Processing*.

M. Tomalin, J. Park, F. Diehl, M.J.F. Gales, and P.C. Woodland, "Recent Improvements to the Cambridge Arabic Speech-to-Text Systems," *ICASSP 2010*, Dallas, TX, 2010.

C. Longworth and M.J.F. Gales, "Combining Derivative and Parametric Kernels for Speaker Verification," *IEEE Transactions on Audio Speech and Language Processing*, May 2009.

M.J.F. Gales and F. Flego, "Discriminative Classifiers with Adaptive Kernels for Noise Robust Speech Recognition," to appear *Computer Speech and Language*, 2010.

C.J. Leggetter and P.C. Woodland, "Maximum likelihood linear regression for speaker adaptation of continuous density Hidden Markov models," *Computer speech and language*, 1995.

L.F. Uebel and PC Woodland, "Speaker adaptation using lattice-based MLLR," *ITRW on Adaptation Methods for Speech Recognition*, 2001.

L. Wang and P.C. Woodland, "MPE-based discriminative linear transforms for speaker adaptation," *Computer Speech & Language*, Vol. 22, No. 3, 2008.

K. Yu, M. Gales and P.C. Woodland, "Unsupervised adaptation with discriminative mapping transforms," *IEEE Trans. ASLP*, Vol. 17, No. 4, 2009.

D. Povey and P.C Woodland, "Minimum phone error and I-smoothing for improved discriminative training," *Proc. ICASSP*, Orlando, FL, 2002.

R. Cordoba, P.C. Woodland, and M. Gales, "Improved cross-task recognition using MMIE training," *Proc. ICASSP*, Orlando, FL, 2002.

M. Tomalin, J. Park, F. Diehl, M.J.F Gales and P.C. Woodland "Recent improvements to the Cambridge Arabic speech-to-text systems" *Proc. ICASSP'2010*, Dallas, TX, 2010.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

8.7.5 Brno University of Technology

N. Brümmer, L. Burget, J. Černocký, O. Glembek, F. Grézl, M. Karafiát, D. van Leeuwen, P. Matějka, P. Schwarz, and A. Strasheim, "Fusion of heterogeneous speaker recognition systems in the STBU submission for the NIST speaker recognition evaluation 2006," *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 7, pp. 2072-2084, 2007.

N. Brümmer, A. Strasheim, V. Hubeika, P. Matějka, L. Burget, and O. Glembek, "Discriminative Acoustic Language Recognition via Channel-Compensated GMM Statistics," *Proc. Interspeech*, Brighton, UK, pp. 2187-2190, 2009.

L. Burget, P. Matějka, and J. Černocký, "Discriminative Training Techniques for Acoustic Language Identification", *Proc. ICASSP*, Toulouse, FR, pp. 209-212, 2006.

L. Burget, P. Matějka, P. Schwarz, O. Glembek, and J. Černocký, "Analysis of feature extraction and channel compensation in GMM speaker recognition system", *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 7, pp. 1979-1986 2007.

L. Burget, N. Brümmer, D. Reynolds, P. Kenny, J. Pelecanos, R. Vogt, F. Castaldo, J. Dehak, R. Dehak, O. Glembek, Z Karam, J. Noecker Jr., Y. Na Hye, C. Costin Ciprian, V. Hubeika, S. Kajarekar, N. Scheffer, and J. Černocký, "Robust Speaker Recognition Over Varying Channels", *JHU* , Baltimore, MD, p. 81, 2008.

L. Burget, P. Matějka, V. Hubeika, and J. Černocký, "Investigation into variants of Joint Factor Analysis for speaker recognition", *Proc. Interspeech*, Brighton, UK, pp. 1263-1266, 2009.

L. Burget, M. Fapšo, V. Hubeika, O. Glembe, M. Karafiát, M. Kockmann, P. Matějka, P. Schwarz, and J. Černocký, "BUT system for NIST 2008 speaker recognition evaluation", *Proc. Interspeech*, Brighton, UK, pp. 2335-2338, 2009.

O. Glembek, L. Burget, N. Dehak, N. Brümmer and P. Kenny: "Comparison of scoring methods used in speaker recognition with joint factor analysis", accepted to *ICASSP*, 2009.

Z. Jancik, O. Plchot, N. Brümmer, L. Burget, O. Glembek, V. Hubeika, M. Karafiat, P. Matejka, T. Mikolov, A. Strasheim and J. Černocký, "Data selection and calibration issues in automatic language recognition", submitted to *Odyssey 2010*, Brno, 2010.

M. Kockmann, L. Burget, "Contour modeling of prosodic and acoustic features for speaker recognition", *Proc. IEEE Workshop on Spoken Language Technology*, p. 4, 2008.

P. Matějka, P. Schwarz, L. Burget, and J. Černocký, "Use of anti-models to further improve state-of-the-art PRLM language recognition system", *Proc. ICASSP*, Toulouse, FR, pp. 197-200, 2006.

P. Matějka, L. Burget,, P. Schwarz, O. Glembek, M. Karafiát, F. Grézl, J. Černocký, D. van Leeuwen, N. Brümmer, A. Strasheim, "STBU system for the NIST 2006 speaker recognition evaluation", *Proc. IEEE International Conference on Acoustics, Speech and Signal Processing*, Honolulu, HI, pp. 221-224, 2007.

P. Matejka, L. Burget, O. Glembek, P. Schwarz, V. Hubeika, M. Fapso, T. Mikolov, and O. Plchot, "BUT system description for NIST LRE 2007", *NIST LRE 2007 Workshop*, Orlando, FL, 2007.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

T. Mikolov, O. Plchot, O. Glembek, P. Matějka, L. Burget, and J. Černocký, "PCA-based Feature Extraction for Phonotactic Language Recognition", submitted to *Odyssey 2010, Brno*, 2010.

P. Oldrich, V. Hubeika, L. Burget, P. Schwarz, P. Matějka, and J. Černocký, "Acquisition of Telephone Data from Radio Broadcasts with Applications to Language Recognition: Technical Report", *technical report*, BUT, January 2009.

P. Schwarz, "Phoneme recognition based on long temporal context", *PhD Thesis, Brno*, BUT, 2009.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

9 Project Management and Interaction Plan

9.1 Project Management

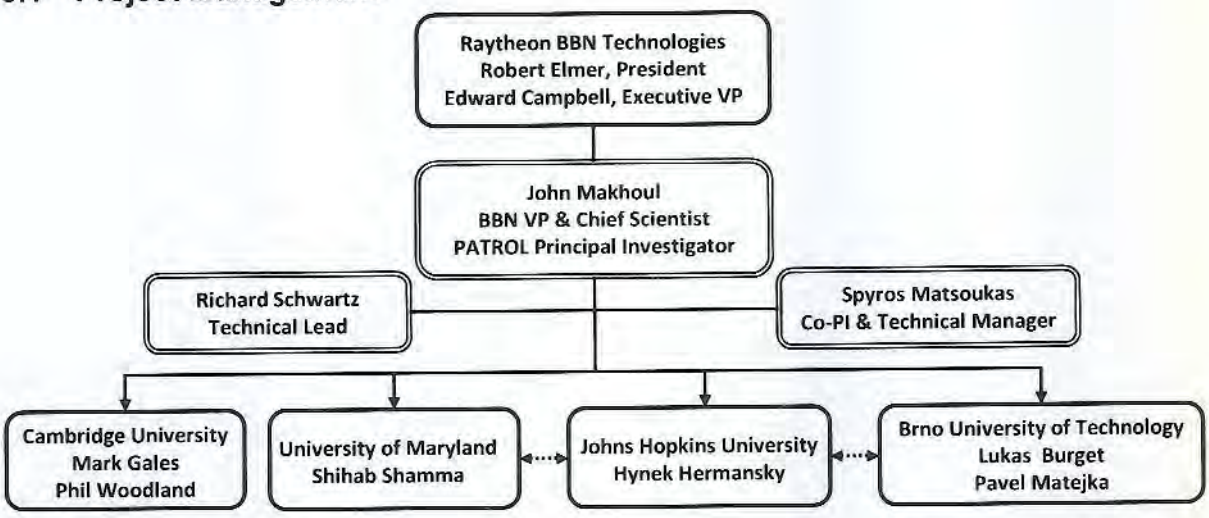


Figure 9-1 Organizational chart for the PATROL Team.

Figure 9-1 shows the management organization for the project. Dr. John Makhoul will serve as Principal Investigator and the overall Program Manager for this effort; he will coordinate and manage the activities of the BBN-led PATROL Team in the RATS Program. Dr. Makhoul is an experienced project leader and program manager. He has served as PI on a large number of DARPA and other government projects. Most recently, he managed and coordinated the BBN-led team under the DARPA EARS program and is currently the PI on the BBN-led team under the DARPA GALE program. In both programs, the BBN-led teams have been consistently top performers.

Under the BBN GALE effort, Dr. Makhoul oversees the activities of a dozen subcontractors, distributed geographically across multiple locations, from California to Europe. The work under the GALE program is quite diverse; it includes work on speech recognition, machine translation, distillation, the linguistic annotation of text, and the integration of technologies into operational systems that take speech or text from multiple languages and produce an English translation. The resulting technologies have been transitioned into BBN's Broadcast Monitoring and Web Monitoring products which have been deployed in at least 18 Government locations for 24/7 operation. The GALE program is expected to end in April 2011, but our experience in running such a large and successful project will serve us well in managing the BBN-led PATROL Team in the RATS program.

Dr. Makhoul will be assisted at BBN by two senior researchers, both of whom have had prominent and important roles in the EARS and GALES programs. Spyros Matsoukas, who has demonstrated in the GALES program first-rate skills in performing technical management across many sites, will serve as co-PI and Technical Manager. He will manage the technical activities at BBN and across the other four sites on the PATROL team. Rich Schwartz, who has been a fountain of innovation in various aspects of speech processing and pattern recognition for over three decades, will serve as Technical Lead for the effort.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Having been on the BBN-led team on GALE for the last four and a half years, Professors Mark Gales and Phil Woodland of Cambridge University have worked closely with Makhoul, Matsoukas, and Schwartz, and together have developed one of the most advanced multi-lingual speech recognition systems in the world. That strong working relationship will continue during the RATS project.

Professor Hynek Hermansky of Johns Hopkins University (JHU), who has an ongoing joint appointment with the Brno University of Technology, has a very strong relationship with Drs. Lukas Burget and Pavel Matejka there, both having performed research with him in the past and have attended summer workshops at JHU. Currently, Burget and Matejka serve as consultants on one of Hermansky's projects in the US.

Professor Shihab Shamma of the University of Maryland (UMD) currently has a joint project on speaker identification with Professor Hermansky, sponsored by IARPA under the BEST program. Furthermore, both of them work on auditory modeling of sound and will be able to join up to provide the PATROL team unparalleled expertise in that area, which we expect to be very important in isolating the speech part of the noisy speech signal. Shamma and Hermansky will be providing auditory-based feature extraction software to the other team members for use in building systems for the four RATS applications.

BBN and UMD have also had a very strong working relationship under the GALE program. Rich Schwartz from BBN, who is physically located in Virginia, has supervised the work of several students from UMD. Schwartz has been meeting with the students at the BBN office in Columbia, MD, on a weekly basis. Under RATS, Schwartz will continue his relationship with UMD and his interaction with the students there.

The above described working relationships that already exist among the PATROL team members will be a very good starting point for establishing a multilateral, open, and cooperative working relationship among all team members.

9.2 Team Size and Composition

From our experience in the EARS and GALE programs, we have found that, in programs with challenging goals, it is important to have the most qualified people at multiple sites work together, bounce off ideas against each other, and collaborate in implementing and integrating ideas and solutions that lead to the best performance and accuracy. By any measure, the goals of the RATS program are very challenging and, therefore, require diverse ideas from a team that has the requisite expertise and proven record of innovation and superior performance. We believe we have assembled such a team. Going with the old adage of "two heads are better than one", we have endeavored to ensure that, in each primary technical area, we have at least two sites that focus on each area of research.

Table 9-1 lists the five sites on the PATROL team and the areas of primary concentration in the proposed effort. In the area of robust feature extraction and speech enhancement, Dr. Hermansky at JHU and Dr. Shamma at UMD have each developed audio processing methods that tend to separate the speech from the noise. They will make their methods and software available to the PATROL team in order to see which features work best with large levels of noise.

Even with the use of robust features, there will still be the need for noise modeling. BBN and Cambridge University are both very experienced in this area and will continue to innovate

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

solutions to this thorny problem. Brno University of Technology (BUT) will use their joint factor analysis (JFA) work to render the ID tasks more immune to noise.

	Robust Feature Extraction/ Enhancement	Noise Modeling	Speech Activity Detection	Language ID	Speaker ID	Keyword Spotting	Software Integration
BBN		X	X	X	X	X	X
Johns Hopkins U	X				X		
U Maryland	X		X				
Cambridge U		X				X	
Brno U Tech		X		X	X		

Table 9-1 Organizations of the PATROL team and their primary areas of concentration in RATS.

Dr. Shamma at UMD has obtained very good results in performing speech activity detection (SAD) in various types of noise by using his cortical representation of sound. He will continue to work on this problem while BBN will use advanced statistical methods to perform SAD.

BBN and BUT will work on language ID and speaker ID, both of which are the primary areas of strength at BUT. In addition, Dr. Hermansky will work on speaker ID, especially that he currently is working on that problem on an IARPA effort.

BBN is a recognized leader in the area of keyword spotting and has delivered systems to the Government for keyword spotting in various languages. Cambridge University – well known for their expertise in speech recognition – will work with BBN on keyword spotting.

Sites would be expected to deliver to BBN working software modules (a task that is not shown explicitly in the above table). BBN will need to integrate all this software into a working system that can be delivered to the Government. Naturally, BBN will be the only site performing integration and delivery.

Even though each site will concentrate on the tasks listed in Table 9-1, the sites, through joint meetings, will also be available to consult on the other tasks, as their expertise permits.

As one can see from the above, BBN has built a team with some built-in redundancy, but with maximum effectiveness. We strongly believe that this team, both in size and composition, is both necessary and sufficient to meet the program objectives.

9.3 Working/Meeting Models

BBN has established teaming relationships with each of the PATROL team members. The Statement of Work for each team member contains specific technical requirements by task, project management reporting responsibilities, anticipated support for meetings and presentations, and the delivery of software. The teaming strategy draws upon unique capabilities of each participant to form a comprehensive execution plan for the RATS effort.

BBN plans to have a face-to-face kickoff meeting early in the project. The kick-off meeting will concentrate on working out an overall plan to be accomplished during Phase 1 of the project, and a timeline for executing the plan. After the kickoff meeting, most of the day-to-day coordination of activities will take place through regular weekly teleconferences for each of the major tasks. The technical reviews for each task will include coordination of activities across sites, planning for the research to be performed, the assignment of specific tasks, and the review of progress against the work plan. In addition to the frequent teleconferences, periodic face-to-face meetings will be scheduled at various locations, taking advantage of major conferences and DARPA

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

meetings. Such team meetings will strengthen the ties among team members and allow for in-depth technical discussions and planning. In addition to the DARPA requested meetings and workshops specified upon award of the contract, meetings and teleconferences with the DARPA RATS Program Manager will be scheduled as needed.

9.4 Software/Code Management

9.4.1 Software Integration/Management

BBN is highly experienced in the integration of technology from other sites into functioning systems, with appropriate GUIs, which have been delivered to the Government, either for internal use or for deployment in the field for 24/7 operations. Our expertise includes developing COTS Human Language Technology (HLT) software systems, such as the Broadcast Monitoring System (BMS), which integrates software from two other sites; it operates with minimal human oversight at over 18 Government locations. In addition to commercial products, BBN delivers state-of-the-art HLT software suites directly to Government customers for their use in research and operations. One such system is Byblos, BBN's state-of-the-art, *trainable* speech-to-text (STT) system. Byblos is used by Government staff to run STT and to create new STT models using Byblos tools to train on noisy, real-world telephone data in a variety of languages. BBN also delivers keyword search (KWS) systems to the Government, including word and phone-based systems that operate at very high speeds. The KWS software suites are designed for use in larger Government systems, so they are callable via an API jointly designed by the Government and BBN. BBN has extensive experience supporting the Government customer to help them make STT, KWS, and other HLT work in a wide variety of languages and domains. We have also delivered a GUI-driven *trainable* OCR system to the Government (the OCR system is based on the BBN Byblos STT system).

In the RATS project, the integration of the PATROL software system will be performed at BBN. Well before integration, BBN, in consultation with the team members, will create APIs that clearly define interfaces of the modules each site is developing. Early integration of prototype modules in Phase 1 will take place to test API soundness and ensure successful final integration for delivery to the Evaluation Team. Our teammates will contribute software in the form of source code, enabling BBN to create a single tightly integrated software suite. Having the source code in a centralized location will also allow us to rapidly debug and fix any issues discovered during integration. Whenever sites submit new versions of software modules, they will also submit test programs that exercise API functionality and the results of test runs demonstrating the module's conformance to the defined API.

All source code will be checked into a central repository at BBN dedicated to the RATS project. This repository will be used for integration of all subcontractor components into the complete system. In addition, this repository will support the usual revision control functions necessary for BBN staff to efficiently conduct research exploring alternative software algorithms. Specifically, BBN researchers will check in to the repository all changes to the research code base, typically checking into a private branch of the repository until a change has been shown to be beneficial experimentally, at which time the change is merged into the repository's "main trunk." The RATS project will adopt a number of software practices that BBN currently uses to maintain a consistent software repository. This includes creation of a symbolically tagged, date-stamped "local release" every time the trunk of the repository is updated. Software will also be tagged, built, and released whenever new software contributions are received from subcontractors. The symbolic tags ensure our ability to analyze and identify the source of any

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

unexpected changes in system behavior. Other software practices we will adopt include the use of automated nightly builds and regression tests, with emails automatically sent when any error or warning occurs in the process, along with information about the specific module and originating site. We will then work with that site to fix any problems.

BBN's software system will also adhere to the APIs and input and output formatting conventions defined by the Evaluation Team for periodic evaluation deliveries and by military end users for the trainable system deliveries. BBN will test all software components extensively prior to delivery to ensure robustness and conformance to external APIs.

BBN will support the Evaluation Team by providing access to all developed system and component software; addressing software instrumentation needs; identifying failure causes; meeting required delivery dates to support evaluation events; and responding to any other test-related issues. BBN's extensive experience supporting Byblos and other Human Language Technology systems at remote customer sites gives us the confidence to ensure that the Evaluation Team can successfully run our RATS system.

9.4.2 Secure Processing

The work under Technical Area 1 of the RATS program does not require the processing of any classified data. However, the Evaluation Team is slated to test the delivered systems on classified data, possibly at the SCI level. In the event that the Evaluation Team should meet with any issues when running the PATROL system on classified data, BBN has numerous staff in its speech processing group who are cleared at the SCI level and who will be able to interact with the Evaluation Team or any future military end users to ensure that the software suite is correctly configured and runs successfully on their data. In addition, BBN has a TS/SCI SCIF equipped with a secure connection to TS/SCI Government networks and, therefore, will be able to process the same classified data in its SCIF, should the need arise to do that.

9.5 University Participation

In this project, we are fortunate to have as partners four universities that have excelled in their areas of expertise which are relevant to the RATS program, as mentioned above. The Key Personnel professors in those universities will be supervising a mix of research staff, including post-docs, and graduate students who will be participating as research assistants. The staff and students, under the guidance of their professors, will be writing software and performing experiments to improve accuracy in the four RATS applications in noise.

In addition to developing new research ideas and writing code, these universities will be expected to deliver to BBN software that will be integrated into the BBN system for delivery to the Evaluation Team and to potential users. Under an existing IARPA contract, JHU, UMD, and BUT will be delivering software for an external evaluation team as well. So, some of their software will have already undergone hardening for delivery to third parties. Furthermore, BUT has a lot of experience in transitioning their software for government and commercial applications. We plan to take advantage of all this experience to integrate the software from these sites with ours for delivery. The speech group at Cambridge University has a long history of building solid software (especially their HTK speech recognition toolkit) for external use.

For over 20 years, BBN has had a cooperative relationship with Northeastern University, where Dr. Makhoul has an appointment as Adjunct Professor. In this relationship, graduate students from Northeastern work as research assistants and do their graduate thesis work at BBN,

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

supervised by senior BBN staff, with Dr. Makhoul acting as their official thesis advisor. We expect to continue this arrangement, so we have budgeted one research assistant from Northeastern to work on this project.

9.6 Government's Role

We expect that the Government will provide us with the necessary data with which to build our models for the four application areas. We will work closely with the Program Manager to ensure that our work is well aligned with the Government's expectations.

9.7 Project Management at BBN

BBN has an extensive history of program management success that is predicated on tailoring the management approach to the needs of the project. Our project activities range from very large, formal programs to small, short-term projects with only a few engineers. Our project activities range from research and development to technical transition to commercial offerings. We have the flexibility and trained talent within BBN to work at all these levels. BBN is dedicated to the principle that program management must take a balanced and comprehensive approach to the achievement of technical, cost, and schedule objectives. To meet these objectives, BBN has developed a program organization and management approach proven effective by prior successful contract performances with multiple team members.

Every project has a designated program manager who is responsible for the overall management leading to the successful completion within time and budget. The program manager is responsible for day-to-day management, staffing plans, financial tracking and reporting, subcontractor relationships, and tracking of milestones and deliverables. Every project is required to have a project plan detailing that project's tasks, work organization, and budgeted costs. The plan, schedule, assignments and budgets are communicated to all program team members. Monthly tracking and control procedures measure the progress and cost against the plan. Each project is subjected to monthly reviews that report on project status, so that any difficulties or risks that may impact project completion can be identified and addressed in a timely manner.

The team holds regular project meetings to coordinate work, identify any potential problems, and ensure the project is proceeding on schedule. Program risks will be identified and discussed, and mitigation plans implemented. In a program of the magnitude of RATS the team meetings are required at both the program and task level.

A project web site is established to store, retrieve, and share technical and management documents. All work products undergo an appropriate level of peer review. We provide status reports prepared and submitted in accordance with procedures in the award document. At the end of each phase, we prepare and submit a Final Report that summarizes the project and tasks at the end of that phase.

BBN has the infrastructure to support the management of large government contracts. A number of systems (such as BBN's time reporting system, Microsoft Project for schedule creation and management, and PRISM for project finances) are in place to monitor project status and facilitate project control. Most importantly, we have customizable project management practices. We understand and appreciate the need to be flexible, efficient, and highly responsive to the evolving goals of the RATS Program.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

10 Cost Summaries

The following four pages provide a complete cost summary for the proposed PATROL effort under the RATS program.

The first page contains a top-level view of the proposed costs, broken down by phase, major task, and performing site. There are seven major technical tasks: Feature Extraction, Noise Modeling/Compensation, Speech Activity Detection, Language Identification, Speaker Identification, Keyword Spotting, and System Integration. In addition to the seven major technical tasks, we have also priced Program Management and Program Infrastructure as additional major tasks. Program Management includes program administration, all BBN travel, and data services (such as low-level annotations, which will be made available to the other RATS performing sites). Program Infrastructure includes materials (computer equipment), computer system administration, and data acquisition. For Phases 2 and 3, we also show at the bottom of the page the costing for the proposed Speech Transcription option.

The second page contains the same costing information for Phase 1, but broken down by month, for the 18-month duration. The third and fourth pages contain the corresponding monthly costing information for Phases 2 and 3, each lasting 12 months.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

RATS PATROL SUMMARY

Base Proposal	Phase 1	Phase 2	Phase 3
Feature Extraction			
Johns Hopkins University	$ 191,501	$ 128,763	$ 131,141
University of Maryland	$ 129,261	$ 95,315	$ 94,500
Noise Modeling/Compensation			
BBN	$ 664,876	$ 442,559	$ 441,829
Brno University	$ 51,520	$ 38,011	$ 37,687
Cambridge University	$ 134,649	$ 54,900	$ 55,978
Speech Activity Detection (SAD)			
BBN	$ 475,039	$ 276,875	$ 243,125
University of Maryland	$ 129,261	$ 95,315	$ 94,500
Language ID (LID)			
BBN	$ 644,170	$ 461,689	$ 495,500
Brno University	$ 103,040	$ 76,036	$ 75,372
Speaker ID (SID)			
BBN	$ 486,644	$ 398,574	$ 419,302
Brno University	$ 103,040	$ 76,036	$ 75,372
Johns Hopkins University	$ 191,501	$ 128,763	$ 131,140
Keyword Spotting (KWS)			
BBN	$ 616,423	$ 461,364	$ 483,307
Cambridge University	$ 134,650	$ 54,904	$ 55,978
Northeastern University	$ 43,063	$ 44,782	$ 46,573
System Integration			
BBN	$ 313,267	$ 342,757	$ 377,522
Program Management *			
BBN	$ 328,681	$ 247,608	$ 234,355
Program Infrastructure **			
BBN	$ 552,786	$ 470,811	$ 400,930
Total Base	**$ 5,293,372**	**$ 3,895,062**	**$ 3,894,111**
Speech Transcription Option			
BBN		$ 325,701	$ 337,100
Cambridge University		$ 157,084	$ 162,459
Total Option		**$ 482,785**	**$ 499,559**

* Program Management includes program administration, travel and data services.

** Program Infrastructure includes system administration, material, and data acquisition.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Phase 1 by Month

Phase 1 Base	Aug-10	Sep-10	Oct-10	Nov-10	Dec-10	Jan-11	Feb-11	Mar-11	Apr-11	May-11	Jun-11	Jul-11	Aug-11	Sep-11	Oct-11	Nov-11	Dec-11	Jan-12	Feb-12	Total
Feature Extraction																				
Johns Hopkins University	$ 5,341	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,581	$ 10,413	$ 10,412	$ 5,206	**$ 191,501**
University of Maryland	$ 3,588	$ 7,182	$ 7,182	$ 7,182	$ 7,182	$ 7,180	$ 7,182	$ 7,182	$ 7,182	$ 7,181	$ 7,181	$ 7,182	$ 7,181	$ 7,181	$ 7,181	$ 7,181	$ 7,182	$ 7,182	$ 3,587	**$ 129,261**
Noise Modeling/Compensation																				
BBN	$ 14,523	$ 30,623	$ 30,707	$ 30,707	$ 30,159	$ 39,235	$ 38,122	$ 38,427	$ 38,391	$ 36,348	$ 38,202	$ 37,214	$ 38,549	$ 37,101	$ 36,202	$ 38,146	$ 35,798	$ 40,396	$ 36,026	**$ 664,876**
Brno University							$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,294	$ 4,290	$ 4,290		**$ 51,520**
Cambridge University	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,086	$ 7,086	$ 7,086	$ 7,086	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	**$ 134,649**
Speech Activity Detection (SAD)																				
BBN	$ 11,264	$ 19,976	$ 20,855	$ 21,532	$ 21,532	$ 30,169	$ 29,800	$ 29,356	$ 29,026	$ 28,644	$ 28,666	$ 26,743	$ 24,889	$ 26,371	$ 25,860	$ 22,616	$ 24,335	$ 25,303	$ 28,101	**$ 475,038**
University of Maryland	$ 3,588	$ 7,182	$ 7,182	$ 7,182	$ 7,182	$ 7,180	$ 7,182	$ 7,182	$ 7,182	$ 7,181	$ 7,181	$ 7,182	$ 7,181	$ 7,181	$ 7,181	$ 7,181	$ 7,182	$ 7,182	$ 3,587	**$ 129,261**
Language ID (LID)																				
BBN	$ 15,246	$ 29,236	$ 29,057	$ 29,441	$ 29,479	$ 40,152	$ 38,798	$ 38,392	$ 36,002	$ 37,709	$ 35,452	$ 35,087	$ 35,002	$ 37,956	$ 37,338	$ 34,463	$ 33,061	$ 37,269	$ 34,030	**$ 644,170**
Brno University							$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,583	$ -	**$ 103,040**
Speaker ID (SID)																				
BBN	$ 13,565	$ 23,807	$ 24,100	$ 24,393	$ 24,535	$ 27,232	$ 26,620	$ 26,229	$ 28,679	$ 27,344	$ 26,574	$ 26,883	$ 27,838	$ 28,073	$ 27,379	$ 25,266	$ 24,187	$ 27,650	$ 26,290	**$ 486,644**
Brno University							$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,587	$ 8,583	$ -	**$ 103,040**
Johns Hopkins University	$ 5,341	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,682	$ 10,683	$ 10,581	$ 10,412	$ 10,412	$ 5,206	**$ 191,501**
Keyword Spotting (KWS)																				
BBN	$ 12,112	$ 29,376	$ 30,852	$ 30,252	$ 30,218	$ 34,581	$ 35,157	$ 33,786	$ 36,233	$ 35,536	$ 35,132	$ 35,497	$ 35,175	$ 34,924	$ 34,242	$ 34,154	$ 33,756	$ 34,460	$ 30,980	**$ 616,423**
Cambridge University	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,088	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,083	$ 7,087	$ 7,087	$ 7,087	$ 7,087	$ 7,087	**$ 134,650**
Northeastern University							$ 1,826	$ 1,826	$ 1,826	$ 1,826	$ 8,876	$ 8,876	$ 8,876	$ 1,826	$ 1,826	$ 1,826	$ 1,827	$ 1,826		**$ 43,063**
System Integration																				
BBN	$ 6,464	$ 13,777	$ 13,777	$ 13,777	$ 13,777	$ 20,329	$ 19,041	$ 18,367	$ 19,392	$ 17,871	$ 17,685	$ 17,359	$ 16,900	$ 18,291	$ 17,508	$ 17,157	$ 16,355	$ 18,484	$ 16,956	**$ 313,267**
Program Management																				
BBN	$ 19,031	$ 24,678	$ 12,546	$ 12,375	$ 12,071	$ 13,200	$ 12,930	$ 12,071	$ 17,290	$ 27,703	$ 25,211	$ 29,330	$ 15,766	$ 29,912	$ 12,020	$ 12,733	$ 16,993	$ 11,901	$ 10,920	**$ 328,681**
Program Infrastructure																				
BBN	$ 11,975	$ 12,161	$ 11,975	$228,757	$ 12,161	$ 11,975	$ 12,160	$ 11,976	$115,174	$ 12,562	$ 12,370	$ 12,371	$ 12,562	$ 12,370	$ 12,563	$ 12,370	$ 12,563	$ 12,370	$ 12,372	**$ 552,787**
Total Phase 1 Base	**$136,212**	**$233,536**	**$223,771**	**$441,136**	**$223,835**	**$266,771**	**$285,824**	**$281,800**	**$393,383**	**$296,911**	**$299,535**	**$300,729**	**$287,924**	**$298,187**	**$276,307**	**$269,897**	**$268,702**	**$280,477**	**$227,435**	**$5,293,372**

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Phase 2 by Month

Phase 2 Base	Feb-12	Mar-12	Apr-12	May-12	Jun-12	Jul-12	Aug-12	Sep-12	Oct-12	Nov-12	Dec-12	Jan-13	Feb-13	Total
Feature Extraction														
Johns Hopkins University	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,903	**$ 128,763**
University of Maryland	$ 4,039	$ 8,079	$ 8,079	$ 8,079	$ 7,977	$ 7,875	$ 7,875	$ 7,874	$ 7,875	$ 7,875	$ 7,875	$ 7,875	$ 3,938	**$ 95,315**
Noise Modeling/Compensation														
BBN	$ 34,140	$ 33,039	$ 34,196	$ 34,256	$ 35,020	$ 35,589	$ 35,832	$ 35,589	$ 34,196	$ 34,256	$ 34,196	$ 33,057	$ 29,192	**$ 442,558**
Brno University	$ 3,222	$ 3,222	$ 3,222	$ 3,222	$ 3,140	$ 3,141	$ 3,141	$ 3,141	$ 3,140	$ 3,140	$ 3,140	$ 3,140		**$ 38,011**
Cambridge University	$ 2,317	$ 4,635	$ 4,635	$ 4,635	$ 4,635	$ 4,635	$ 4,554	$ 4,519	$ 4,519	$ 4,519	$ 4,519	$ 4,519	$ 2,259	**$ 54,900**
Speech Activity														
BBN	$ 20,859	$ 21,188	$ 21,244	$ 22,243	$ 22,099	$ 21,716	$ 22,595	$ 22,128	$ 21,814	$ 21,959	$ 20,959	$ 20,124	$ 17,947	**$ 276,875**
University of Maryland	$ 4,039	$ 8,079	$ 8,079	$ 8,079	$ 7,976	$ 7,875	$ 7,875	$ 7,875	$ 7,875	$ 7,875	$ 7,875	$ 7,875	$ 3,938	**$ 95,315**
Language ID (LID)														
BBN	$ 33,190	$ 33,827	$ 34,352	$ 34,823	$ 36,090	$ 36,315	$ 37,156	$ 36,315	$ 35,836	$ 36,217	$ 35,491	$ 37,761	$ 34,316	**$ 461,689**
Brno University	$ 6,444	$ 6,444	$ 6,444	$ 6,444	$ 6,293	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281		**$ 76,036**
Speaker ID (SID)														
BBN	$ 29,956	$ 30,559	$ 30,720	$ 31,628	$ 31,289	$ 32,012	$ 31,603	$ 31,599	$ 31,289	$ 31,216	$ 30,720	$ 28,302	$ 27,679	**$ 398,572**
Brno University	$ 6,444	$ 6,444	$ 6,444	$ 6,444	$ 6,295	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281		**$ 76,038**
Johns Hopkins University	$ 9,905	$ 9,903	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	$ 9,905	**$ 128,763**
Keyword Spotting														
BBN	$ 34,624	$ 35,723	$ 36,315	$ 36,689	$ 36,660	$ 36,660	$ 35,478	$ 36,248	$ 35,836	$ 36,149	$ 35,492	$ 33,107	$ 32,383	**$ 461,364**
Cambridge University	$ 2,318	$ 4,635	$ 4,635	$ 4,636	$ 4,637	$ 4,637	$ 4,556	$ 4,519	$ 4,518	$ 4,518	$ 4,518	$ 4,518	$ 2,259	**$ 54,904**
Northeastern University	$ 1,899	$ 1,899	$ 1,899	$ 1,900	$ 9,229	$ 9,229	$ 9,230	$ 1,900	$ 1,899	$ 1,900	$ 1,899	$ 1,899		**$ 44,782**
System Integration														
BBN	$ 23,176	$ 22,803	$ 24,400	$ 23,703	$ 24,298	$ 28,507	$ 30,360	$ 28,507	$ 28,689	$ 29,251	$ 29,948	$ 24,364	$ 24,751	**$ 342,757**
Program Management														
BBN	$ 21,263	$ 28,165	$ 19,079	$ 14,460	$ 26,018	$ 30,577	$ 16,978	$ 20,653	$ 14,336	$ 14,048	$ 18,179	$ 11,663	$ 12,189	**$ 247,608**
Program Infrastructure														
BBN	$228,892	$ 12,110	$ 12,500	$ 12,499	$ 12,499	$ 12,499	$ 12,302	$105,017	$ 12,499	$ 12,497	$ 12,499	$ 12,500	$ 12,499	**$ 470,812**
Total **Phase 2** Base	**$476,632**	**$280,659**	**$276,053**	$273,550	$293,965	**$303,639**	$291,907	**$378,256**	$276,693	$277,792	$279,682	$263,076	**$223,158**	**$3,895,062**
Speech Transcription Option														
BBN	$ 24,153	$ 24,698	$ 24,492	$ 25,847	$ 24,998	$ 25,562	$ 24,492	$ 25,562	$ 24,998	$ 25,847	$ 24,492	$ 25,562	$ 24,998	**$ 325,701**
Cambridge University	$ 12,084	$ 12,083	$ 12,083	$ 12,083	$ 12,084	$ 12,084	$ 12,083	$ 12,084	$ 12,084	$ 12,083	$ 12,083	$ 12,083	$ 12,083	**$ 157,084**
Total Option	**$ 36,237**	**$ 36,781**	**$ 36,575**	**$ 37,930**	**$ 37,082**	**$ 37,646**	**$ 36,575**	**$ 37,646**	**$ 37,082**	**$ 37,930**	**$ 36,575**	**$ 37,645**	**$ 37,081**	$ 482,785

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Phase 3 by Month

Phase 3 Base	Feb-13	Mar-13	Apr-13	May-13	Jun-13	Jul-13	Aug-13	Sep-13	Oct-13	Nov-13	Dec-13	Jan-14	Feb-14	Total
Feature Extraction														
Johns Hopkins University	$ 10,087	$ 10,088	$ 10,087	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,087	**$ 131,141**
University of Maryland	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,270	$ 7,270	$ 7,270	$ 7,269	$ 7,269	$ 7,269	$ 7,269	**$ 94,500**
Noise Modeling/Compensation														
BBN	$ 33,219	$ 33,050	$ 36,019	$ 35,945	$ 37,967	$ 37,465	$ 38,067	$ 36,339	$ 35,235	$ 35,030	$ 34,382	$ 30,318	$ 18,793	**$ 441,829**
Brno University	$ 3,141	$ 3,141	$ 3,141	$ 3,141	$ 3,140	$ 3,140	$ 3,140	$ 3,140	$ 3,140	$ 3,141	$ 3,141	$ 3,141		**$ 37,687**
Cambridge University	$ 2,332	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,664	$ 2,332	**$ 55,978**
Speech Activity Detection (SAD)														
BBN	$ 18,525	$ 17,938	$ 19,823	$ 19,347	$ 20,058	$ 21,630	$ 20,233	$ 20,980	$ 19,146	$ 18,697	$ 19,174	$ 16,962	$ 10,612	**$ 243,125**
University of Maryland	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,269	$ 7,270	$ 7,270	$ 7,270	$ 7,269	$ 7,269	$ 7,269	**$ 94,500**
Language ID (LID)														
BBN	$ 37,082	$ 39,102	$ 40,033	$ 42,064	$ 40,750	$ 41,854	$ 38,053	$ 40,148	$ 40,750	$ 40,116	$ 38,085	$ 34,602	$ 22,861	**$ 495,500**
Brno University	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281		**$ 75,372**
Speaker ID (SID)														
BBN	$ 31,519	$ 31,006	$ 32,623	$ 32,222	$ 32,196	$ 32,517	$ 32,947	$ 32,517	$ 32,196	$ 32,222	$ 32,623	$ 32,091	$ 32,623	**$ 419,302**
Brno University	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281	$ 6,281		**$ 75,372**
Johns Hopkins University	$ 10,087	$ 10,088	$ 10,088	$ 10,088	$ 10,087	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,088	$ 10,087	$ 10,087	**$ 131,140**
Keyword Spotting (KWS)														
BBN	$ 35,766	$ 36,323	$ 37,506	$ 37,594	$ 37,018	$ 37,269	$ 37,831	$ 37,269	$ 37,018	$ 37,594	$ 37,506	$ 37,594	$ 37,018	**$ 483,306**
Cambridge University	$ 2,332	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,665	$ 4,664	$ 4,665	$ 4,665	$ 2,332	**$ 55,978**
Northeastern University	$ 1,976	$ 1,976	$ 1,976	$ 1,976	$ 9,599	$ 9,599	$ 9,599	$ 1,976	$ 1,974	$ 1,974	$ 1,974	$ 1,974		**$ 46,573**
System Integration														
BBN	$ 28,420	$ 27,877	$ 29,415	$ 29,091	$ 29,177	$ 29,090	$ 29,415	$ 29,090	$ 29,175	$ 29,090	$ 29,414	$ 28,853	$ 29,415	**$ 377,522**
Program Management														
BBN	$ 21,095	$ 22,568	$ 18,482	$ 13,860	$ 13,305	$ 29,977	$ 16,382	$ 25,949	$ 13,305	$ 13,860	$ 18,406	$ 13,097	$ 14,069	**$ 234,355**
Program Infrastructure														
BBN	$208,835	$ 12,499	$ 12,903	$ 12,901	$ 12,901	$ 12,901	$ 12,697	$ 50,784	$ 12,901	$ 12,901	$ 12,903	$ 12,903	$ 12,902	**$ 400,931**
Total Phase 3 Base	$471,516	$282,086	$288,526	$284,747	$292,716	$312,048	$294,971	$334,800	$281,448	$281,231	$284,214	$268,139	$217,669	$3,894,111
Speech Transcription Option														
BBN	$ 24,998	$ 25,562	$ 25,349	$ 26,752	$ 25,873	$ 26,457	$ 25,349	$ 26,457	$ 25,873	$ 26,752	$ 25,348	$ 26,457	$ 25,873	$ 337,100
Cambridge University	$ 12,497	$ 12,497	$ 12,496	$ 12,497	$ 12,497	$ 12,497	$ 12,496	$ 12,497	$ 12,497	$ 12,497	$ 12,497	$ 12,497	$ 12,497	$ 162,459
Total Option	$ 37,495	$ 38,059	$ 37,845	$ 39,249	$ 38,370	$ 38,954	$ 37,845	$ 38,954	$ 38,370	$ 39,249	$ 37,845	$ 38,954	$ 38,370	$ 499,559

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

11 Organizational Conflict of Interest Affirmations and Disclosure

None.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

12 Human Use

None.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

13 Animal Use

None.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

14 Statement of Unique Capability Provided by Government or Government-funded Team Member

Not Applicable.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

15 Government or Government-funded Team Member Eligibility

Not Applicable.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Appendix A: References

[1] P. Mermelstein, "Distance measures for speech recognition, psychological and instrumental," in *Pattern Recognition and Artificial Intelligence*, C. H. Chen (Ed.), pp. 374–388, Academic, New York, 1976.

[2] H. Hermansky, "Perceptual Linear Prediction (PLP) Analysis for Speech," *J. Acoust. Soc. Amer.*, Vol. 87, pp. 1738-1752, April 1990.

[3] H. Hermansky and N. Morgan, "RASTA Processing of Speech," *IEEE Transactions on Speech and Audio Processing*, Vol. 2, No. 4, pp. 587-589, October, 1994.

[4] M. Berouti, R. Schwartz, and J. Makhoul, "Enhancement of speech corrupted by acoustic noise," *IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 208-211, April, 1979.

[5] S. Ganapathy, S. Thomas, and H. Hermansky, "Robust Spectro-temporal Features based on Autoregressive Models of Hilbert Envelopes," *ICASSP*, Dallas, TX, March 2010.

[6] H. Hermansky, "The Modulation Spectrum in Automatic Recognition of Speech," *IEEE Workshop on Automatic Speech Recognition and Understanding*, S. Furui, B.-H. Juang and W. Chou (Eds.), 1997.

[7] T. Chi, P. Ru, and S. Shamma, "Multiresolution spectrotemporal analysis of complex sounds," *J. Acoust. Soc. Amer.*, Vol. 118, pp. 887-906, 2006.

[8] M. Mirbagheri, N. Mesgarani, and S. Shamma "Nonlinear filtering of spectrotemporal modulations in speech enhancement," *ICASSP*, Dallas, TX, March 2010.

[9] B. Zhang, S. Matsoukas, and R. Schwartz, "Discriminatively Trained Region Dependent Feature Transforms for Speech Recognition," *ICASSP*, Vol. 1, pp.I-313 - I-316, May 2006.

[10] H. Hermansky, D.P.W. Ellis, and S. Sharma, "Connectionist Feature Extraction for Conventional HMM Systems," *ICASSP*, 2000.

[11] H. Hermansky and S. Sharma, "Temporal Patterns (TRAPS) in ASR of Noisy Speech," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Phoenix, AZ, 1999.

[12] F. Valente and H. Hermansky, "Hierarchical and Parallel Processing of Modulation Spectrum for ASR applications," *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2008.

[13] J. Pinto, Y. Bayya, H. Hermansky, and M. Magimai-Doss, "Exploiting Contextual Information For Improved Phoneme Recognition," *IEEE Int. Conf. on Acoustics, Speech, and Signal Processing*, 2008.

[14] B. Zhang, S. Matsoukas, and R. Schwartz, "Recent progress on the discriminative region-dependent transform for speech feature extraction," *Interspeech*, Pittsburgh, PA, September 2006.

[15] J. Ma and R. Schwartz, "Unsupervised versus supervised training of acoustic models," *Interspeech*, 2008.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

[16] S. Novotney, R. Schwartz, and J. Ma, "Unsupervised acoustic and language model training with small amounts of labeled data," *ICASSP*, 2009.

[17] S. Novotney, and R. Schwartz, "Analysis of Low-Resource Acoustic Model Self-Training," *ICASSP*, 2009.

[18] H. Gish, M. Siu, A. Chan, and W. Belfield, "Unsupervised training of an HMM-based speech recognizer for topic classification," *ICASSP*, 2009.

[19] N. Mesgarani, M. Slaney, S. A. Shamma, "Discrimination of speech from nonspeech based on multiscale spectro-temporal modulations," *IEEE Transactions on Audio, Speech, and Language Processing*, vol.14, no.3, pp. 920- 930, May 2006.

[20] L. Deng, A. Acero, M. Plumpe, and X.D. Huang, "Large vocabulary speech recognition under adverse acoustic environments," *ICSLP*, Vol.3, pp. 806-809, October, 2000.

[21] M.J.F. Gales and S.J. Young, "An Improved Approach to the Hidden Markov Model Decomposition of Speech and Noise," *ICASSP*, Vol. 1, pp. 233-236, 1992.

[22] M.J.F. Gales, S.J. Young, "A Fast and Flexible Implementation of Parallel Model Combination," *ICASSP*, pp. 133-136, 1995.

[23] A. Acero, L. Deng, T. Kristjansson, and J. Zhang, "HMM adaptation using vector Taylor series for noisy speech recognition," *ICSLP*, September 2000.

[24] F. Flego and M.J.F., Gales, "Discriminative Adaptive Training with VTS and JUD," *IEEE Automatic Speech Recognition and Understanding Workshop*, December 2009.

[25] H. Xu, M.J.F., Gales, and K.K, Chin, "Improving Joint Uncertainty Decoding Performance by Predictive Methods for Noise Robust Speech Recognition," *IEEE Automatic Speech Recognition and Understanding Workshop*, December 2009.

[26] D.K, Kim and M.J.F., Gales, "Noisy Constrained Maximum Likelihood Linear Regression for Noise-Robust Speech Recognition," submitted to *IEEE Transactions on Audio Speech and Language Processing*, 2010.

[27] Z. Ghahramani and M. I. Jordan, "Factorial hidden Markov models," *Machine Learning*, 29:245–275, 1997.

[28] P. Kenny, G. Boulianne, P. Ouellet, and P. Dumouchel, "Joint factor analysis versus eigenchannels in speaker recognition," *IEEE Trans. Audio, Speech, Lang. Processing*, Vol. 15, no. 4, May 2007.

[29] G. Shafer, *A Mathematical Theory of Evidence*, MIT Press, 1976.

[30] F. Valente and H. Hermansky, "Combination of Acoustic Classifiers based on Dempster-Shafer Theory of Evidence," *IEEE International Conference on Acoustics, Speech and Signal Processing*, Honolulu, HI, 2007.

[31] R. Schwartz, A. Derr, A. Wilgus, and J. Makhoul, "Speaker Verification IR&D Final Report," *BBN Technical Report* No. 6661, Nov. 1987.

[32] E. Singer and D.A. Reynolds, "Analysis of multitarget detection for speaker and language recognition," *IEEE Odyssey: The Speaker and Language Recognition Workshop*, pp. 301-308, 2004.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

[33] H. Jin, R. Schwartz, S. Sista, and F. Walls, "Topic Tracking for Radio, TV Broadcast, and Newswire," *Eurospeech*, Budapest, Hungary, pp. 2439-2442, September 1999.

[34] H. Jin, R. Schwartz, F. Walls, and S. Sista, "Method and apparatus for score normalization for information retrieval applications," *U.S. Patents* 7,062,485 and 6,651,057.

[35] D.R.H. Miller, M. Kleber, C. Kao, O. Kimball, T. Colthurst, S.A. Lowe, R.M. Schwartz, and H. Gish, "Rapid and Accurate Spoken Term Detection," *Interspeech*, Antwerp, Belgium, 2007.

[36] S. Matsoukas, L. Nguyen, J. Davenport, J. Billa, F. Richardson, M. Siu, D. Liu, R. Schwartz, and J. Makhoul, "The 1998 BBN Byblos Primary System Applied to English and Spanish Broadcast News Transcription," *DARPA Broadcast News Workshop*, Herndon, VA, Morgan Kaufmann Publishers, pp. 255-260, Feb. 28-Mar. 3, 1999.

[37] S. Galligan, "Automatic speech detection and segmentation of air traffic control audio using parametric trajectory models," *IEEE Aerospace Conference*, Big Sky, Montana, March 2007.

[38] A. Stolcke, L. Ferrer, S. Kajarekar, E. Shriberg, and A. Venkataraman, "MLLR Transforms as Features in Speaker Recognition," *Eurospeech*, pp. 2425-2428, September 2005.

[39] E. E. Shriberg "Higher level features in speaker recognition," in *Speaker Classification I*, C. Muller (Ed.), Vol. 4343 of *Lecture Notes in Computer Science*, Artificial Intelligence, Springer, pp. 241-259, 2007.

[40] C.J. Leggetter and P.C. Woodland, "Maximum likelihood linear regression for speaker adaptation of continuous density HMMs," *Computer Speech and Language*, Vol. 9, pp. 171-186, 1995.

[41] M.A. Zissman, "Comparison of four approaches to automatic language identification of telephone speech," *IEEE Trans. on Speech and Audio Processing*, 4(1):31-44, 1996.

[42] L. Burget, P. Matějka, and J. Černocký, "Discriminative Training Techniques for Acoustic Language Identification," *ICASSP*, Toulouse, France, pp. 209-212, 2006.

[43] T. Anastasakos, J. McDonough, R. Schwartz, and J. Makhoul, "A compact model for speaker-adaptive training," *Int. Conf. Speech Language Processing*, vol. 2, Philadelphia, PA, pp. 1137–1140, 1996.

[44] H. Li and B. Ma, "A Phonotactic Language Model for Spoken Language Identification," *43rd Annual Meeting of the ACL*, Ann Arbor, MI, pp. 515–522, June 2005.

[45] S. Sista, R. Schwartz, T. Leek, and J. Makhoul, "An Algorithm for Unsupervised Topic Discovery from Broadcast News Stories," *Human Language Technology Workshop*, San Diego, CA, 2002.

[46] S. F. Boll, "Suppression of acoustic noise in speech using spectral subtraction," *IEEE ASSP-27*, pp. 113-120, April 1979.

[47] H. Gish, Y.-L. Chow, and J. R. Rohlicek, "Probabilistic Vector Mapping of Noisy Speech parameters for HMM word spotting," *ICASSP*, 1990.

[48] K. Ng, H. Gish, and J. R. Rohlicek, "Robust Mapping of noisy speech parameters for HMM word spotting," *ICASSP*, 1992.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

[49] X. Zhu, "Multi-Stage speaker diarization for conference and lecture meetings," http://www.itl.nist.gov/iad/mig/tests/rt/2007/workshop/RT07-LIMSI-SPKR-Presentation_v2.pdf

[50] C. Wooters and M. Huijbregts, "The ICSI 2007 speaker diarization system," http://www.itl.nist.gov/iad/mig/tests/rt/2007/workshop/RT07-SRIICSI-SPKR-Presentation.pdf

[51] X. Anguera, M. Aguilo, C. Wooters, C. Nadeu, and J. Hernando, "Hybrid Speech/non-speech detector applied to Speaker Diarization of Meetings," *IEEE Odyssey Speaker and Language Recognition Workshop*, pp. 1-6, 2006.

[52] E. Scheirer and M. Slaney, "Construction and evaluation of a robust multifeature speech/music discriminator," *ICASSP*, 1997.

[53] B. Kingsbury, G. Saon, L. Mangu, M Padmanabhan and R. Sarikaya, "Robust speech recognition in noisy environments: The 2001 IBM SPINE evaluation system," *ICASSP*, 2002.

[54] D. A. Reynolds, R. C. Rose, and M. J. T. Smith, "A Mixture Modeling Approach to Text-Independent Speaker Identification," *Journal of the Acoustical Society of America*, Suppl. 1, Vol. 87, 1990.

[55] L. Burget, P. Matějka, P. Schwarz, O. Glembek, and J. Černocký, "Analysis of feature extraction and channel compensation in GMM speaker recognition system," *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 7, 2007.

[56] W. Campbell, D. E. Sturim, D. A. Reynolds, "Support Vector Machines Using GMM Supervectors for Speaker Verification," *IEEE Signal Processing Letters*, Vol. 13, No. 5, pp. 308–311, May 2006.

[57] L. Burget, P. Matějka, V. Hubeika, and J. Černocký, "Investigation into variants of Joint Factor Analysis for speaker recognition," *Interspeech*, Brighton, UK, pp. 1263-1266, 2009.

[58] A. Solomonoff, W. M. Campbell, and C. Quillen, "Nuisance Attribute Projection," *Speech Communication*, Amsterdam, The Netherlands, May 2007.

[59] W. M. Campbell, J. P. Campbell, D. A. Reynolds, D. A. Jones, and T. R. Leek, "High-Level Speaker Verification with Support Vector Machines," *ICASSP*, Montreal, Quebec, Canada, pp. I: 73-76, May 2004.

[60] J. P. Campbell, D. A. Reynolds, and R. B. Dunn, "Fusing High- and Low-Level Features for Speaker Recognition," *Eurospeech*, 2003.

[61] N. Brümmer, L. Burget, J. Černocký, O. Glembek, F. Grézl, M. Karafiát, D. van Leeuwen, P. Matějka, P. Schwarz, and A. Strasheim, "Fusion of heterogeneous speaker recognition systems in the STBU submission for the NIST speaker recognition evaluation 2006," *IEEE Transactions on Audio, Speech, and Language Processing*, Vol. 15, No. 7, pp. 2072-2084, 2007.

[62] P. Matějka, P. Schwarz, L. Burget, and J. Černocký, "Use of anti-models to further improve state-of-the-art PRLM language recognition system," *ICASSP*, Toulouse, France, pp. 197-200, 2006.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

[63] J.R. Rohlicek, P. Jeanrenaud, K. Ng, H. Gish, B. Musicus, M. Siu, "Phonetic training and language modeling for word spotting," *ICASSP*, 1993.

[64] J.G. Fiscus, J. Ajot, J.S. Garofolo, and G. Doddington, "Results of the 2006 Spoken Term Detection Evaluation," *Special Interest Group on Information Retrieval* (SIGIR-07) *Workshop in Searching Spontaneous Conversational Speech*, 2007.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

Appendix B: Effect of Target Size

In this Appendix, we examine the effect of target set size on the probability of miss, P_{miss}, and the probability of false alarm, P_{FA}. We will examine the effect for the speaker ID problem, but the same analysis will apply to other problems, such as keyword spotting and language ID. This effect has been well documented in the literature [Singer and Reynolds, 2004; Zigel and Wasserblat, 2006].

For this analysis, we shall assume that the target speaker set, S, is a random set of N target speakers and that each acceptance or rejection of a speaker is independent. A hit is defined as the correct identification of any of the speakers in the set S. So, a miss is the false rejection of a speaker who is in the set S as none of the speakers in that set. Similarly, a false alarm is defined as the false acceptance of a speaker not in the set S as one of the speakers in the set S. We will now derive the probability of a miss and of a false alarm for the set of N speakers in terms of the corresponding probabilities for a single target speaker. For simplicity, we will assume that each of the target speakers has the same miss and false alarm probability.

Let

P_{miss} = probability of a miss (false reject) for a single target speaker, and

P_{FA} = probability of a false alarm (false accept) for a single target speaker.

In other words, P_{miss} is the probability that a sample from the target speaker is rejected in favor of the set of all speakers outside of the target set, and P_{FA} is the probability that the system recognizes a sample from a speaker outside the target set as the target speaker. We then conclude that the probability of correctly recognizing a sample from one of the speakers outside the target set as ***not*** from the target speaker, is given by:

P(correct reject) = 1- P_{FA} = probability of correctly recognizing a sample not from the target speaker as such.

For a miss in the set S to take place, the system must decide that a sample from our target speaker is neither from that speaker (false reject) nor is it from any of the other N-1 speakers in the set S (correct reject). Because of the independence assumption, the joint probability of the miss event is the product of the individual probabilities. The probability of falsely rejecting the target speaker is just P_{miss}. The probability of correctly rejecting each of the other N-1 speakers in the set S is 1- P_{FA}. Therefore, the probability of a miss in the set is just the product of the component probabilities:

$$P(\text{miss-set}) = P_{miss}(1 - P_{FA})^{N-1} \qquad (1)$$

Now, for a false alarm in the set to take place, a test sample from outside the target set must be falsely accepted as one of the target speakers. So, the P(false-alarm-set) is one minus the probability of being correctly rejected by all of the N target models. In other words,

$$P(\text{false-alarm-set}) = 1 - (1 - P_{FA})^{N} \qquad (2)$$

The table below shows an example with P_{miss} =5% and P_{FA} =1% for different values of set size N. As one would expect, P(miss-set) does decrease as N increases, but very slowly. However, we also see that P(false-alarm-set) increases almost linearly with N, because there are N ways to have a false alarm. That can be seen from equation (2), which for small P_{FA} and values of N less than 1/ P_{FA}, can be approximated by:

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.

$P(\text{false-alarm-set}) \approx N * P_{FA}$, for $P_{FA} << 1$, and $N << 1/P_{FA}$. (3)

P(miss-set) could also be approximated by:

$P(\text{miss-set}) \approx P_{miss}[1 - (N-1)P_{FA}]$, for $P_{FA} << 1$, and $N << 1/P_{FA}$.

So, although P(miss-set) decreases as N increases, the effect of the decrease is not as dramatic as the relative increase in P(false-alarm-set). Our conclusion is that having a larger set of target speakers does make a test somewhat easier in terms of probability of miss, but it makes it much, much harder in terms of probability of false alarm.

N	P(miss)	P(FA)
1	5.0%	1.0%
5	4.8%	4.9%
10	4.6%	9.6%
25	3.9%	22.2%
100	1.8%	63.4%

Now, the above analysis made a number of simplifying assumptions, so the question is whether the above described effect occurs in practice or not. We first present an experimental result and then we give a different scenario of interest.

In a project on speaker ID sponsored by AFRL [Gish and Bellfield, 2004], we performed experiments with target sets of multiple speakers (N=15, 30, 45) using data from the TIMIT database, supplied to us by AFRL. The speaker ID experiments we performed showed that the experimental ROC curves for each of the three conditions very much followed the expected ROC curves as would be predicted from (1) and (2) above.

The independence assumptions made in the above analysis might not hold for certain applications. For example, there might be interest in a set of speakers who use the same communication channel, so you would expect those speakers to sound more similar to each other and, therefore, be more correlated. In keyword spotting, the set of keywords may be similar to each other, e.g., variations of pronunciations of the same word in different dialects. In such cases, one would not expect the false alarm rate to increase as rapidly with increased set size.

E. Singer and D. Reynolds, "Analysis of Multitarget Detection for Speaker and Language Recognition," *IEEE Odyssey: The Speaker and Language Recognition Workshop*, Toledo, OH, 2004.

Y. Zigel and M. Wasserblat, "How to Deal with Multiple Targets in Speaker Identification Systems," *IEEE Odyssy: The Speaker and Language Recognition Workshop*, pp. 1–7, 2006.

H. Gish and W. Bellfield, "Improved Methods for Open Set Speaker Identification," *Final Report*, AFRL Contract F30602-03-C-0096, 2004.

Use or disclosure of the data contained on this page is subject to the restriction on the title page of this proposal.