



STO MEETING PROCEEDINGS

MP-IST-178

Big Data Challenges: Situation Awareness and Decision Support

(Les défis des méga-données (Big Data): Appréciation de situation
et aide à la décision)

STO Information Systems and Technology (IST) Panel Workshop
held on 15 – 16 October 2019 in Budapest, Hungary.



Published October 2019

The NATO Science and Technology Organization

Science & Technology (S&T) in the NATO context is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models, namely a collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange, and secondly an in-house delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The mission of the NATO Science & Technology Organization (STO) is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives, and contributing to NATO's ability to enable and influence security and defence related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies.

The total spectrum of this collaborative effort is addressed by six Technical Panels who manage a wide range of scientific research activities, a Group specialising in modelling and simulation, plus a Committee dedicated to supporting the information management needs of the organization.

- AVT Applied Vehicle Technology Panel
- HFM Human Factors and Medicine Panel
- IST Information Systems Technology Panel
- NMSG NATO Modelling and Simulation Group
- SAS System Analysis and Studies Panel
- SCI Systems Concepts and Integration Panel
- SET Sensors and Electronics Technology Panel

These Panels and Group are the power-house of the collaborative model and are made up of national representatives as well as recognised world-class scientists, engineers and information specialists. In addition to providing critical technical oversight, they also provide a communication link to military users and other NATO bodies.

The scientific and technological work is carried out by Technical Teams, created under one or more of these eight bodies, for specific research activities which have a defined duration. These research activities can take a variety of forms, including Task Groups, Workshops, Symposia, Specialists' Meetings, Lecture Series and Technical Courses.

The content of this publication has been reproduced directly from material supplied by STO or the authors.

Published October 2019

Copyright © STO/NATO 2019
All Rights Reserved

ISBN 978-92-837-2319-6

Single copies of this publication or of a part of it may be made for individual use only by those organisations or individuals in NATO Nations defined by the limitation notice printed on the front cover. The approval of the STO Information Management Systems Branch is required for more than one copy to be made or an extract included in another publication. Requests to do so should be sent to the address on the back cover.

Big Data Challenges: Situation Awareness and Decision Support

(STO-MP-IST-178)

Executive Summary

The aim of this two day Inter-Panel / Inter-Group Workshop was to establish raised awareness of common and different problems, available and foreseen technologies, capabilities and efforts, and also to support the development of collaborations, across the many Research Task Groups (RTGs) and Exploratory Teams (ETs) operating under the six NATO panels (AVT, HFM, IST, SAS, SCI and SET) and the NATO Modelling and Simulation Group (NMSG). Some targeted outcomes were the sharing of data, tools and tasks as well as the generation of new research ideas for RTGs and /or ET Technical Activity Proposals (TAPs).

The workshop addressed, in particular, how the availability of big data may assist the military in understanding the situations they encounter and will encounter, and how this may support them in making better decisions.

The workshop call elicited a large number of paper submissions which were down selected by peer review. The workshop was extremely well attended by members of the different RTGs and ETs as well as IST panel members (approximately 100 participants from across the NATO and Partner for Peace countries). In addition to presentation and demonstration sessions, a core aspect of the workshop was a series of break-out sessions in which sub-groups of attendees worked together on selected topics (enabling participants to explore and benefit from each other's expertise, their work as well as to generate new ideas).

It was considered that the workshop fully achieved its aims and resulted in very fruitful and informative discussions on a broad spectrum of potential (military) domains that had relevance in NATO and beyond, in particular in respect to:

- Artificial intelligence
- Challenges in Big Data exploration / exploitation
- Cyber defence and security
- Evaluation frameworks, measures and metrics
- Genomics
- Human Factors
- Information warfare
- Knowledge engineering
- Mission planning
- Modelling and simulation
- Situation awareness and Decision Support
- Social media exploitation
- System concepts and integration
- Visual Analytics and visualization
- War gaming
- etc

Les défis des méga-données (Big Data): Appréciation de situation et aide à la décision

(STO-TR-IST-178)

Synthèse

L'objectif de ce séminaire inter-commission/inter-groupe de deux jours était de sensibiliser les participants à des problèmes communs et différents, aux technologies, capacités et travaux disponibles et prévus, ainsi que de soutenir le développement de collaborations entre les nombreux groupes de recherche (RTG) et équipes exploratoires (ET) des six commissions de l'OTAN (AVT, HFM, IST, SAS, SCI et SET) et le Groupe OTAN sur la modélisation et la simulation (NMSG). Ce séminaire visait notamment au partage de données, d'outils et de tâches, ainsi qu'à la création de nouvelles idées de recherche pour la proposition d'activités techniques (TAP) des RTG et/ou des ET.

Le séminaire s'est intéressé en particulier à la manière dont la disponibilité des données massives pourrait aider les militaires à comprendre les situations qu'ils rencontrent et rencontreront et l'aide qu'ils pourraient tirer de ces données massives pour prendre de meilleures décisions.

La convocation du séminaire a suscité l'envoi d'un grand nombre d'articles, que nous avons sélectionnés au moyen d'un comité de lecture. La participation au séminaire a été extrêmement bonne, puisqu'une centaine de membres des différents RTG et ET et de la Commission IST se sont réunis (venant des pays de l'OTAN et du Partenariat pour la paix). En plus d'exposés et de démonstrations, le séminaire se composait principalement de réunions en petits groupes, pendant lesquelles les participants ont travaillé ensemble sur des sujets choisis, ce qui leur a permis d'explorer et bénéficier de l'expertise des autres et de leur travail et de produire de nouvelles idées.

Il a été considéré que le séminaire avait pleinement atteint ses objectifs et engendré des discussions informatives très fructueuses dans un large spectre de domaines (militaires) potentiels, importants au sein de l'OTAN et au-delà, concernant notamment les thèmes suivants:

- Intelligence artificielle ;
- Défis de l'exploration/exploitation des données massives ;
- Cyberdéfense et sûreté ;
- Cadres d'évaluation, mesures et indicateurs ;
- Génomique ;
- Facteurs humains ;
- Guerre de l'information ;
- Génie de la connaissance ;
- Planification des missions ;
- Modélisation et simulation ;
- Connaissance de la situation et aide à la décision ;
- Exploitation des médias sociaux ;
- Concepts et intégration des systèmes ;
- Analyse visuelle et visualisation ;
- Jeux de guerre ;
- Etc.



SCIENCE & TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

WORKSHOP

BIG DATA CHALLENGES: SITUATION AWARENESS AND DECISION SUPPORT

IST-178

organised by the

Information Systems and Technology Panel

to be held at

Radisson Blu Beke Hotel, Budapest, Hungary

Tuesday 15 October - Wednesday 16 October 2019

This Workshop is open to NATO nations, PfP nations, Mediterranean Dialogue, ICI nations and Global Partners



ENROLMENTS: Participants are requested to enrol on line at
<https://events.sto.nato.int>

Deadline for enrolments is 1 October 2019

There are no conference fees.

Collaboration Support Office – Point of Contact:

Mrs. Aysegül APAYDIN Tel: +33 (0)1 5561 2282
IST Panel Assistant Fax: +33 (0)1 5561 9626



Background

Information Systems Technology Panel (IST) is one of the seven Panels whose role it is to implement, on behalf of the Science & Technology Board, the STO Mission with respect to Information Systems Technology. The advancement and exchange of techniques and technologies to provide timely, affordable, dependable, secure and relevant information to war fighters, planners and strategists, as well as enabling technologies for modelling, simulation, and training are the focus of this Panel. The Information Systems Technology Panel covers the fields of Information Warfare and Assurance, Information and Knowledge Management, Communications and Networks and Architecture and Enabling Technologies.

Theme - Objectives –Topics

There are many Research Task Groups (RTGs) and Exploratory Teams (ETs) under different NATO Panels and the NATO Modelling and Simulation Group (NMSG) all working on different aspects of defence, military and intelligence challenges, e.g. big data exploration, decision support, mission planning, human factors, medicine, modelling and simulation and visual analytics etc. The IST-178 Workshop invites members of the different RTGs and ETs together in a workshop that will enable them to explore and benefit from each other's work, and to develop collaborations to complement and enhance their respective efforts, as well as facilitating the sharing of data, tasks and tools.

The main objective of this two-day Inter-Panel / Inter-Group Workshop is to provide the means for the RTG/ET members to discuss views, challenges, and work on the use of big data to support the military with obtaining situation awareness and informed decision making.

The aim of the Workshop is to establish a raised awareness of common and different problems, available and foreseen technologies, capabilities and efforts, as well as to support the development of collaborations. This may lead to outcomes such as joint Inter-Panel / Inter-Group collaborations; sharing of data, tools and tasks; and the generation of new research ideas for RTGs and /or ET Technical Activity Proposals (TAPs).

TOPICS:

The theme of this meeting is how the availability of big data may assist the military in understanding the situations they will encounter, and how this may support them in making better decisions. The challenges for development and opportunities for the exploitation of Big Data are likely to vary from domain to domain. This workshop intends to host the discussion for a broad spectrum of potential (military) domains, such as Human Factors, Artificial Intelligence, Information Warfare, Situation Awareness and Decision Support, Visual Analytics and visualization, etc...

WORKSHOP FORMAT and ORGANIZATION:

The workshop will commence with selected presentations to inspire and provide directions for the workshop discussion. We plan an interactive workshop, consisting of several working groups, each with a different theme, topic, and/or domain. The working groups will be arenas for in-depth discussions. Each participant in the workshop will be assigned to a working group, with the aim to have participants from the various communities in each group. These working groups will be formed on the first day. Each working group is to produce a summary report of their discussion, and to present it to the other groups at the end of the workshop. Attendance will be limited to 50 people.

Programme Committee

Programme Co-Chairs

Dr. Margaret VARGA

Seetru Ltd. / Oxford University
United Kingdom

Email: margaret.varga@zoo.ox.ac.uk

Dr. Karel van den BOSCH

TNO Behavioural and Social Sciences
The Netherlands

Email: karel.vandenbosch@tno.nl

Members

Dr. Elena CAMOSSÌ

Center for Maritime Research & Experimentation
CMRE

Email: elena.camossi@cmre.nato.int

Mrs. Francesca de ROSA

Center for Maritime Research & Experimentation
CMRE

Email: francesca.derosa@cmre.nato.int

Prof. Wim MEES

Royal Military Academy
Belgium

Email: wim.mees@rma.ac.be

Col.Dr.Eng. Nikolai STOIANOV

Defence Institute
Bulgaria

Email: n.stoianov@di.mod.bg

Dr. Cathy BOSCARINO

Defence Research & Development Canada
Canada

Email: cathy.boscarino@drdc-rddc.gc.ca

Dr. Bruce FORRESTER

Defence Research & Development Canada
Canada

Email: bruce.forrester@drdc-rddc.gc.ca

Ms. Valérie LAVIGNE

Defence Research & Development Canada
Canada

Email: valerie.lavigne@drdc-rddc.gc.ca

Prof. Bob MADAHAR

Dstl
United Kingdom

Email: bkmadahar@dstl.gov.uk

Prof. Cyril RAY

IRENav, Ecole Navale
France

Email: cyril.ray@ecole-navale.fr

Mr. Albert PRITZKAU

Fraunhofer-FKIE

Germany

Email: albert.pritzkau@fkie.fraunhofer.de**Dr. Carsten WINKELHOLZ**

Fraunhofer-FKIE

Germany

Email: carsten.winkelhorz@fkie.fraunhofer.de**Dr. Andras BENCZUR**

Hungarian Academy of Sciences

Hungary

Email: benczur@sztaki.mta.hu**Prof. Tomas KRILAVICIUS**

Baltic Institute of Advanced Technology

Vytautas Magnus University

Lithuania

Email: tomas.krilavicius@bpti.lt**Dr. Petter BIVALL**

Swedish Defence Research Agency

Sweden

Email: petter.bivall@foi.se**Dr. Liz BOWMAN**

US Army Futures Command

United States

Email: elizabeth.k.bowman.civ@mail.mil**Dr. Alexander KOTT**

Army Research Laboratory

United States

Email: alexander.kott1.civ@mail.mil**Dr. Kristen LIGGETT**

Air Force Research Laboratory

United States

Email: kristen.liggett@us.af.mil**Dr. Michael STREET**

NCIA

The Netherlands

Email: michael.street@ncia.nato.int**Ms. Ivana Ilic MESTRIC**

NCIA

The Netherlands

Email: ivana.ilicmestric@ncia.nato.int**Science and Technology Organization in NATO**

In NATO, Science & Technology (S&T) is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models:

- The Collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange.
- The In-House delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The Science and Technology Organization - STO

The mission of the NATO STO is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by:

- Conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives;
- Contributing to NATO's ability to enable and influence security- and defence-related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies;
- Supporting decision-making in the NATO Nations and NATO.

Acknowledgements

We wish to thank our hosts from Hungary for organising and hosting this event in the facilities they have chosen in their lovely city, Budapest.



<http://www.sto.nato.int>



IST-178 Workshop on Big Data Challenges: Situation Awareness and Decision Support Programme

Tuesday 15 October 2019

08:30 REGISTRATION

09:00 OPENING CEREMONY

Host Welcome Speech

Introduction: Dr.Eng. Michael WUNDER, Chairman IST Panel

Introduction to the Workshop: Dr. Margaret VARGA, GBR

Dr. Karel van den BOSCH, NLD

Administrative Announcements: LtCol. Attila HAJDU, HUN

SESSION 1 – INTRODUCTION

09:30 KEYNOTE SPEECH 1:

Big Data Challenges – Like War

by Dr. Ing. Michael WUNDER, IST Panel Chairman, DEU

10:15 BREAK

SESSION 2– AI and MACHINE LEARNING

10:45 1 Extracting Value from NATO Data Sets through Machine Learning and Advanced Data Analytics
by I.I. MESTRIC, A. KOK, G. VALIYEV, M. STREET,
P. LENK, NCIA, NLD, M. RACOVITA, F. VIEIRA, JALLC,
PRT

11:15 2 Living S&T Roadmap for Artificial Intelligence and Big Data for Military Decision Making
by F. DESHARNAIS, DR&D Canada, CAN, A. BARROS,
P. KERBUSCH, TNO, NLD

11:35 3 Exploitation of Sensor Data Using Artificial Intelligence for Battlefield Sensemaking
by V. LAVIGNE, M. MOKHTARI, M. BRETON, E.
MARTINEAU, J. FOURNIER, DRDC Canada, CAN

11:55 4 Ukrainian Conflict in Media: Two Approaches to Narrative Analysis
by T. KRILAVICIUS, J. MANDRAVICKAITE, BPTI,
Vytautas Magnus University, LTU

12:25 LUNCH

SESSION 3 – HUMAN FACTORS

13:40 5 Developing Transparent Conversational Agent Systems for Intelligence Analysis
by S. HEPENSTAL, DSTL, GBR

14:10 6 Determining the Effectiveness of Representing Intrusion Detection System Log Files with Visualization Techniques
by G. THOMAS, K. LIGGETT, Wright Patterson AFB, USA

14:30 7 Meaningful Human Control of AI-Based Systems: What are the Implications for Situational Awareness and Decision Support?
by F. BUTCHER, DSTL, GBR

14:50 BREAK & SESSION 4 – DEMONSTRATION

15:50 SESSION 5 – BREAKOUT GROUP DISCUSSIONS

17:05 SESSION 6 – DAY SUMMARY

17:25 END

19:00 HOST NATION RECEPTION (for all attendees)

Wednesday 16 October 2019

09:00 KEYNOTE SPEECH 2:

What's Next for Visual Analytics?

by Ms. Kristin COOK, Pacific Northwest National Laboratory, USA

SESSION 7 – INFORMATION WARFARE AND SOCIAL MEDIA

09:45 8 Security Perspectives on Social Media Exploitation
by N. BIETENIECE, NATO STRATCOM COE,
V. DRAGOS, ONERA, FRA, B. FORRESTER, DRDC
Canada, T. KRILAVICIUS, BPTI, LTU, A. PRITZGAU,
Fraunhofer FKIE, DEU

10:05 9 Propaganda Filters: Tracking Malign Foreign Interventions on Social Media in Real Time
by B. FORRESTER, DRDC Canada, Z. DEVEREAUX, A.
BACOVICIN, S. BEDOYA, NEXALOGY, CAN

10:25 BREAK

10:45 10 People, Bots or Cyborgs? Analysis of a Hashtag War
by R. GOOLSBY, US Navy, USA

11:05 11 Towards Big Data in Tactical Domain
by F-T. JOHNSEN, FFI, NOR

11:25 12 Assessment of IoT Data Ingest Reliability for Urban Environments
by J.R. MICHAELIS, J. FREEMAN, A. RAGLIN, US Army

Research Lab., USA

SESSION 8 – VISUAL ANALYTICS AND VISUALIZATION

11:45 13 Storytelling Exploratory Visual Analytics for Counter-Improvised Explosive Device Incidents
by V. LAVIGNE, S. JAYARAM, M. PANGE, DRDC
Canada, CAN

12:15 14 A Preliminary Experimentation for Meteorological Data Visual Analytics
by F.M. PISANO, C. ASSUNTINA, R. GIANPAOLO, CIRA,
ITA

12:35 15 Genomic Data Analysis for Infectious Diseases
by M. VARGA, K. JOLLEY, H. BRATCHER,
C. RODRIGUES, J. MACLENNAN, O. HARRISON,
F. COLLES, J. BRAY, M. MAIDEN, University of Oxford,
GBR

12:55 LUNCH

14:00 SESSION 9 – BREAKOUT DISCUSSIONS

15:00 BREAK

15:20 SESSION 10 – BREAKOUT GROUPS' PRESENTATIONS

16:20 CONCLUDING REMARKS AND WORKSHOP CLOSING

16:30 END OF MEETING

Programme Committee

Co-Chairs

Dr. Margaret VARGA
Seetru Ltd. / Oxford University
UNITED KINGDOM

Email: margaret.varga@zoo.ox.ac.uk

Dr. Karel van den BOSCH
TNO Behavioural and Social Sciences
THE NETHERLANDS

Email: karel.vandenbosch@tno.nl

Members

Dr. Elena CAMOSSÌ
Center for Maritime Research & Experimentation
CMRE

Email: elena.camossi@cmre.nato.int

Dr. Carsten WINKELHOLZ
Fraunhofer-FKIE
GERMANY

Email: carsten.winkelhorz@fkie.fraunhofer.de

Mrs. Francesca de ROSA
Center for Maritime Research & Experimentation
CMRE

Email: francesca.derosa@cmre.nato.int

Dr. Andras BENCZUR
Hungarian Academy of Sciences
HUNGARY

Email: benczur@sztaki.mta.hu

Prof. Wim MEES
Royal Military Academy
BELGIUM

Email: wim.mees@rma.ac.be

Prof. Tomas KRILAVICIUS
Baltic Institute of Advanced Technology
Vytautas Magnus University
LITHUANIA

Email: tomas.krilavicius@bpti.lt

Col.Dr.Eng. Nikolai STOIANOV
Defence Institute
BULGARIA

Email: n.stoianov@di.mod.bg

Dr. Petter BIVALL
Swedish Defence Research Agency
SWEDEN

Email: petter.bivall@foi.se

Dr. Cathy BOSCARINO
Defence Research & Development Canada
CANADA

Email: cathy.boscarino@drdc-rddc.gc.ca

Dr. Liz BOWMAN
US Army Futures Command
UNITED STATES

Email: elizabeth.k.bowman.civ@mail.mil

Dr. Bruce FORRESTER
Defence Research & Development Canada
CANADA

Email: bruce.forrester@drdc-rddc.gc.ca

Dr. Alexander KOTT
Army Research Laboratory
UNITED STATES

Email: alexander.kott1.civ@mail.mil

Ms. Valérie LAVIGNE
Defence Research & Development Canada
CANADA

Email: valerie.lavigne@drdc-rddc.gc.ca

Prof. Bob MADAHAR
Dstl
UNITED KINGDOM

Email: bkmadahar@dstl.gov.uk

Prof. Cyril RAY
IRENav, Ecole Navale
FRANCE
Email: cyril.ray@ecole-navale.fr

Mr. Albert PRITZKAU
Fraunhofer-FKIE
GERMANY
Email: albert.pritzkau@fkie.fraunhofer.de

Dr. Kristen LIGGETT
Air Force Research Laboratory
UNITED STATES

Email: kristen.liggett@us.af.mil

Dr. Michael STREET
NCIA
THE NETHERLANDS
Email: michael.street@ncia.nato.int

Ms. Ivana Ilic MESTRIC
NCIA
THE NETHERLANDS
Email: ivana.ilicmestric@ncia.nato.int

Technical Evaluation Report

G.Wyman
UNITED KINGDOM

glyn.wyman@gmail.com

ABSTRACT

This paper presents the findings of the Technical Evaluator relating to the workshop on Big Data Challenges: Situation Awareness and Decision Support held in Budapest 15-16 October 2019. A summary of each invited presentation is included as an annex. Data capture remains a critical element for presenting succinct information, AI can assist in this aim but the users must be aware of the limitations. Confidence in the data presented and the conditions of use should be displayed. Visual Analytics need further research to supplement the existing tool set, evidence of progress in visualisation techniques was noted. It concludes that the vast amount of available data is already exploited for commercial and state purposes and that users must be educated to be aware of the impact that third parties can exert.

INTRODUCTION

The proliferation of data continues apace and the distribution is feasible over the internet for commercial and other enterprise. Data from a wide spectrum of sources is becoming readily available and requires analysis to present the necessary and sufficient information to the operators. Users need relevant information in a timely fashion but defining what is relevant can be difficult and will vary with the background, environmental context and psychological state of that person. AI development is now sufficiently mature to achieve some of the desires which coupled with advances in Visual Analytics gives enhanced understanding of the information displayed. A whole raft of AI tools have evolved and continue to be developed, critical to the models is the training data which needs to be extensive. In the military field some data can be sparse or highly sensitive, in this instance the training data could be supplemented with synthetic data derived from other models or through data farming. The ethical and legal implications need extensive thought but should not be insurmountable. The field of Visual Analytics and AI has attracted academic and commercial interest which encourages innovation in, for example, reinforced learning tools. The desire is to retain information superiority over an adversary.

The Workshop was called and invited scientists in the area to combine their knowledge with experts involved in current NATO projects. It is noted that the 'Call for Papers' restricted the attendance to 50 but it is unknown how the attendees were selected.

THEME

The theme of this meeting is how the availability of big data may assist the military in understanding the situations they will encounter, and how this may support them in making better decisions. The challenges for development and opportunities for the exploitation of Big Data are likely to vary from domain to domain the following were identified for consideration:

Applied vehicle technology	Human cognition and perception	Situation awareness and Decision Support
Artificial intelligence	Information warfare	System analysis and studies
Challenges in Big Data exploration / exploitation	Knowledge engineering	System concepts and integration
Cyber defence and security	Medicine	Visual Analytics and visualization
Evaluation frameworks, measures and metrics	Mission planning	War gaming
Genomics	Modelling and simulation	
Human Factors	Sensors and electronic technology	

PERCEIVED MILITARY ISSUES

Information superiority is paramount to achieve dominance, regrettably the military appear to lag the development and exploitation of big data observed in other fields. Information warfare has already been observed. Decisions are critical at all levels of the Military Command where succinct, timely information is essential presented with a high level of confidence. The significant increase in raw data obtained from, e.g. ad hoc sensors and social media will only increase. Further enlargement is envisioned as a result of the proliferation of small low earth orbiting satellites. Filtering of these massive amounts of data is essential to avoid overload of an individual viewing the information, demanding a level of data analysis with the ability to identify false sets of information whether explicit or subliminal. Trust in the systems is critical with any derived actions displayed to the users, ideally coupled to the reasoning for taking that approach. Visual Analytics is already employed with a significant array of commercial tools. The use of extant modules to ease the development of new systems can incur penalties introducing uncertainty and thus reducing trust; this can be mitigated through transparency since formal analysis is impractical.

The increased use of Artificial Intelligence (AI) to identify the salient aspects of a situation exposes issues of verification impacting on the confidence. AI tools are evolving with a further level of AI employed to establish the model structure. Research and Development in commercial applications have the potential to be exploited but care needs to be exercised with respect to the security implications. Third parties can readily influence the outcome both in terms of the raw data but also the data employed during the training process, detection of these activities is non trivial. Third parties can also have an impact on communities through social media which will have an impact on any interface with that group. The aspect of using social media in a pro active manner was raised but not pursued.

Information specific to the operator taking cognisance of environment, background, ability and psychological disposition requires development. Collection of data using autonomous systems to supplement identified deficiencies need to be integrated. A pragmatic position on risk assessment must be understood by the decision makers made easier if recommended action made by intelligent agents are displayed.

EVALUATION

Structure

A Call for Papers was issued to solicit contributions and participation from scientists who have been engaged in this field. From the submissions, fifteen were selected with an additional two keynote papers designed to give a background with a general assessment of the research in hand. The technical papers were divided into four sessions: 1) AI and Machine Learning, 2) Human Factors, 3) Information Warfare and Social Media, and 4) Visual Analytics and Visualisation. All papers were relevant and benefited from the selection process, the quality was generally good. The papers and presentations were, in some instances, endorsed by demonstrations of the tools during the breaks. The inclined reader can access the full papers on the CSO website with a synopsis of each provided as an annex to this report. Leverage from other working groups was evident both from within the IST and other STO groups. It is strongly recommended that groups or individuals submit technical activity proposals (TAP) to establish research areas which can be supported by the IST or other STO Panels.

Overview

To successfully engage in any conflict information superiority is critical. Data relating to situational awareness proliferates and is compounded by the improvements in resolution and quantity of sensors. The internet of things (IoT) allows searches and is exploited in the commercial field. Social media also plays a significant role by targeting perceived weaknesses in a population or dividing opinions within a coalition. The so called Digital Influence Machine (DIM) allows the use of information as an effective weapon. DIM can search data associated with individuals and target their activities and derived views. This then allows propaganda to be focused, subtly modifying their views and changing their actions. Evidence was shown in which third parties are exploiting social media to this effect. Not all these aspects need be negative and the military could mobilise support; an example was cited in which the Boston Police force defused a volatile situation during the Boston Marathon by using social media. The use of Deep Fake Videos was raised as a means to bias opinion, it is reported that this media can have a strong influence. Lax regulations over the use of the internet allow fake news to proliferate which is no longer countered by reputable journalists; the issue to detect and remove false information is non trivial.

AI and Machine Learning

Trust was an issue raised extensively during the workshop and a strong recommendation to display a measure of confidence was offered. The size of the supporting tools and the nature of neural networks makes formal analysis impractical. The inherent nature of neural networks makes ascribing a confidence interval very difficult with a quantitative assessment at best. The degree of trust can be improved by knowing the provenance and having active trusted data organisations. If in specific instances the error bounds and distribution of the parameters is derived a figure can be computed but in most instances it is the result of a coupling of entities with unknown characteristics. AI and Deep Learning are trained in specific aspects and show promise but should not be overestimated, if applied within the design constraints they can assist operators and filter raw data rapidly. In principle training based machine intelligence can use generic architectures (several exemplar templates are available). Its functioning critically relies on the availability of data that is reliable and selected to be carefully in tune with the targeted objective. A strong recommendation was made to retain the human as part of the process.

Natural language processing has been advanced with appropriate dictionaries available, discussion relating to establishing an ontology for this aspect did not reach a conclusion but received support. Sentiment analysis where groups are identified with similar views is a potential enabler to direct propaganda but could be refined.

Data is the corner stone of this topic and the dissemination is crucial, the nature of military operations demands that some data will be classified and have a restricted distribution, however the wider the dissemination the better the models are trained. A limited set of training data can be mitigated by introducing simulated data but care must be exercised to avoid any bias which will skew the outcome. It is also feasible that third parties can poison the data which could modify the weightings within the model. The ability of the tools to identify trends was observed and the ability to extrapolate to make predictions. This is enhanced with visual analytics to display the results in a coherent fashion.

Evidence was provided to isolate subsets and outliers to enable the analysts to focus on specific effects, the concept of coupling notionally independent parameters to identify causality remains heuristic. If causality could be identified a solution space rapidly opens, false correlations remain an issue where similar characteristics are incorrectly correlated.

Human Factors

The aim is simply expressed to present the available raw data in a form readily assimilated, underlying this requirement is a whole raft of processes. AI is one element where assistance is available, other areas include conversational analysis which contains its own analysis history relevant to the topic. The display selected will be: a) user specific; adjusted to reflect ability, experience, and psychological state b) context specific; adjusted to match the environment and c) transparent. Software can assist in this by applying smart filters, applying fast querying, and identifying gaps. The importance of trust is again significant where not only a confidence indication is present but the reasoning behind the display. It is important to understand the demands of the user and reflect their requirements; some users will have different priorities.

Diversity should be considered to enhance confidence not only from the perspective of different views but ideally to employ different methods to derive the display. It is widely acknowledged that AI has considerable potential for Defence. However, its current state of progress, with its operational requirements, anticipated benefits and risks, requires that its users exercise a high level of critical judgement and awareness if it is to be used as a component in critical decision making. Decision makers can only assimilate a small quantity of data so that any displayed must be relevant and succinct.

Information Warfare and Social Media

Social media is experiencing a rapid expansion and is accessible in a large number of forms, a representative list is available in the papers but the list is not static and subject to evolution. The use of social media is complex and ranges from an individual informing friends to states attempting to influence targeted sections of the public. Evidence of use for nefarious purposes is readily available and should be regarded as a potential battle ground. NATO will need to question the actions it needs to take with a recommendation to consider a proactive approach. This could entail funding of bots to promoting trusted users. It was suggested that prominent scientists should be active to improve the public perception of the organisation. It should also be identifying clusters with particular views and accumulating information on detected bots to hold a library to identify future instances. Evidence was demonstrated of a current capability. The internet as a conduit is not only capable of distributing information but is a rich source to accumulate situational awareness where interested observers can identify an adversary preparing an event.

Visual Analytics and Visualisation

A wide variety of visual analytic tools were discussed and a selected set demonstrated. Decision makers have an inherent limitation on the information which can be absorbed. An attempt was made to quantify the figure during the breakout discussions with a figure of tens of kilohertz indicated. This emphasises the need to focus on the salient points. Further research is required to establish the degree of visual occlusion required to extract information created by the visualisation methods.

Adaptive screens were advocated allowing the users to drill down for specific detail or use their initiative to interrogate aspects observed in the initial view. Enhanced relational data structures were advocated and demonstrated to reduce the reaction interval to display the association. Humans are adapt at extracting patterns from displays particularly if they have practice in that scenario.

Facilities

The room was well appointed with the delegates sat in rows with a limited number of tables giving all delegates full visibility of the screen. WiFi access was provided giving access to the cso website and other internet facilities, however the nature of the WiFi provided by the Hotel required security considerations. Coffee was served during the breaks with ample opportunity to engage with the presenters to elaborate on their research and to seek points of clarification. I observed very productive discussions during the breaks.

CONCLUSIONS

The Workshop achieved its objectives in informing the invited subject experts of the issues affecting NATO and provided an interface with scientists working for The Organisation. The knowledge associated with the discipline was advanced with a number of aspects identified which require further research. Raw data is expanding at a rapid rate dictating that AI assistance is required for decision makers to comprehend the information. Humans in the loop remains essential but smart filters and visual analytics play a significant role in assisting the users. A raft of tools have been presented and demonstrated with the recommendation that more research is undertaken to allow tools to evolve.

Trust of the information displayed is paramount but recognised as difficult to achieve. Defining the source and provenance would assist as would only accepting data from a known repository, but this may be too restrictive. Identifying corrupted data particularly for training neural nets remains critical.

Social media coupled with the internet of things permits rapid access and dissemination of information to a wide audience. It also permits adversaries to influence the opinion of communities by giving fake news to a wide audience with bots relaying the data as if it were a true actor. Another tactic is to drive a wedge through the community to create division. NATO needs to consider a proactive stance.

I commend the co-chairs and their technical committee for a successful outcome.

ANNEX

Papers/Presentations

The inclined reader is directed to the CSO website where a file of the papers and power point slides are available, the following are extracts of the salient points, with comments.

Keynote 1: Big Data Challenges-Like War: M. Wunder The keynote speaker provided a sound background to the issues concluding that information has already been observed as a weapon. Civilian companies have unintentionally provided the means to implement the war and NATO should be vigilant and be aware of the consequences. The STO has over 100 sponsored activities associated with this discipline with additional aspects still requiring research, in particular the ability to transfer data whilst maintaining trust. AI needs new training of personnel and availability of relevant data to establish the models

Paper 1: Extracting Value from NATO Data Sets through Machine Learning and Advanced Data Analytics: I. Mestric A paper describing the activities within the NATO Agencies to enhance the visualisation of events during exercises and to enhance the products through lessons learned. The specific exercise is 'Trident

Junction 2018' It concludes that Military specific ontology would be a considerable advantage but that NATO specific research is necessary.

Presentation 2: Mission-Oriented Research for AI & BD for Military Decision Making (AIBD4MDM)
Theme: F. Desharnais The presentation was given by the chair of IST 173, a full report on the activities of the group will be published in March 2020. It addresses the need to ensure AI and Big Data can be both operational and cost effective. It advocates a mission orientated approach from a multidisciplinary perspective coupling stakeholders and scientists. The need for cross panel communication was emphasised.

Paper 3 Exploitation of Sensor Data Using Artificial Intelligence for Battlefield Sensemaking: V. Lavigne
The paper describes the work undertaken in Canada to augment situational awareness. The emphasis was on labelling objects to use as training data. They identified challenges specific to military applications and encouraged common practice between nations to share best practice.

Paper 4: Ukraine Conflict in Media: Two Approaches to Narrative Analysis: T. Krilavicius The use of natural language extraction and the inference drawn from data obtained during the Ukrainian Crisis. The primary aim was to track and understand public opinion. The need to automate the process became evident. Visualisation of the occurrence of words were displayed from the various sources which conform with the anticipated outcome through reputation and political alliance of the sources.

Paper 5: Developing Transparent Conversational Agent Systems for Intelligence Analysis: S. Hepenstal The presenter has produced a prototype tool to assist analysts to retrieve information using natural interactions. The work modelled the actions of analysts derived by undertaking a cognitive task analysis of four experienced analysts. A practical demonstration of the tool should be available in the near future.

Paper 6: Determining the Effectiveness of Representing Intrusion Detection System Log Files with Visualisation Techniques: G. Thomas The paper is an attempt to optimise the display of data to identify a cyber attack. It is recognised that identifying the structure of an attack in near real time is hard. A comparison was made between a) scatter plot and b) parallel coordinates. The results show a slight preference for the use of scatter plots.

Paper 7: An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It: M. Boardman A paper presenting the findings of HFM-ET-178 looking at the increased use of AI in robotics and the legal, moral and ethical questions which arise. The recommendation that a level of human control is maintained was presented because of the complexity and the adaptability to the situation. The human will however need support tools to ensure meaningful control.

Keynote 2: What's next for Visual Analytics? K. Cook The second keynote speaker presented the flavour of the current position with a view to extrapolation. No major step functions were identified but evolutionary progress predicted. New approaches are envisioned by combining computational techniques with interactive visualisation whilst acknowledging the limitations of human resolution. Aspects of trust to enable a measure of calibration were addressed. No solution was offered but mitigated by the ability to drill into the provenance. Assistance to visualise development and exploration was advocated which would be an extension of the common aids associated with, for example, smart speakers. Visualisation coupled with new relational data libraries may be worth pursuing. The new generation of machine learning will allow humans to create smarter systems removing the mundane elements and aiming for a shared understanding between humans and machines.

Paper 8: Security Perspectives on Social Media Exploitation: B. Forrester The work reports on the activities within IST-177. The number of platforms in which individuals can access social media is extensive and continues to grow with very little regulation this paper provides a snapshot taken a short time ago. Three significant aspects were considered a) the content b) the meta-data and c) the links and networks formed.

Content analysis and natural language processing are both applied in this context to create situational awareness and undertake predictive analysis. The presentation gives evidence of significant advances in natural language processing and identifying sentiments expressed in the media.

Paper 9: Propagation Filters: Tracking Malign Foreign Interventions on Social Media in Real Time: B. Forrester Suspicion had been aroused as to whether foreign agents were influencing public opinion in Canada prior to an election. Models were applied to identify the agents and to isolate the bots to populate the library of known networks. Narratives propagated by the identified agents were similar to those detected in the US elections of 2016. The technique is to drive a wedge between groups to diversify opinion. NATO should ensure that we remain steadfast and promote unity.

Paper 10: People, Bots or Cyborgs? Analysis of a Hashtag War: R. Goolsby Information threats can arise from a combination of artificial and enhanced actors who manipulate community structures, content and code. The aim of the project presented is to create a common framework to analyse the data and assess the potential impact and prepare countermeasures. Attacks have been identified which manipulate perception, emotions, stance and opinion. The counter is to build social trust by engaging with topic groups. The framework 'BEND' has been postulated to build community resilience. Concern was expressed about a amygdala hijack forcing individuals from rational thought. Understanding the mechanism and making early recognition of an attack is crucial.

Paper 11: Towards Big Data in the Tactical Domain: F. Johnsen The paper addresses the benefits and challenges of the Internet of Things (IoT) which can contain physical and virtual objects. It is assumed that in general use the communications will be available on demand, for use in a military environment intermittent service should be anticipated. The rapid expansion of available entities is driven by miniaturisation of both sensors and processing equipment coupled with the power of cloud storage. The military should be able to take advantage of the leverage from commercial products but will require interoperability.

Paper 12: Assessment of IoT Data Ingest Reliability for Urban Environments: J. Michaelis This paper expands on paper11 to encompass the Internet of Battlefield Things (IoBT). Again using the leverage from the development of smart cities in the civilian field. A protocol was selected for use in an urban environment where communications could not be guaranteed with test results presented. Additional research is recommended to establish the necessary network bandwidth, the computing power demanded and the cognitive load on the soldiers.

Paper 13: Storytelling Exploratory Visual Analytics for Counter-Improvised Explosive Device Incidents: V. Lavigne The paper and presentation give the results of using a tool to display IED events in Ukraine. The tool is web based and is readily available. It provides four viewing perspectives: geospatial, data set, incident type and text analysis. The raw data is from a NATO source through work undertaken by SAS-117, a complementary research group to IST within the Science and Technology Organisation (STO).

Paper 14: A Preliminary Experimentation for Meteorological Data Visual Analytics: F. Pisano The paper focuses on meteorological data but could readily be applied to other fields. The chaotic nature of weather makes forecasting difficult but benefits from the extensive observation sites and the accumulation of historic data. The aim is to assist meteorologists predict extreme events to allow adequate preparation to protect life. The tool was demonstrated during the breaks and forms part of the evolutionary progress by refining a relational database. Different views can be displayed allowing the user to interact and focus on specific aspects.

Paper 15: Microbial Genomic Data Analysis for Infectious Diseases: M. Varga Advances in visual analytics has had a positive impact in a wide set of disciplines, this paper identifies the significant advances in diagnosing infectious diseases with the opportunity to assess pathogenesis. The impact on the military is

significant because of the heightened exposure both deployed and in barracks. A library of tools is available to the developers enabling the right information to be available at the right time in the right format.

Workshop 1 Data Challenges.

- Data key factor
- Need to establish infrastructure
- Machine Learning Eco system
- Build common data sets
- Establish repository for data.

Workshop 2 Human Factors.

- AI to assist Decision Making with Big Data
- Training and Education (Gaming)
- Transparent
- Recommend interdisciplinary workshop
- Trust calibration

Workshop 3 Social Media.

- Ethics of Social Media (Define unethical !!)
- Can NATO be proactive? (Define the constraints)
- Information wargame
- Educate regarding Social Media
- Detect inflation of ethnic tension

Workshop 4 Visual Analytics.

- Chaotic data, fractal characteristics (identify transform to display strange attractor)
- Display Risk/ Uncertainty
- Listen to users' requirements
- Display source



SCIENCE & TECHNOLOGY ORGANIZATION
COLLABORATION SUPPORT OFFICE

WORKSHOP

BIG DATA CHALLENGES: SITUATION AWARENESS AND DECISION SUPPORT

IST-178

organised by the

Information Systems and Technology Panel

to be held at

Radisson Blu Beke Hotel, Budapest, Hungary

Tuesday 15 October - Wednesday 16 October 2019

This Workshop is open to NATO nations, PfP nations, Mediterranean Dialogue, ICI nations and Global Partners



ENROLMENTS: Participants are requested to enrol on line at
<https://events.sto.nato.int>

Deadline for enrolments is 1 October 2019

There are no conference fees.

Collaboration Support Office – Point of Contact:

Mrs. Aysegül APAYDIN Tel: +33 (0)1 5561 2282
IST Panel Assistant Fax: +33 (0)1 5561 9626



Background

Information Systems Technology Panel (IST) is one of the seven Panels whose role it is to implement, on behalf of the Science & Technology Board, the STO Mission with respect to Information Systems Technology. The advancement and exchange of techniques and technologies to provide timely, affordable, dependable, secure and relevant information to war fighters, planners and strategists, as well as enabling technologies for modelling, simulation, and training are the focus of this Panel. The Information Systems Technology Panel covers the fields of Information Warfare and Assurance, Information and Knowledge Management, Communications and Networks and Architecture and Enabling Technologies.

Theme - Objectives –Topics

There are many Research Task Groups (RTGs) and Exploratory Teams (ETs) under different NATO Panels and the NATO Modelling and Simulation Group (NMSG) all working on different aspects of defence, military and intelligence challenges, e.g. big data exploration, decision support, mission planning, human factors, medicine, modelling and simulation and visual analytics etc. The IST-178 Workshop invites members of the different RTGs and ETs together in a workshop that will enable them to explore and benefit from each other's work, and to develop collaborations to complement and enhance their respective efforts, as well as facilitating the sharing of data, tasks and tools.

The main objective of this two-day Inter-Panel / Inter-Group Workshop is to provide the means for the RTG/ET members to discuss views, challenges, and work on the use of big data to support the military with obtaining situation awareness and informed decision making.

The aim of the Workshop is to establish a raised awareness of common and different problems, available and foreseen technologies, capabilities and efforts, as well as to support the development of collaborations. This may lead to outcomes such as joint Inter-Panel / Inter-Group collaborations; sharing of data, tools and tasks; and the generation of new research ideas for RTGs and /or ET Technical Activity Proposals (TAPs).

TOPICS:

The theme of this meeting is how the availability of big data may assist the military in understanding the situations they will encounter, and how this may support them in making better decisions. The challenges for development and opportunities for the exploitation of Big Data are likely to vary from domain to domain. This workshop intends to host the discussion for a broad spectrum of potential (military) domains, such as Human Factors, Artificial Intelligence, Information Warfare, Situation Awareness and Decision Support, Visual Analytics and visualization, etc...

WORKSHOP FORMAT and ORGANIZATION:

The workshop will commence with selected presentations to inspire and provide directions for the workshop discussion. We plan an interactive workshop, consisting of several working groups, each with a different theme, topic, and/or domain. The working groups will be arenas for in-depth discussions. Each participant in the workshop will be assigned to a working group, with the aim to have participants from the various communities in each group. These working groups will be formed on the first day. Each working group is to produce a summary report of their discussion, and to present it to the other groups at the end of the workshop. Attendance will be limited to 50 people.

Programme Committee

Programme Co-Chairs

Dr. Margaret VARGA

Seetru Ltd. / Oxford University
United Kingdom

Email: margaret.varga@zoo.ox.ac.uk

Dr. Karel van den BOSCH

TNO Behavioural and Social Sciences
The Netherlands

Email: karel.vandenbosch@tno.nl

Members

Dr. Elena CAMOSSÌ

Center for Maritime Research & Experimentation
CMRE

Email: elena.camossi@cmre.nato.int

Mrs. Francesca de ROSA

Center for Maritime Research & Experimentation
CMRE

Email: francesca.derosa@cmre.nato.int

Prof. Wim MEES

Royal Military Academy
Belgium

Email: wim.mees@rma.ac.be

Col.Dr.Eng. Nikolai STOIANOV

Defence Institute
Bulgaria

Email: n.stoianov@di.mod.bg

Dr. Cathy BOSCARINO

Defence Research & Development Canada
Canada

Email: cathy.boscarino@drdc-rddc.gc.ca

Dr. Bruce FORRESTER

Defence Research & Development Canada
Canada

Email: bruce.forrester@drdc-rddc.gc.ca

Ms. Valérie LAVIGNE

Defence Research & Development Canada
Canada

Email: valerie.lavigne@drdc-rddc.gc.ca

Prof. Bob MADAHAR

Dstl
United Kingdom

Email: bkmadahar@dstl.gov.uk

Prof. Cyril RAY

IRENav, Ecole Navale
France

Email: cyril.ray@ecole-navale.fr

Mr. Albert PRITZKAU

Fraunhofer-FKIE

Germany

Email: albert.pritzkau@fkie.fraunhofer.de**Dr. Carsten WINKELHOLZ**

Fraunhofer-FKIE

Germany

Email: carsten.winkelhorz@fkie.fraunhofer.de**Dr. Andras BENCZUR**

Hungarian Academy of Sciences

Hungary

Email: benczur@sztaki.mta.hu**Prof. Tomas KRILAVICIUS**

Baltic Institute of Advanced Technology

Vytautas Magnus University

Lithuania

Email: tomas.krilavicius@bpti.lt**Dr. Petter BIVALL**

Swedish Defence Research Agency

Sweden

Email: petter.bivall@foi.se**Dr. Liz BOWMAN**

US Army Futures Command

United States

Email: elizabeth.k.bowman.civ@mail.mil**Dr. Alexander KOTT**

Army Research Laboratory

United States

Email: alexander.kott1.civ@mail.mil**Dr. Kristen LIGGETT**

Air Force Research Laboratory

United States

Email: kristen.liggett@us.af.mil**Dr. Michael STREET**

NCIA

The Netherlands

Email: michael.street@ncia.nato.int**Ms. Ivana Ilic MESTRIC**

NCIA

The Netherlands

Email: ivana.ilicmestric@ncia.nato.int**Science and Technology Organization in NATO**

In NATO, Science & Technology (S&T) is defined as the selective and rigorous generation and application of state-of-the-art, validated knowledge for defence and security purposes. S&T activities embrace scientific research, technology development, transition, application and field-testing, experimentation and a range of related scientific activities that include systems engineering, operational research and analysis, synthesis, integration and validation of knowledge derived through the scientific method.

In NATO, S&T is addressed using different business models:

- The Collaborative business model where NATO provides a forum where NATO Nations and partner Nations elect to use their national resources to define, conduct and promote cooperative research and information exchange.
- The In-House delivery business model where S&T activities are conducted in a NATO dedicated executive body, having its own personnel, capabilities and infrastructure.

The Science and Technology Organization - STO

The mission of the NATO STO is to help position the Nations' and NATO's S&T investments as a strategic enabler of the knowledge and technology advantage for the defence and security posture of NATO Nations and partner Nations, by:

- Conducting and promoting S&T activities that augment and leverage the capabilities and programmes of the Alliance, of the NATO Nations and the partner Nations, in support of NATO's objectives;
- Contributing to NATO's ability to enable and influence security- and defence-related capability development and threat mitigation in NATO Nations and partner Nations, in accordance with NATO policies;
- Supporting decision-making in the NATO Nations and NATO.

Acknowledgements

We wish to thank our hosts from Hungary for organising and hosting this event in the facilities they have chosen in their lovely city, Budapest.



<http://www.sto.nato.int>



IST-178 Workshop on Big Data Challenges: Situation Awareness and Decision Support Programme

Tuesday 15 October 2019

08:30 REGISTRATION

09:00 OPENING CEREMONY

Host Welcome Speech

Introduction: Dr.Eng. Michael WUNDER, Chairman IST Panel

Introduction to the Workshop: Dr. Margaret VARGA, GBR

Dr. Karel van den BOSCH, NLD

Administrative Announcements: LtCol. Attila HAJDU, HUN

SESSION 1 – INTRODUCTION

09:30 KEYNOTE SPEECH 1:

Big Data Challenges – Like War

by Dr.Ing. Michael WUNDER, IST Panel Chairman, DEU

10:15 BREAK

SESSION 2– AI and MACHINE LEARNING

10:45 1 Extracting Value from NATO Data Sets through Machine Learning and Advanced Data Analytics
by I.I. MESTRIC, A. KOK, G. VALIYEV, M. STREET,
P. LENK, NCIA, NLD, M. RACOVITA, F. VIEIRA, JALLC,
PRT

11:15 2 Living S&T Roadmap for Artificial Intelligence and Big Data for Military Decision Making
by F. DESHARNAIS, DR&D Canada, CAN, A. BARROS,
P. KERBUSCH, TNO, NLD

11:35 3 Exploitation of Sensor Data Using Artificial Intelligence for Battlefield Sensemaking
by V. LAVIGNE, M. MOKHTARI, M. BRETON, E.
MARTINEAU, J. FOURNIER, DRDC Canada, CAN

11:55 4 Ukrainian Conflict in Media: Two Approaches to Narrative Analysis
by T. KRILAVICIUS, J. MANDRAVICKAITE, BPTI,
Vytautas Magnus University, LTU

12:25 LUNCH

SESSION 3 – HUMAN FACTORS

13:40 5 Developing Transparent Conversational Agent Systems for Intelligence Analysis
by S. HEPENSTAL, DSTL, GBR

14:10 6 Determining the Effectiveness of Representing Intrusion Detection System Log Files with Visualization Techniques
by G. THOMAS, K. LIGGETT, Wright Patterson AFB, USA

14:30 7 Meaningful Human Control of AI-Based Systems: What are the Implications for Situational Awareness and Decision Support?
by F. BUTCHER, DSTL, GBR

14:50 BREAK & SESSION 4 – DEMONSTRATION

15:50 SESSION 5 – BREAKOUT GROUP DISCUSSIONS

17:05 SESSION 6 – DAY SUMMARY

17:25 END

19:00 HOST NATION RECEPTION (for all attendees)

Wednesday 16 October 2019

09:00 KEYNOTE SPEECH 2:
What's Next for Visual Analytics?
by Ms. Kristin COOK, Pacific Northwest National Laboratory, USA

SESSION 7 – INFORMATION WARFARE AND SOCIAL MEDIA

09:45 8 Security Perspectives on Social Media Exploitation
by N. BIETENIECE, NATO STRATCOM COE,
V. DRAGOS, ONERA, FRA, B. FORRESTER, DRDC
Canada, T. KRILAVICIUS, BPTI, LTU, A. PRITZGAU,
Fraunhofer FKIE, DEU

10:05 9 Propaganda Filters: Tracking Malign Foreign Interventions on Social Media in Real Time
by B. FORRESTER, DRDC Canada, Z. DEVEREAUX, A.
BACOVICIN, S. BEDOYA, NEXALOGY, CAN

10:25 BREAK

10:45 10 People, Bots or Cyborgs? Analysis of a Hashtag War
by R. GOOLSBY, US Navy, USA

11:05 11 Towards Big Data in Tactical Domain
by F-T. JOHNSEN, FFI, NOR

11:25 12 Assessment of IoT Data Ingest Reliability for Urban Environments
by J.R. MICHAELIS, J. FREEMAN, A. RAGLIN, US Army

Research Lab., USA

SESSION 8 – VISUAL ANALYTICS AND VISUALIZATION

11:45 13 Storytelling Exploratory Visual Analytics for Counter-Improvised Explosive Device Incidents
by V. LAVIGNE, S. JAYARAM, M. PANGE, DRDC
Canada, CAN

12:15 14 A Preliminary Experimentation for Meteorological Data Visual Analytics
by F.M. PISANO, C. ASSUNTINA, R. GIANPAOLO, CIRA,
ITA

12:35 15 Genomic Data Analysis for Infectious Diseases
by M. VARGA, K. JOLLEY, H. BRATCHER,
C. RODRIGUES, J. MACLENNAN, O. HARRISON,
F. COLLES, J. BRAY, M. MAIDEN, University of Oxford,
GBR

12:55 LUNCH

14:00 SESSION 9 – BREAKOUT DISCUSSIONS

15:00 BREAK

15:20 SESSION 10 – BREAKOUT GROUPS' PRESENTATIONS

16:20 CONCLUDING REMARKS AND WORKSHOP CLOSING

16:30 END OF MEETING

Extracting Value from NATO Data Sets through Machine Learning and Advanced Data Analytics

Ivana Ilic Mestric, Arvid Kok, Giavid Valiyev, Michael Street, Peter Lenk
NATO Communications and Information Agency (NCIA), Data Analytics and Innovation
Oude Waalsdorperweg 61
2597 AK The Hague, NETHERLANDS
Ivana.IlicMestric@ncia.nato.int; Arvid.kok@ncia.nato.int; Giavid.Valiyev@ncia.nato.int; Michael.Street@ncia.nato.int; Peter.Lenk@ncia.nato.int;

Mihaela Racovita, Filipe Vieira,
Joint Analysis and Lessons Learned Centre
Avenida Tenente Martins
1500-589 Lisboa, PORTUGAL
Mihaela.racovita@jallc.nato.int;
Filipe.vieira@jallc.nato.int

ABSTRACT

Exercises and operations generate large quantities of data, but until recently this data was not exploited to enhance decision support and situational awareness, or for analysis and assessment. This paper describes the steps taken to collect, prepare and extract knowledge from such data collected at exercise Trident Juncture in 2018, using advanced analytics [1] and machine learning [2] to extract high level perspectives. Results show initial macro level assessments of exercise performance including response times; visualisation of information flows across locations and systems; and identification of other items of interest in this “big data” collection.

The paper also presents results using natural language processing [3] to identify situations and incidents of interest to the lessons learned community from exercises and operations. The data sources for this assessment are both the raw data from exercises as well as from lessons learned reports about exercises and operations.

The paper also describes the technical challenges of using advanced data analytics and machine learning on large, classified data sets, together with solutions for the hardware and software challenges encountered.

1.0 INTRODUCTION

The work described in this paper makes use of two distinct data sets; one from exercise Trident Juncture 2018 (TRJE 2018) and another of reports from the NATO Lessons Learned Portal (NLLP). During exercise Trident Juncture, NCIA stored and repatriated around 3.9 TB of data from six different locations operating on two security domains. This heterogeneous collection consists of data relating to:

- Collaboration services (over 23,000 chat conversations)
- Functional area services for: Land, Maritime, Air and Joint Command & Control systems, Planning Logistics, Intelligence, Geographic Information Systems, Incident management, Collaboration Core services: Exchange (Email), Tasker Tracker, Active Directory, SharePoint Portal, Document Handling System
- Exercise management: 1161 confirmed events & 373 injects

Extracting value from Large Exercise Data Sets through Machine Learning & Advanced Data Analytics

Figure 1 lists the file sizes and file counts which make up this data set.

FAS File Size (GB)		FAS File Count	
FAS	Size GB	FAS	Count
lc2is	14,655.51	lc2is	1625361
logfas	2,589.38	jchat	1202972
sql server	1,177.87	logfas	288510
ncop scc	190.21	ncop scc	132434
exchange	178.31	coregis	73130
sharepoint	139.58	jocwatch	46193
jocwatch	129.60	domain controller	6311
coregis	110.77	sharepoint	4620
jchat	39.38	exchange	3778
domain controller	36.96	sql server	1890
intel-fs	8.67	intel-fs	647
niris	7.97	niris	420
icc	3.60	icc	312
Total	19,267.82	Total	3386578

Figure 1 File Sizes & Counts repatriated from TRJE 2018

The NLLP is a repository of over 400 documents which capture lessons learned, best practices identified and other assessments of NATO activities in operations and exercises. This repository also provides a rich dataset where data science can be applied to generate valuable insights, and enable topics and trends to be extracted.

Two toolsets were used for this work, in some cases both were applied to address similar problems, allowing a comparison to be made between the toolsets. One tool was IBM Watson Explorer; the other was a combination of the open-source KNIME analytics platform [4], in conjunction with Microsoft's PowerBI for data visualisation.

2.0 EXERCISE ANALYSIS

Initial analysis of data collected from Trident Juncture focussed on data generated by functional area services (FAS) used for intelligence (IntelFS) and for collaboration (JCHAT). Analysing this data at an exercise level revealed a high-level picture of how these tools were used during the exercise to support analysis and reporting.

2.1 Analysis of collaboration services

Analysis of all messages sent during the exercise using the JCHAT (joint tactical chat) service for collaboration showed that there were 23,123 unique group conversations, involving 693 unique users. Applying language recognition to the messages showed that while the majority of the messages were in English, a significant portion were in French, Norwegian, Italian and Swedish plus very limited amounts of other NATO and Partner languages.

Activity patterns were generated by comparing activity through the day and the week, as shown in Figure 2. This shows that while collaboration takes place throughout the week and largely reflects the exercise schedule with most activity taking place during normal working hours, but some coordination activity continuing into the night. This figure also shows the most active organisational elements participating in most JCHAT conversations.

Extracting value from Large Exercise Data Sets through Machine Learning

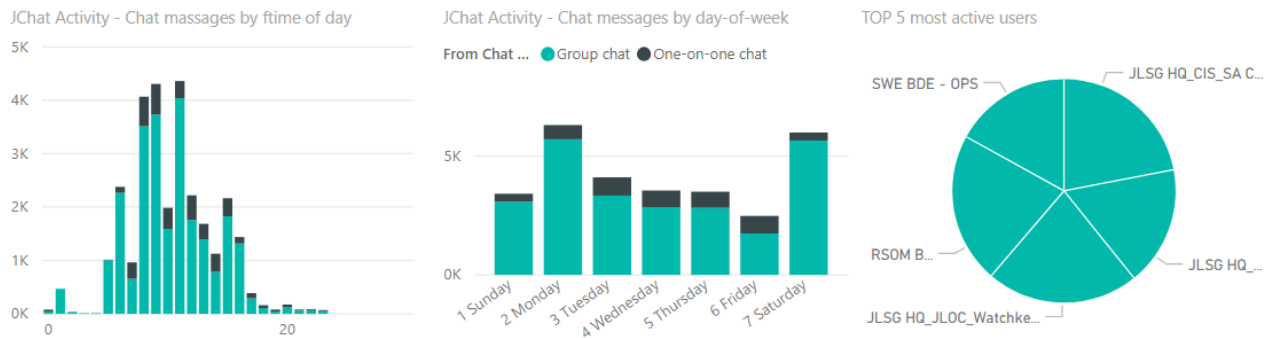


Figure 2: JCHAT activity

Figure 3 shows the level of JCHAT activity – for collaboration – throughout the exercise (upper plot) against the overall activity in other functional area services (lower plot). While there is some level of correlation, to be expected as coordination often precedes action, these results merit more detailed analysis.

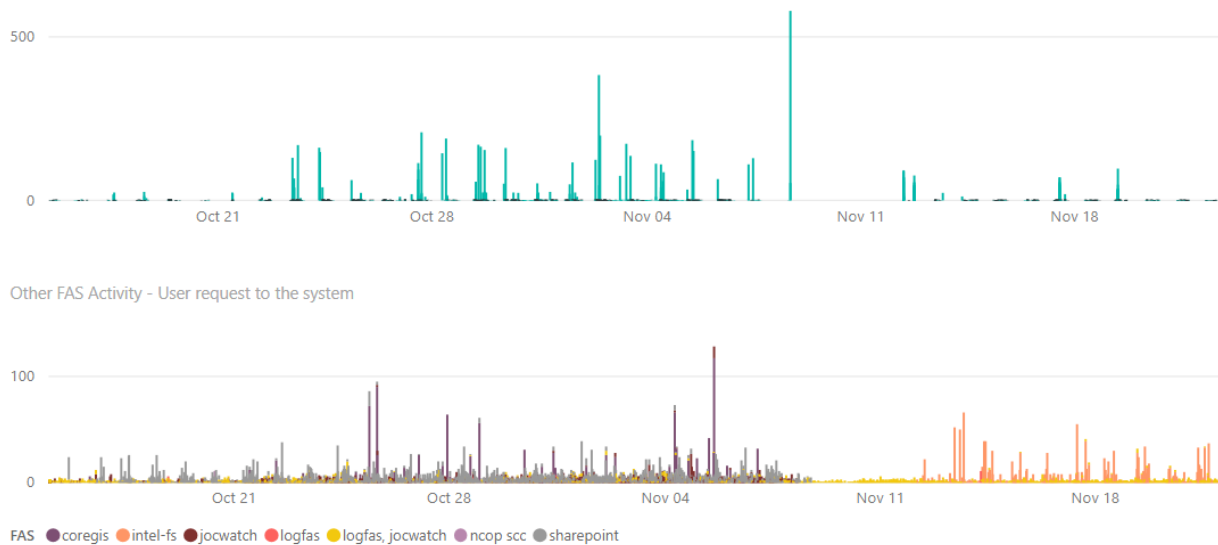


Figure 3: Activity levels for collaboration and other functions during the exercise

2.2 Analysis of information flows in IntelFS

Initial analysis of information flows in Intel FS focussed on requests for information and the associated response. Figure 4 shows the main sources of intelligence requests for information (RFI), the main responders and the routing. It is unsurprising that the main source of responses is the Joint Warfare Centre, which is responsible for running the exercise. Figure 4 also shows that the majority of requests are routed through the Joint Force Command in Naples (JFCNP), or through its Mission Information Room (JFCNP-MIR). On the right of Figure 4 are also shown the prime sources and consumers of RFIs.

Extracting value from Large Exercise Data Sets through Machine Learning & Advanced Data Analytics

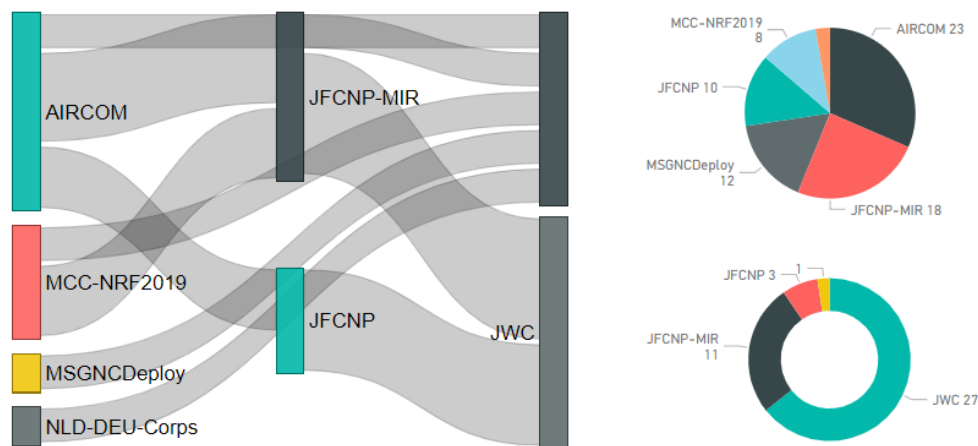


Figure 4: Information flows in IntelFS

Natural language processing (NLP) and machine learning (ML) techniques were used to process a number of sources of text data including intelligence reports and requests for information (RFIs), JChat conversations and Exercise events. NLP was able to extract entities e.g. locations, people, organisations and times. From these entities, independently extracted from exercise information in many different systems, it was possible to recreate events which occurred during the exercise and which included the same entities (location) and actors (individuals, organisations) all being reported within the same time period. These are visualised in a dashboard of Figure 5, where source reports are shown in the top left, content of a selected report is shown in the top right, relationships between entities – such as the location and people – are shown in the bottom left, linked by bars which are proportional to the strength of relationship identified in the data. The bottom right corner shows all entities extracted from the selected documents. In Figure 6 we can see the same entity (Arvidsjaur) extracted from conversations in JChat collaboration service during the same time period.

Figure 5 shows entities identified by entity extraction and machine learning techniques as being relevant at a particular time. At this time, the exercise featured an event at Arvidsjaur airport involving a number of key actors shown in the diagram. This demonstrates the ability of machine learning to find entities and relationships of relevance to assess – or conduct – the exercise.



Figure 5: Named entity extraction Intelligence system

Extracting value from Large Exercise Data Sets through Machine Learning

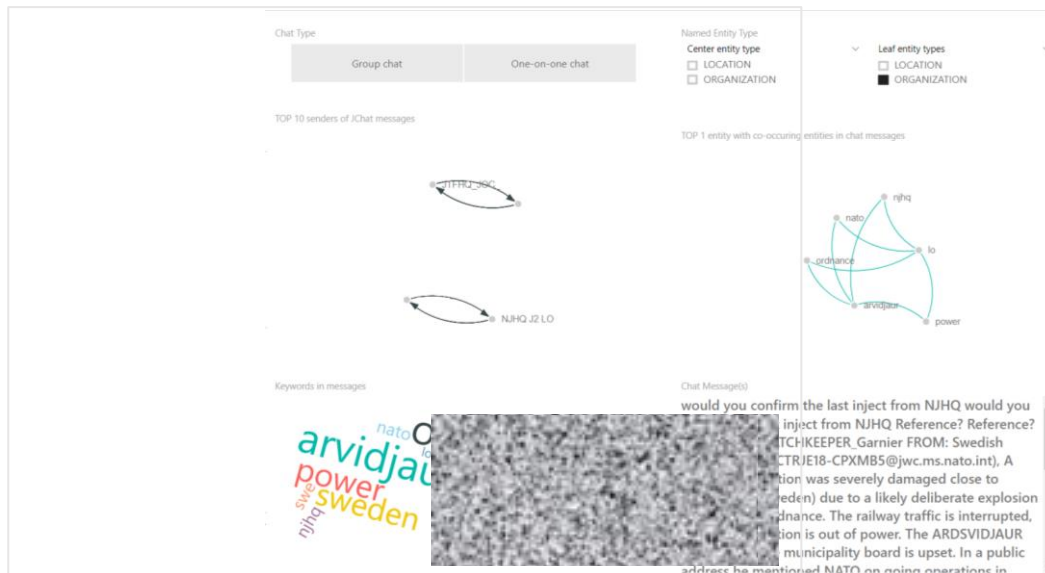


Figure 6 Named Entity extraction from collaboration service (JChat)

3.0 MINING TEXT FOR LESSONS LEARNED

Manually extracting lessons learned from such a huge exercise is a resource intensive activity. Data science tools offer the possibility to handle datasets with large volume and variety, to speed up the process and to sift through large quantities of material, highlighting items worthy of further investigation by lessons learned analysts.

In addition to the FAS data described above, exercises generate large quantities of text, from formal documents, email and text collaboration tools. Text data from Trident Juncture has also been analysed, taking data from a number of sources:

- Documents (from SharePoint)
- Emails and email attachments
- Skype for Business (chat) conversations
- JChat conversations

Results from analysis of classified data are not suitable for publication in this format. But similar text analytics has been performed on unclassified text documents from the NLLP.

3.1 Extracting Lessons Learned from reports

NCIA, in conjunction with the JALLC, compared two toolsets for text mining, analysis and visualisation. These were IBM's Watson Explorer (WEX) and KNIME Analytics platform coupled with Microsoft Power BI. The dataset for this activity comprised 397 lessons learned reports from the unclassified NATO Lessons Learned Portal (NLLP). The majority of these documents were PDF reports.

In workshops held between JALLC analysts and NCIA data scientists, several data science techniques were identified to help analysts extract valuable information from this large document collection. These covered

Extracting value from Large Exercise Data Sets through Machine Learning & Advanced Data Analytics

NLP (Natural Language Processing), ML (Machine Learning) and advance analytics techniques, with the most useful being:

- Sentiment analysis [5] of documents – help analysts to identify Lessons Learned (strong negative sentiment) and Best Practices (strong positive sentiment)
- Topic and trend Analysis for selected entities
- Entity recognition – help analysts to identify the most frequent entities (people, organizations and locations) over period of time and documents
- Extraction of common phrases
- Relationship identification between entities and terms in documents

The initial task was to identify key topics and trends in lessons learned reports. This was done using a combination of entity extraction, custom dictionaries related to sentiment around lessons learned as well as organisations and locations of interest. Results were visualised to show prevalent terms over time; with the ability for lessons learned analysts to find individual references in documents in order to support their SME-driven assessment. Results from WEX are shown in **Error! Reference source not found.** and from KNIME, visualised in Power BI, in Figure 8.

Error! Reference source not found. shows the distribution of documents over a selected period of time for a list of topics which can represent entities, phrases, keywords and dictionary terms. The trend analysis shows sharp and unexpected increases in frequency of the topics over time. The trend index (the dotted line) measures the ratio of actual frequency of a topic compared to the expected average frequency based on the previous data. This expected change in frequency is estimated by using a modified Poisson distribution and can quickly highlight anomalies for a number of topic values on a single window, giving lessons learned analysts a clear indication of changes in importance of specific topic e.g. locations, activities, organisations etc.

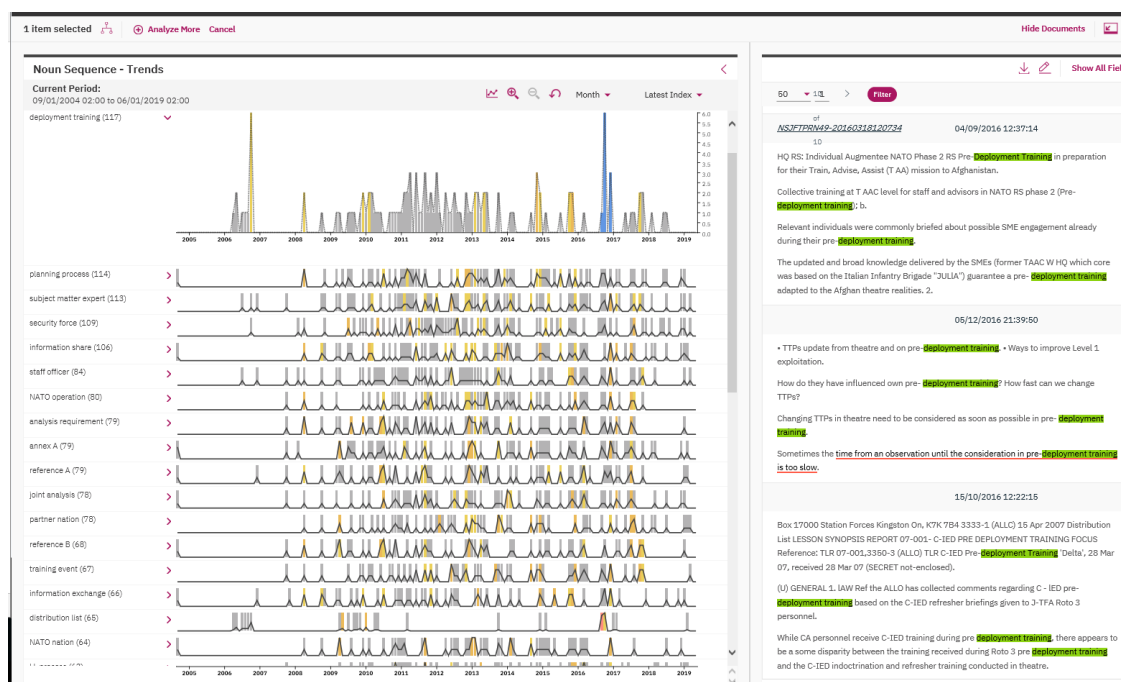


Figure 7 Topic and trend analysis (IBM WEX)

In **Error! Reference source not found.** the grey bars show the actual frequency of a topic, unexpectedly high frequencies are highlighted in yellow. A given period can be selected (here highlighted in blue) and then

Extracting value from Large Exercise Data Sets through Machine Learning

documents from that period will be displayed on the right of the screen, with the current topic/term highlighted (in green). In Figure 7 we can see described scenario for selected topic e.g. “deployment training”. This provides a rapid mechanism to identify periods of interest within the dataset for a given topic, and to rapidly home in on the relevant documents and text.

Figure 8 shows a trend analysis of terms associated with best practices (top) and lessons identified (bottom) across 15 years of lessons learned reports. The bar charts on the right of this figure show simplistic terms which indicate a lesson identified or a best practice to be adopted e.g. “maintain”, “facilitate”, “need” etc. The bar chart on the far right shows the documents with the greatest prevalence of such terms. Without data science tools and custom dictionaries tracking such trends and identifying key documents would be almost impossible on such a document set.

Although the visualisation appears superficially different, the more interesting aspect to both the data scientist and lessons learned analyst is the underlying information feeding that dashboards. Both tools were able to automatically extract valuable insights, such as entities, expressions and phrases accompanied with frequency and relevance measure. Apart from using custom dictionaries, which help analysts to focus on area of interest, tools were able to extract new knowledge and insights not yet known to analyst through functions for open information extraction such as subject-verb-objective analysis.

The ability to perform and display such analysis across a dataset of over 400 documents spanning many years is difficult for human analysts alone to perform. Data science tools provide an ability to search (and remember) volumes of material which humans alone cannot, to identify “hot topics” and to display trends. All of these functions bring additional value to the lessons learned process.

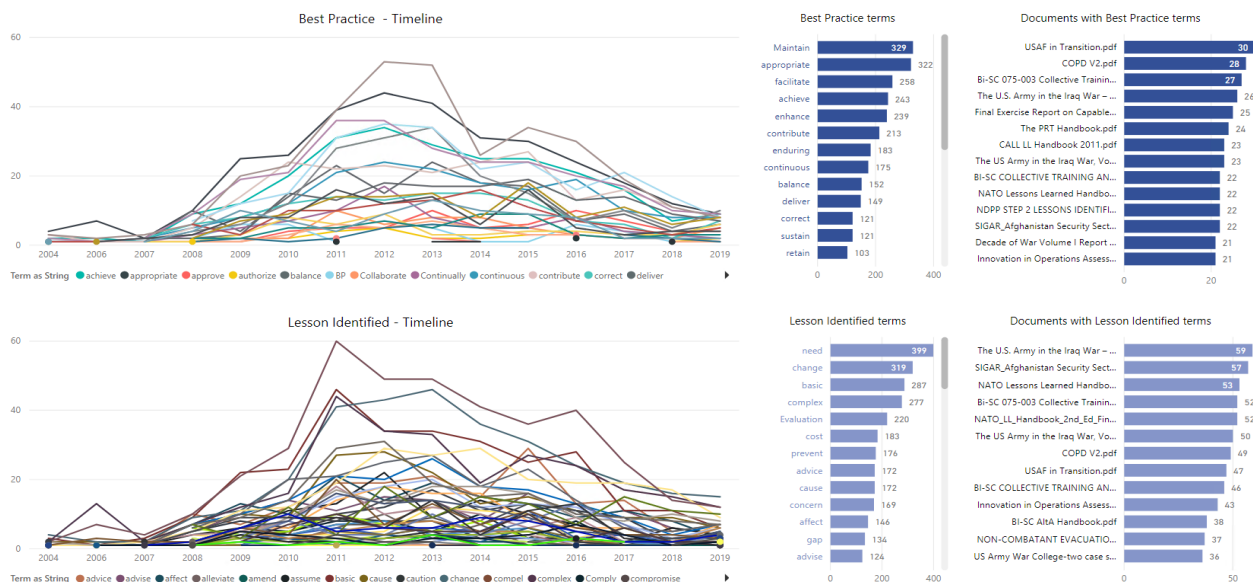


Figure 8: Topic and trend analysis (KNIME & MS Power BI)

5.0 CONCLUSION

Big data is often referred to as having four “V” characteristics, of having volume, variety, velocity and veracity, which are each too great for a human analyst to address. The datasets described in this work certainly demonstrate high volume and variety. The initial work described in this paper shows that data science tools and techniques have the ability to process large datasets – such as document sets spanning several years, or

Extracting value from Large Exercise Data Sets through Machine Learning & Advanced Data Analytics

exercise datasets of terabytes;- and to extract value from them – such as identification of trends and relationships of entities as well as numerical analysis at scale to show information flows.

Data collection and preparation has been a challenge in this area, as in almost every application of data science. But the time taken to prepare data for machine learning and analysis is low, compared to the impact on the lessons learned process. The ability of data science tools to process a variety of structured (databases), semi-structured (logfiles) and unstructured (email, documents) data and apply comparable techniques to them also provides a new approach to the extraction of value from large datasets.

Initial results of combining different data sources from an exercise (cross-domain analysis), result in more valuable insights and reveal connections and collaboration between different units which may be difficult to identify without big data analytics of all exercise data. Examples of this include the relationship between collaboration and intelligence activities in the exercise, or the automatic extraction of entities and their relationships which can reproduce events which occurred during the exercise.

Data science tools to identify and visualise key terms within large document sets have also been demonstrated to provide value to the lessons learned community, discovering and displaying trends in lessons identified and in best practices employed.

However, further work is needed in this area, for example the use of different languages for collaboration noted in section 2.1 does not itself prove the need for machine translation tools in NATO exercises unless further analysis of entities extracted from non-English conversations shows that the same exercise-related subjects are being coordinated in multiple languages. Additional work is also planned to extend the initial analysis of lessons learned material to operate on multi-word concepts rather than just on single words.

Further work is planned in this area as the authors continue to refine the data science techniques, tools and workflows needed to extract maximum value from large datasets in order to improve the effectiveness of NATO's existing activities and apply data science to better exploit NATO data.

ACKNOWLEDGEMENTS

This paper and the work within it would not be possible without the support of many colleagues at NCIA to collect the data from Trident Juncture, or the support of colleagues at JALLC, in particular Ms Jackie Eaton.

REFERENCES

- [1] Street M., Lenk P., Ilic-Mestric, I. and Richter P., "Lessons learned from initial exploitation of big data and AI to support NATO decision making", STO experts meeting on "Big data and AI to support military decision making", Bordeaux, 2018.
- [2] M. Richter, M. Street & P. Lenk, "Deep Learning NATO document labels: a preliminary investigation", Int. Conf. on Military CIS, Warsaw, May 2018
- [3] Kok A, Ilic Mestric I and Street M. "Named entity extraction in a military context", Int. Conf on Military CIS, Budva, May 2019.
- [4] M. R. Berthold, N. Cebon, F. Dill, T. R. Gabriel, T. Kotter, T. Meinl, P. Ohl, C. Sieb, K. Thiel, B. Wiswedel, "KNIME: The Konstanz Information Miner," in Studies in Classification, Data Analysis, and Knowledge Organization, Springer, 2007.

Extracting value from Large Exercise Data Sets through Machine Learning

- [5] R. Blunt, C. Riley, M. Richter, M. Street, D. Drabkin, “Using data analytics and machine learning to assess NATO’s information environment”, IST-160 specialists meeting on *Big data and artificial intelligence for military decision making*, Bordeaux, May 2018.



Exploitation of Sensor Data using Artificial Intelligence for Battlefield Sensemaking

Valérie Lavigne Marielle Mokhtari Mélanie Breton Étienne Martineau Jonathan Fournier

valerie.lavigne
@drdc-rddc.gc.ca

marielle.mokhtari
@drdc-rddc.gc.ca

melanie.breton
@drdc-rddc.gc.ca

etienne.martineau
@drdc-rddc.gc.ca

jonathan.fournier
@drdc-rddc.gc.ca

Defence R&D Canada
2459, route de la Bravoure, Québec (QC)
CANADA

ABSTRACT

Although Artificial Intelligence (AI) techniques based on Machine Learning (ML) have shown a lot of progress in the recent years, leveraging these capabilities for military operations is still a challenging endeavour. At Defence Research and Development Canada (DRDC), we perform R&D activities aiming at leveraging ML for analyzing and making sense of a variety of sensor data. We use computer vision algorithms for analyzing Electro-Optical (EO) and Infrared (IR) Full Motion Video (FMV), 3D Light Detection and Ranging (LiDAR) and Wide Area Motion Imagery (WAMI) sensor data.

Academic AI research and datasets are often not adapted to military sensors and operational conditions. This leads to several challenges when transferring promising technologies to military applications. For instance, we need to handle noisy and blurry images, to process in real-time sensor modalities other than visible EO imagery, and to work with point clouds having different spatial resolutions.

In this paper, we discuss our work on creating military relevant labelled datasets, merging and aligning multiple training datasets, optimizing models for real-time analysis, and defining military relevant evaluation metrics.

GETTING LABELLED DATASETS FOR MODEL TRAINING

The first challenge that is faced when we want to train a model for object detection is the lack of labelled data. Academic models are often trained on large datasets such as ImageNet [1] and OpenImage [2], which both contain over 1 million images. For a dataset to be relevant for training our detection models, it should:

- Contain labelled instances of the objects of interest;
- Offer a perspective similar to the imagery it is intended for (ground-based vs. aerial perspective); and
- Have similar image characteristics (visible spectrum vs. infrared imagery, image degradation).

Large datasets that meet all of these requirements are not available. In order to create our training datasets, we decided to adopt a mixed approach. First, we combined multiple available open source datasets. Second, we performed labelling on our own sensor data to cover the visual aspects that were missing from the open source datasets.

Semantic Alignment for Leveraging Multiple Open Source Datasets

Our first approach was to leverage the open source datasets that are available. Merging multiple datasets comes with the following challenges:

- Multiple labelling formats are used by different datasets;
- Each dataset has its own label names and definitions; and finally,
- Some objects that are labelled in some datasets, but are present in other datasets but not labelled in the latter.

In order to handle the different file formats, we opted to create an intermediary labelling file format where we standardized all the labelling information, but kept the original label names. This allowed us to create configurations specific to each model to be trained. For each model, we specify which datasets should be included in the final training dataset and the desired label conversion to apply. This conversion states which labels from the original datasets should be kept and the mapping to our model object types. With this method, we can create a separate set of label files that is specific to each model. For now, the semantic alignment of the various labels is performed manually, but eventually, a measure of similarity based on detection confusion matrices and visual similarity [3] could be used to partially automate the label conversion process.

Labelling Sensor Data

The specific objects that we would like to detect with our models include military vehicles and equipment which are rarely present in open source labelled datasets. In addition, a vast majority of the datasets available offer ground-based visible imagery. Datasets that contain aerial imagery and other sensor modalities such as infrared can be found, but they are relatively modest and they usually do not combine all the aspects necessary to train the desired model. Although, it has been shown that Generative Adversarial Networks (GANs) can be trained to convert imagery between sensor modalities [4], this was performed from ground-based infrared to visible imagery and the performance is always lower than what we can obtain with a model directly trained on the source modality. Moreover, there is generally not enough labelled data to create GANs for each conversion that we need. This meant that we add to collect our own sensor data and label the objects of interest ourselves.

Labelling objects in images is a tedious and mind-numbing activity. Many open source datasets were created through crowdsourcing sites. We did not adopt this approach because some objects require military expertise to identify correctly.

As we were labelling video data, we speed up the labelling by copying the labels from the previous frame to the next, so that the user only needs to slightly adjust the boxes' localization and size. We were helped by two full-time military personnel for 4 weeks. During that time, around 40 000 objects bounding boxes were labelled in more than 120 000 images. For future labelling sessions, we intend to explore active learning to train a model in parallel with the labelling activities. This model could then suggest labels that users only need to confirm the localization, instead of having to draw all the object bounding boxes.

DEEP LEARNING PROCESSING PERFORMANCE

Gathering training, validation and test datasets, and training models are only the first steps. After that, we need to be able to apply those models in real time to sensor feeds. Both WAMI and FMV feeds are very large and require the development of optimization strategies, if we want the object detection, localization and classification

to happen in real time.

Model Architecture

Selecting the deep neural network architecture has a significant impact on model execution speed. Early on, we decided to use bounding boxes. The model architecture we selected was YOLOv3 [5]. Part of its high-speed performance comes from performing object detection, localization and classification in a single pass, whereas some other models do this in two passes. Some models that achieve better accuracy, but they cannot process information in real time. For most operational contexts, it is better to have imperfect timely detections rather than perfect detections too late.

Tiling Strategy and Object Tracking

To keep YOLOv3 fast, we used the medium detector resolution which is 416x416 pixels. The default approach is to reformat the input image to fit this smaller detector resolution. This causes small objects to sometimes become so small that they cannot be detected anymore, especially in an aerial operating context. In order to ensure that we exploited the full resolution of the sensor input, we opted for a tiling strategy, where we moved our model detection to different parts of the image for each frame. As we are analyzing video feeds, we can leverage temporal continuity between frames for the other parts of the image that are not been analyzed. Objects detected in the image are tracked between frames until the detection model comes back to the area where that object is and confirm its presence again. This strategy has allowed us to achieve real-time analysis (average of 25 fps) on a video feed on a single laptop equipped with a NVIDIA P5000 GPU card.

Detection and Tracking in WAMI Imagery

Images collected with a WAMI camera system proved to be challenging to process using available ML approaches. The frame rate of those systems is very low (e.g. 2-3 Hz) and the image resolution is generally in the order of 0.5 m to 1 m, depending of the selected altitude. In addition, images are generally quite large (e.g. 10,000 x 10,000 pixels) and they need to be processed using a tiling approach as described in the previous subsection.

Traditional ML object detection techniques have been proven to have a low performance with WAMI imagery. This is due to the fact that the number of pixels for each target is generally very low and the contrast of the target with respect to the background is poor. For this reason, hybrid approaches have started to emerge where motion and appearance change detection techniques are combined with a convolutional architecture to perform reliable detections [6]. This type of approach is what DRDC plans to implement into its current WAMI image processing chain.

It is envisioned that good quality datasets will be required to assess the performance of the implemented approaches. As of now, there are a limited number of annotated WAMI datasets openly available. To address this issue, it is planned to rely on manually annotated datasets collected by DRDC and its partners, as well as datasets generated by DRDC's airborne simulation software tools.

Point Cloud Classification

LiDAR laser scanners are the most common instruments used to collect geographic Point Cloud Data (PCD). Even if point cloud classification is an active research study area, the number of tools available for this classification is limited. Popular tools require a paying license in order to be used, and free of use tools are quite

rare. Also, these tools do not necessarily produce the best quality results and most of the time, the choice of labels are limited to ground, building and high vegetation.

Due to their different spatial resolution and number of points, Aerial Laser Scanned (ALS) and Terrestrial Laser Scanned (TLS) need to be classified separately. For example, objects of interest such as people, panel signs and bicycles cannot be extracted from ALS PCD, due to their small spatial resolution, but they can be extracted from TLS PCD. In ALS point clouds, the building roofs are being extracted, but in TLS it is often hard to extract roofs due to their field of view and range limits. ALS PCD can be classified using classic ML algorithms in less time and with fairly good results. On the other hand, TLS PCDs usually require more learning to segment labels and Deep Learning (DL) algorithms work best on this type of data.

We first developed two classification models in order to automatically classify ALS point clouds in urban and rural areas, respectively. ALS classification can be done with standard ML algorithm such as Random Forest. Classification results on ALS point clouds show a 90% global accuracy for the following classes: ground, vegetation, roof, façade, and car.

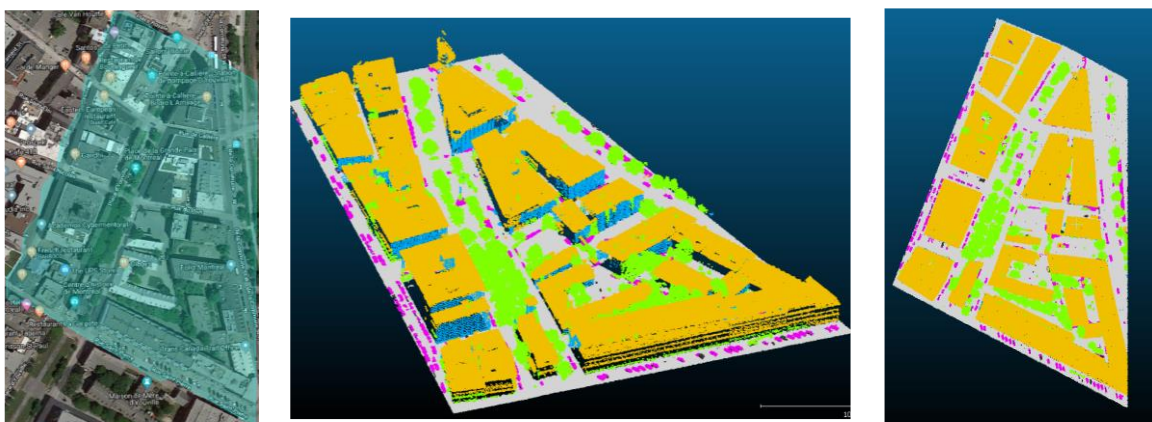


Figure 2: ALS point cloud classification in an urban area (Old Port Area, Montreal). Ground (grey), Vegetation (green), Roof (yellow), Façade (blue), Car (pink), and Unclassified (black)

Second, we investigated automatic classification for TLS point clouds using DL algorithms. Classification results on TLS point clouds show a 95% global accuracy for the following classes: man-made terrain, natural terrain, high vegetation, low vegetation, building, hardscape, scanning artefacts and car. We trained two models for automatic classification of TLS point clouds, one using the colours (Red-Green-Blue) and one without. Out of the two, the model with colour produced the best results.

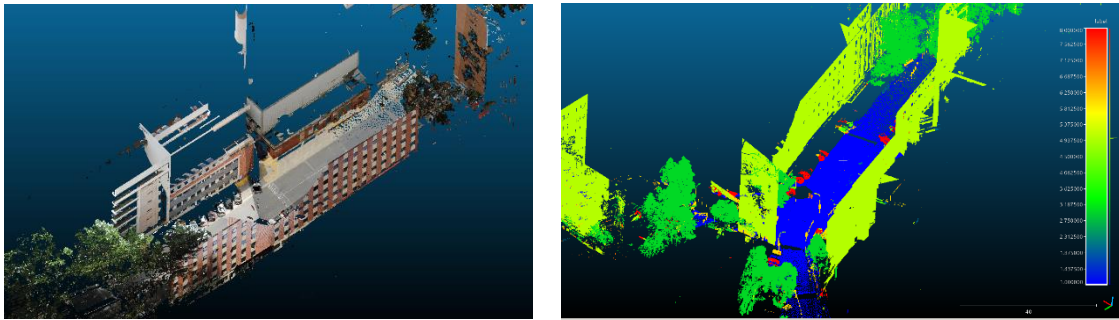


Figure 3: TLS point cloud classification (Adelaide Royal Hospital, Adelaide, Australia) – Man-Made Terrain (blue), Natural Terrain (teal), High Vegetation (dark green), Low Vegetation (green), Building (light green), Hardscape (yellow), Scanning Artefacts (orange) and Car (red).

EXPLOITING AI GENERATED INFORMATION

The exploitation of AI-generated content for sense making is highly application problem specific. We believe this capability can enhance situational awareness. In some cases, alerting strategies can be devised in order to keep an analyst informed of changes in a situation. Automated object counting may produce interesting information for intelligence analysis. Using automated detection and classification can allow operators and analysts to monitor multiple feeds at the same time. Analyzing and tagging large collection of video imagery can allow an analyst to find relevant imagery faster, and this is a tool that we intend to develop.

PERFORMANCE EVALUATION

The mean Average Precision (mAP) metrics [7] is well known in the machine learning community. It is the go-to metrics for evaluating object localization algorithms, and it works well for comparing new models with known models. However, selecting an algorithm to be deployed based on this sole criterion has several shortcomings.

First, the mAP usually evaluates localization-detection success on a specific dataset. It does not represent the detector's performance "in-the-wild", where the sensor might produce noisy, rotated, compressed, and/or blurry images. Also, it does not differentiate between the types of objects, although that could be an important feature if we want to favour performance on specific mission-relevant object categories. Finally, it penalizes highly an algorithm when an object is detected more than once. This hypothesis works well when the intended task relates to object counting within an image. However, if the task requires rare object detection or simply raising a flag when an anomaly is detected, the mAP might discard good algorithms.

For example, let's look at the model output pictured in Figure 4. The algorithm performs poorly at counting each instance of the vehicles class. However, if the task was to raise an alarm in real time, with a noisy and heavily compressed image, then this model might outperform other models.



Figure 4: The bounding box surround two vehicles at a time. For an object counting task perspective, this algorithm performs poorly, however, for a detection task, this algorithm performs correctly.

Problem Specific Metrics

We are in the process of developing specific metrics that include these three criteria: classification accuracy, localization accuracy, and scaling of the bounding box. To that, we add robustness to noise, sensor orientation, and image compression as well as inference time. In the context of Intelligence, Surveillance and Reconnaissance (ISR), these criteria should be weighted depending on the task.

CONCLUSION

In this paper, we discussed the challenges of applying ML-based AI to military problem domains and shared some of our lessons learned in working on different projects involving various types of imagery, as well as point cloud data. There is a need for developing common practices between NATO nations to ensure that we can share and leverage datasets, and that we have military relevant metrics for comparing model performances on common sets of operational problems.

ACKNOWLEDGEMENTS

The authors would like to thank Guillaume Gagné for his involvement in the labelling activity, Maxime Dionne for his technical advices, and Frédérick Lafrance for his work on the point cloud classification prototypes.

REFERENCES

- [1] ImageNet Summary and Statistics, ImageNet. <http://image-net.org/about-stats>. Retrieved 4 September 2019.
- [2] Kuznetsova, A., Rom, H., Alldrin, N., Uijlings, J.R., Krasin, I., Pont-Tuset, J., Kamali, S., Popov, S., Mallocci, M., Duerig, T., & Ferrari, V. (2018). The Open Images Dataset V4: Unified image classification, object detection, and visual relationship detection at scale. ArXiv, abs/1811.00982.
- [3] Brust, C., & Denzler, J. (2018). Not just a matter of semantics: the relationship between visual similarity and semantic similarity. ArXiv, abs/1811.07120.
- [4] Berg, A., Ahlberg, J., & Felsberg, M. (2018). Generating Visible Spectrum Images from Thermal Infrared.

2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 1224-122409.

- [5] Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. ArXiv, abs/1804.02767.
- [6] LaLonde R., Zhang D., Shah, M. Fully Convolutional Deep Neural Networks for Persistent Multi-Frame Multi-Object Detection in Wide Area Aerial Videos. ArXiv:1704.02694v2
- [7] Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S., Huang, Z., Karpathy, A., Khosla, A., Bernstein, M., Berg, A. C., and Fei-Fei, L. (2015), ImageNet Large Scale Visual Recognition Challenge, International Journal of Computer Vision (IJCV), 115(3). Challenge website [image-net.org/challenges/LSVRC/2017].



Ukrainian Conflict in Media: Two Approaches to Narrative Analysis

Justina Mandravickaitė
Tomas Krilavičius
LITHUANIA

tomas.krilavicius@bpti.eu

ABSTRACT

Internet media is one of the most important tools to influence public opinion as well as reflect it. In this paper we analyze reflection of dynamics of Ukrainian conflict in BBC, RussiaToday, DayKiev and delfi.lt (main Lithuanian news portal). We apply two different approaches for the analysis: co-occurrence networks analysis to reflect change of rhetoric in four different media channels during conflict and sentiment-based storyline (syuzhet) analysis to monitor sentiment change in BBC from 2013 to 2014. We split conflict into three stages: beginning (Nov 21, 2013 – Jan 15, 2014), escalation (Jan 16, 2014 - Feb 17, 2014) and occupation of Crimea (Feb 18, 2014 – Feb 28, 2014). These approaches allow visual analysis of the conflict dynamics in media. Such application of Artificial Intelligence, Natural Language Processing and visualization techniques for big data allows better understanding of reflection of conflict dynamics and public mood on specific topics, automation of information analysis.

1.0 INTRODUCTION

In narrative analysis it is essential to understand the sequence in which events occur [1, pp. 2-16]. Recent studies on narrative analysis focus on different types of events, e. g. linguistic-based narrative analysis of YouTube accounts [2]. It was used to describe common attributes of the narratives as well as to identify a list of shared thematic and linguistic characteristics. Study [3] reported how narratives were used by extremist organization. The researchers built a network of stories linked according to their semantic similarity. Moreover, it explored how narrative in social media shapes the vision of future by justification of present and explanation of past [4]. This study also showed how interconnected narratives provide intent and justification for different target audiences. Yet another study presented a case and usage of individual narrative [5]. According to this study, individual narratives help people to make sense of the world as well as shape and express their ideology, including the political functions.

Usually narratives present events in chronological order, e.g., novels, personal retelling of experiences, etc. However, in news articles narratives differ from other types of narratives as they follow a complex time structure. In producing news stories, journalists present the events in “instalments”, thus an event that was introduced in the earlier parts of a story may be described in detail later, sometimes in multiple instances [6, pp. 147-174]. Therefore, events in news stories are usually presented in a non-chronological order, i.e., in the news stories storyline or syuzhet often does not match the chronological order of the events.

In this paper we analyze reflection of dynamics of Ukrainian conflict in BBC, RussiaToday, DayKiev and delfi.lt (main Lithuanian news portal). We apply 2 different approaches for the analysis: word co-occurrence networks to reflect change of rhetoric in four different media channels during conflict as well as sentiment-based storyline (syuzhet) analysis to monitor sentiment change in BBC from 2013 to 2014.

2.0 DATA

We used 2 datasets for our research:

1. results of qualitative discourse analysis of media articles (BBC, RussiaToday, DayKiev and delfi.lt),
2. raw BBC articles.

The first dataset was prepared by the team of students mentored by scientists during the project “Research Meadow / Mokslo pieva”. Media articles were collected from 4 different media sources: BBC, RussiaToday, DayKiev and delfi.lt, from November 21, 2013 till February 28, 2014. This period covered 3 stages of the Ukrainian conflict, see table 2-1 for some of the more important events for each of the Ukrainian conflict.

Table 2-1: Some of the more important events during 3 stages of Ukrainian crisis.

Conflict stage	Events
1 st (2013 11 21-2014 01 15)	● V. Yanukovich postpones the signing of an association agreement with the EU. ● Y. Tymoshenko is prevented from leaving Ukraine. ● Large-scale protests begin. ● Over 100,000 people participate in weekly demonstrations in Kiev. ● First clashes with riot police and arrests occur. ● In early December, over 800,000 people participate in demonstrations in Kiev.
2 nd (2014 01 16-02 17)	<u>January 16-23</u> ● The Supreme Rada adopts anti-demonstration law. ● First deaths. ● Raid of government buildings in western Ukraine. <u>January 28-29</u> ● The Prime Minister resigns. ● The laws directed against the protests are annulled.
3 rd (2014 02 18-28)	● Clashes begin in earnest: casualties on both sides – protesters as well police. ● Agreement between opposition and V. Yanukovich (02 21) ● Y. Tymoshenko is liberated, Yanukovich leaves the country (02 22). <u>February 27-28</u> ● Russian forces occupy the Crimean parliament building. ● Russian troops are deployed to airports and other strategic sites.

Qualitative discourse analysis was performed for these media articles, identifying actors, qualities they were attributed to, predicates of events as well as the qualities they were attributed to. This analysis was performed for BBC, RussiaToday, DayKiev and delfi.lt articles separately. We used the dataset of these results for our narrative analysis via word co-occurrence networks.

The second dataset consists of raw BBC articles, covering the same 3 stages of Ukrainian crisis as described above. The articles were grouped according to the conflict stage they described, and no pre-processing of the texts was performed. We used this dataset for our sentiment-based storyline analysis.

3.0 METHODS

We used 2 methods for narrative analysis in terms of Ukrainian crisis. The first one is word co-occurrence network analysis and the second one – sentiment-based storyline analysis. We used word co-occurrence networks with dataset containing results of qualitative discourse analysis of media articles (BBC, RussiaToday, DayKiev and delfi.lt), while for sentiment-based storyline analysis we used raw BBC articles. In the following sections we describe these 2 methods in more detail.

3.1 Word Co-occurrence Network

Network analysis describes a group of methods that characterize relationships between units to be analysed, i.e. nodes representing units to be analyzed and edges representing connections between nodes. Network analysis is often used in the context of social networks, where nodes usually represent people, while edges refer to friendships, affiliations or other types of relationships [7], [8, p. 10]. Though network analysis is most often used in terms of relationships between people, it can also be applied to represent relationships between words, e. g. [9], [10], [11]. Word co-occurrence network analysis is text length insensitive method thus it is suitable even when dataset is made of short texts or just texts of uneven length [12]. For word co-occurrence network analysis we used *textnets*¹ package for R².

Word co-occurrence network analysis was performed for our first dataset, made of results of qualitative discourse analysis of media articles (BBC, RussiaToday, DayKiev and delfi.lt). The textual data was tokenized and then lemmatized i.e., each word was replaced with its most basic syntactic form [11]. Then part-of-speech tagger (integrated into *textnets*, based of language models of Universal Dependencies project³) was applied in order to identify nouns and noun phrases which most likely describe content of the text [12]. In the next step term frequency-inverse document frequency (TF-IDF) was characteristic was calculated for each noun and noun phrase. This characteristic was calculated for each stage of the conflict (see Table 2-1) separately. Then a bipartite word co-occurrence networks were generated for each stage of the conflict. These bipartite networks linked media sources (BBC, RussiaToday, DayKiev and delfi.lt) based on terms (nouns and noun phrases) in their messages. Finally the community detection algorithm was applied in order to detect similarities in the selected media sources in terms of reporting on the Ukrainian conflict [11], [12]. The results of word co-occurrence networks regarding Ukrainian conflict is presented in Section 4.1.

3.2 Sentiment-Based Storyline Analysis

Sentiment analysis is important task of Natural Language Technologies and Artificial Intelligence, with applications such as automated analysis of reviews and social media, monitoring of political issues, etc. [13]. Sentiment analysis assumes that emotions are important in communication among humans. Some recent work in terms of analysis of literary texts has suggested that shifts in sentiment can be used for analysis of plot development [14, pp. 73-110]. As the raw sentiment time series is very noisy, a smoothing filter is usually applied for the raw data [15]. The result is a fairly smooth curve which represents generalized trend of the sentiment development of the text. This method is based on controlled vocabulary and data compression. For sentiment-based storyline analysis we used package *syuzhet*⁴ for R.

¹ <https://github.com/cbail/textnets>

² R is a software environment (free) for statistical computing and graphics, see: <https://www.r-project.org/>

³ <https://universaldependencies.org/>

⁴ <https://cran.r-project.org/web/packages/syuzhet/>

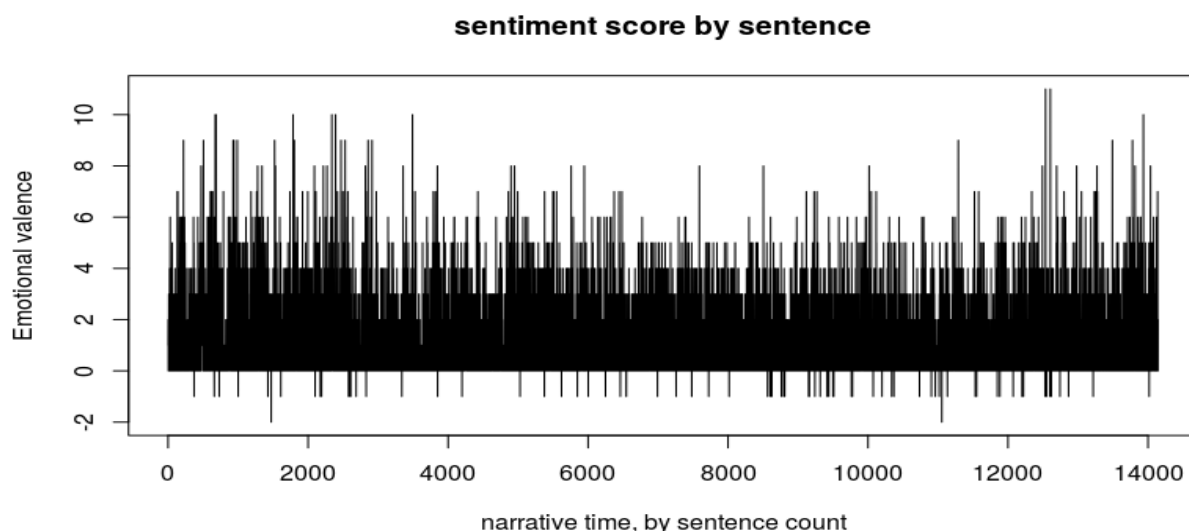


Figure 3-1: Example of raw sentiment plot.

Sentiment-based storyline analysis was performed for our second dataset, made of raw BBC articles, covering 3 stages (see Table 2-1) of the Ukrainian conflict. This analysis was applied for the articles grouped by the stage of the conflict they covered. Sentiment analysis was used to detect latent structure of narrative concerning reporting of Ukrainian conflict. Narrative process (development) in this analysis was marked by emotional shifts. Firstly, texts were tokenized into sentences. Then for each sentence sentiment value was calculated. For these calculations NRC Emotion Lexicon⁵ [16], integrated into R package *syuzhet*, was used. The words in each sentence are assigned a sentiment value according to their emotional intensity and polarity. When a sentiment value is absent, the word is assigned a value of zero (neutral). Then emotional score of each sentence is estimated by aggregating the word-sentiment values.

As we calculate sentiment values for every sentiment, the cumulative results of all sentences belonging to certain Ukrainian conflict stage is “noisy” and non-interpretable for human eyes (see Figure 3-1). For generalized sentiment trajectory Discrete Cosine Transformation (DCT) was applied [17]. After that, the sentiment trajectory was sub-sampled to have 100 elements so that sentiment arcs of BBC news coverage for each stage of Ukrainian conflict could be represented from 0% (beginning of the textual data belonging to certain stage of Ukrainian conflict) to 100% (end of the movie of the textual data belonging to certain stage of Ukrainian conflict). This transformation was used to extract sentiment and sentiment-derived plot arcs from text. Finally, as NRC Emotion Lexicon allows dataset evaluation not only according to polarity scale (positive-neutral-negative), but also according to prevalence of 8 emotions (anger, fear, anticipation, trust, surprise, sadness, joy, disgust), we generated emotional “profiles” for BBC coverage of 3 stages of Ukrainian conflict. The results of sentiment-based storyline analysis of Ukrainian conflict are presented in Section 4.2.

4.0 RESULTS

In this section we present results of 2 approaches to narrative analysis. In Section 4.1 we present results of word co-occurrence networks for news coverage of 3 stages of Ukrainian conflict by 4 media sources – BBC, RussiaToday, DayKiev and delfi.lt. In this approach narrative of the conflict was analyzed based on relations (co-occurrences) of nouns and noun phrases in the messages of these 4 media sources for each stage of the conflict separately. In Section 4.2 we present results of sentiment-based storyline analysis of news coverage of the same 3 stages of Ukrainian conflict by BBC. In this approach narrative of the conflict was

⁵ <http://saifmohammad.com/WebPages/NRC-Emotion-Lexicon.htm>

analyzed based on emotional shifts in the textual data and generation of plot trajectories as well as emotional “profiles”.

4.1 Results of Word Co-occurrence Network Analysis

Figure 4-1 presents word co-occurrence network made of nouns and noun phrases for the 1st stage of Ukrainian conflict (2013 11 21-2014 01 15). As it was presented in Table 2-1, some more important events during the 1st stage of Ukrainian conflict were V. Yanukovych’s postponement to sign an association agreement with the EU, Y. Tymoshenko’s prevention from leaving Ukraine, the beginnings of large-scale protests and demonstrations.

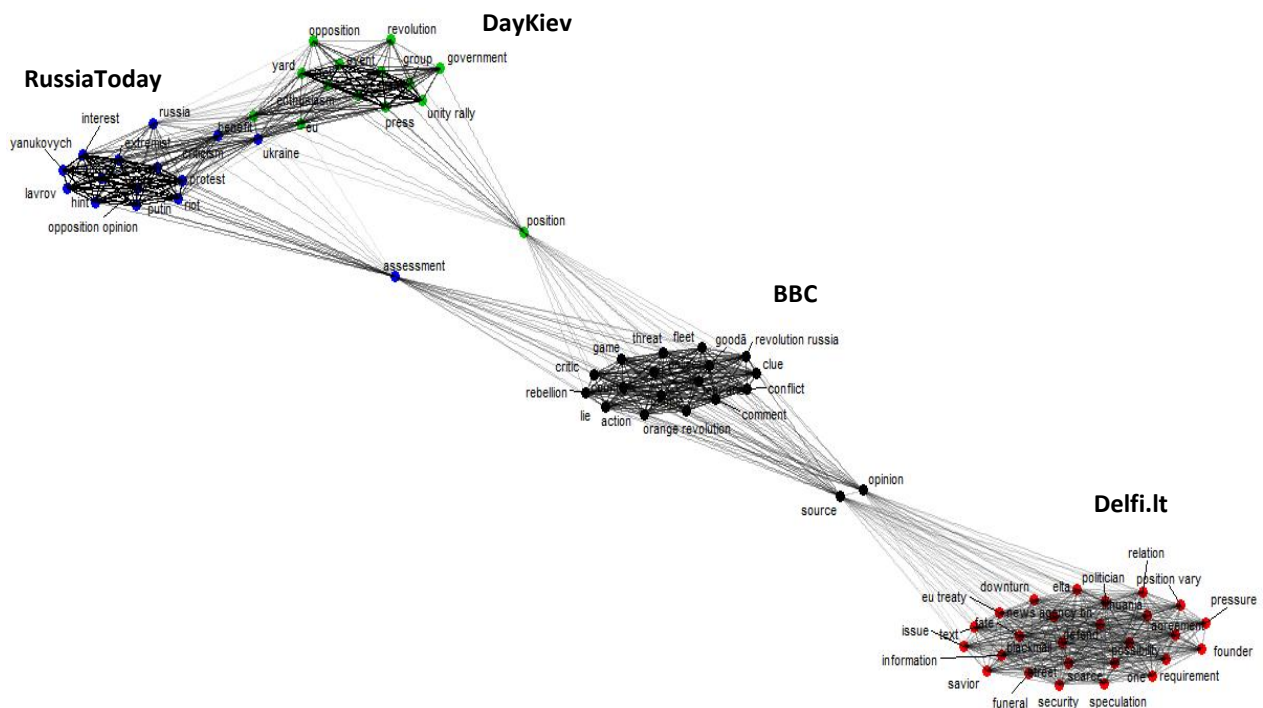


Figure 4-1: Word co-occurrence network: 1st stage of Ukrainian conflict (alpha=0.5).

As Figure 4-1 shows, in RussiaToday messages Nov 21, 2013 – Jan 15, 2014 names of politicians as important actors in the conflict are distinguished (‘yanukovych’, ‘lavrov’, ‘putin’). Naturally, names of the countries involved in the conflict are frequent as well (‘russia’, ‘ukraine’). Events during beginning stage of ‘Euromaidan’ by RussiaToday was described as ‘protest’, ‘riot’, ‘criticism’, ‘extremist’, ‘interest’, ‘benefit’, ‘opposition opinion’. DayKiev spread messages about ‘eu’, ‘enthusiasm’, ‘press’, ‘unity rally’, ‘government’, ‘group’, ‘revolution’, ‘event’, ‘opposition’, ‘yard’. Although these 2 media sources were supposed to be on the opposite sides of the reporting, word clusters that represent the most frequent topics in reporting are close to each other indicating similarities in news coverage. Words ‘ukraine’ (blue color; belongs to RussiaToday cluster) and ‘benefit’ (green color; belongs to DayKiev cluster) could be treated as so called “cultural bridges” [11] that link RussiaToday and DayKiev messages about 1st stage of Ukrainian conflict. “Cultural bridges” are terms (nouns and noun phrases in our case) that lie between clusters and act as bridges by linking disjoint discourses [12].

BBC reporting was distinguished by describing events as ‘rebellion’, ‘conflict’, ‘game’, ‘threat’, ‘lie’, ‘action’, ‘orange revolution’, ‘revolution russia’. Words ‘comment’, ‘clue’, ‘fleet’, ‘critic’ were important in BBC messages as well. BBC and DayKiev messages on the conflict were linked by word ‘position’, while

BBC and RussiaToday messages were linked by ‘assessment’ acting as “cultural bridge”. Finally, delfi.lt messages contain words ‘eu treaty’, ‘downturn’, ‘elta’, ‘politician’, ‘lithuania’, ‘position vary’, ‘relation’, ‘pressure’, ‘agreement’, ‘founder’, ‘requirement’, ‘possibility’, ‘defend’, ‘security’, ‘speculation’, ‘scarce’, ‘street’, ‘funeral’, ‘blackmail’, ‘savior’, ‘information’, ‘issue’, ‘fate’, etc. Words ‘source’ and ‘opinion’ as “cultural bridges” link delfi.lt and BBC messages together.

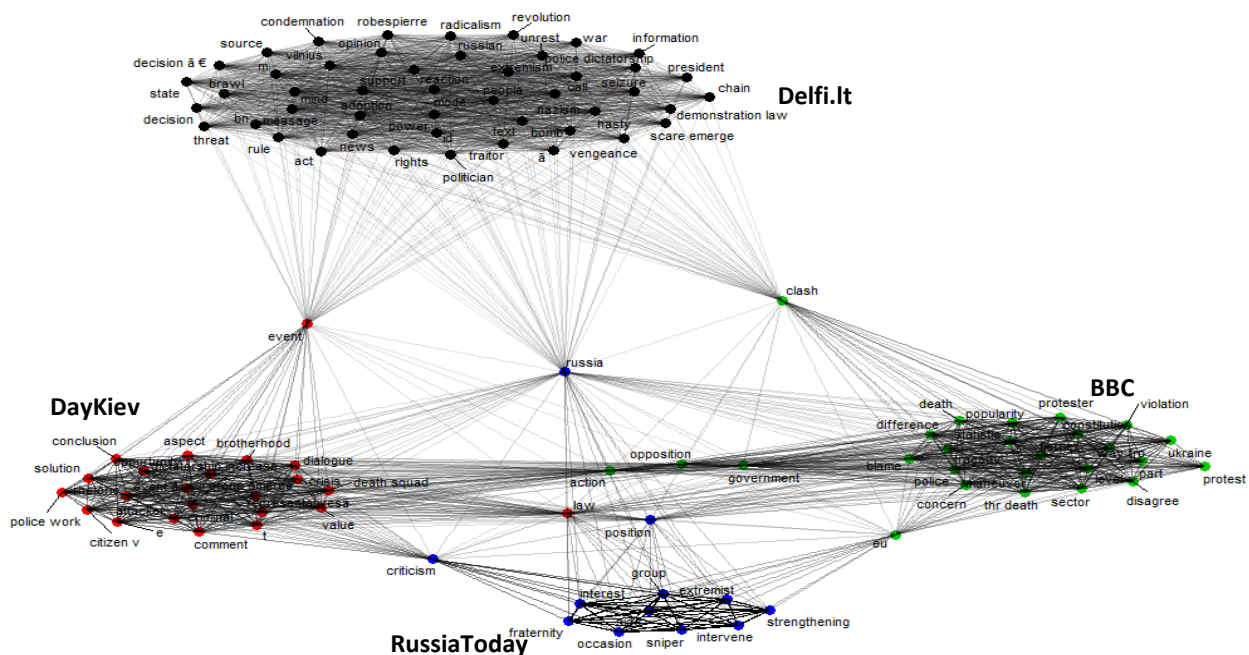


Figure 4-2: Word co-occurrence network: 2nd stage of Ukrainian conflict (alpha=0.6).

2nd stage of Ukrainian conflict (Jan 16, 2014 - Feb 17, 2014) can be described as *escalation* because during this period first deaths occurred, raiding government buildings in western Ukraine took place, etc. (see Table 2-1). As Figure 4-2 shows, RussiaToday during this period spread messages about ‘fraternity’, ‘occasion’, ‘sniper’, ‘intervene’⁶, ‘strengthening’, ‘extremist’, ‘group’, ‘occasion’, ‘interest’, etc. Meanwhile, DayKiev focused on ‘police work’, ‘comment’, ‘solution’, ‘conclusion’, ‘rhetoric’, ‘opportunity’, ‘attacker’, ‘criminal’, ‘value’, ‘crisis’, ‘death squad’, ‘dialogue’, ‘brotherhood’, etc. Word ‘criticism’ links DayKiev and RussiaToday as “cultural bridge”, i. e., discursive theme that is seldom discussed but have a potential to attract feedback and opinion from public [11]. BBC reporting on *escalation* of Ukrainian conflict focused on ‘difference’, ‘blame’, ‘death’, ‘popularity’, ‘tragedy’, ‘maneuver’, ‘statistic’, ‘constitution’, ‘violation’, ‘police’, ‘concern’, ‘disagree’, ‘protest’, ‘protester’, ‘ukraine’, ‘sector’, etc. BBC and RussiaToday narratives are connected by ‘eu’ (green; belongs to BBC). 3 words – ‘event’ (red – DayKiev) , ‘russia’ (blue – RussiaToday) and ‘clash’ (green – BBC) link narratives of all 4 news sources (see Figure 4-2). Narratives of RussiaToday, DayKiev and BBC were linked by ‘law’ (red – DayKiev) and ‘position’ (blue – RussiaToday). DayKiev and BBC narratives had 3 “cultural bridges”: ‘action’, ‘opposition’, ‘government’ (all green, belong to BBC word cluster). Finally, delfi.lt narrative told about ‘state’, ‘decision’, ‘rule’, ‘condemnation’, ‘opinion’, ‘robespierre’, ‘radicalism’, ‘revolution’, ‘unrest’, ‘war’, ‘dictatorship’, ‘rights’, ‘president’, ‘demonstration law’, ‘vengeance’, ‘politician’, ‘traitor’, ‘reaction’, ‘support’, ‘power’, ‘people’, ‘seizure’, etc. Based on Figure 4-2, it might be assumed that delfi.lt narrative was the most diverse while RussiaToday – the least diverse in terms of discursive themes (size of cluster).

⁶ Part-of-speech tagger was not always successful in correctly recognizing nouns.

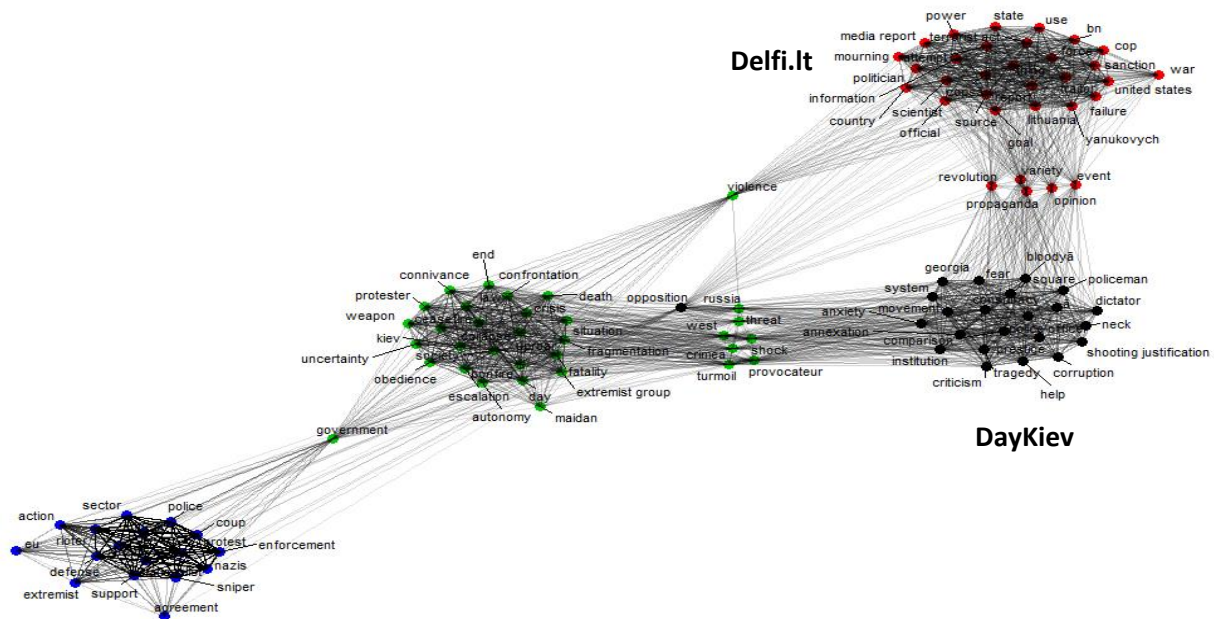


Figure 4-3: Word co-occurrence network: 3rd stage of Ukrainian conflict (alpha=0.4).

3rd stage of Ukrainian conflict (Feb 18 – 28, 2014) is exceptional due to casualties on both sides – protesters as well police, – and agreement between opposition and Yanukovych, liberation of Tymoshenko, Yanukovych ‘s escape from Ukraine and its final event was occupation of Crimea (see Table 2-1).

As Figure 4-3 shows, reporting of RussiaToday featured ‘eu’, ‘extremist’, ‘police’, ‘coup’, ‘protest’, ‘enforcement’, ‘rioter’, ‘defence’, ‘support’, ‘protest’, ‘nazis’, ‘agreement’, ‘sniper’, ‘action’, ‘sector’ discursive themes. BBC coverage of the 3rd stage of Ukrainian conflict spread messages about ‘uncertainty’, ‘obedience’, ‘kiev’, ‘weapon’, ‘protester’, ‘end’, ‘confrontation’, ‘death’, ‘situation’, ‘fragmentation’, ‘extremist group’, ‘fatality’, ‘connivance’, ‘law’, ‘ceasefire’, ‘collapse’, ‘uproar’, ‘autonomy’, ‘maidan’, ‘crisis’, ‘escalation’, etc. Word ‘government’ (green, belongs to BBC cluster) lies between RussiaToday and BBS cluster, acting as a bridge between narratives of these media channels. Furthermore, delfi.lt reported about ‘mourning’, ‘information’, ‘media report’, ‘terrorist act’, ‘country’, ‘scientist’, ‘power’, ‘state’, ‘use’, ‘war’, ‘failure’, ‘cop’, ‘sanction’, ‘united states’, ‘lithuania’, ‘goal’, ‘force’, ‘yanukovych’, etc. Finally, DayKiev featured ‘georgia’, ‘fear’, ‘bloody’⁷, ‘square’, ‘policeman’, ‘conspiracy’, ‘dictator’, ‘neck’, ‘police officer’, ‘shooting justification’, ‘prestige’, ‘corruption’, ‘help’, ‘tragedy’, ‘criticism’, ‘comparison’, ‘institution’, ‘anxiety’, ‘annexation’, ‘movement’, ‘system’. Words ‘russia’, ‘threat’, ‘west’, ‘shock’, ‘crimea’, ‘turmoil’ and ‘provocateur’ (all green, belong to BBC) link narratives of BBC and DayKiev together. BBC and delfi.lt narratives were bridged by discursive theme of ‘violence’ (green dot between BBC and DayKiev clusters). Finally, words ‘revolution’, ‘propaganda’, ‘variety’, ‘opinion’ and ‘event’ (red dots between delfi.lt and DayKiev clusters) act as bridges for delfi.lt and DayKiev narratives.

Beside analysis of the whole networks representing narratives of media coverage of 3 stages of Ukrainian conflict, we also identified communities, i.e., grouped media channels according similarity of their narratives. For this task Louvain community detection algorithm [18] was applied. Thus, for the 1st stage of Ukrainian conflict 2 groups were discerned: 1) DayKiev and delfi.lt, 2) RussiaToday and BBC. Top nouns and noun phrases that distinguish narratives of these two groups are presented in Table 4-1.

⁷ Again part-of-speech tagger did recognize some nouns correctly.

Table 4-1: Similarity of narratives: 1st stage of Ukrainian conflict.

DayKiev & delfi.lt	RussiaToday & BBC
action	benefit
clue	criticism
comment	enthusiasm
conflict	event
country	government
critic	group
fear arise	hesitant
fleet	lack
game	press
lie	unity rally
orange revolution	yard

Interestingly enough, for the 2nd stage of the Ukrainian conflict all 4 media channels fell into the same group. This leads to the assumption that for this stage of the conflict narratives of BBC, DayKiev, delfi.lt and RussiaToday were not that different, or discussions were wide enough to include different options. Another possibility, which slightly complicates things, is application of terms, as extremist to opposing sides by different media channels. Top nouns and noun phrases of collective narrative of 4 media channels are presented in Table 4-2.

Table 4-2: Collective narrative of the 2nd stage of Ukrainian conflict.

DayKiev & delfi.lt & RussiaToday & BBC
protester
extremist
fraternity
group
interest
intervene
occasion
side
sniper
strengthening

Table 4-3: Similarity of narratives: 3rd stage of Ukrainian conflict.

DayKiev & delfi.lt	RussiaToday & BBC
dictator	action
georgia	appear
attempt	coup
cop	defence
country	enforcement
failure	group
force	nationalist
goal	nazis
yanukovych	police

information	protester
media report	rioter
mourning	sniper

For the 3rd stage of the Ukrainian conflict we again discerned the same 2 communities as for the 1st stage of the conflict: 1) DayKiev and delfi.lt, 2) RussiaToday and BBC. Top nouns and noun phrases that distinguish narratives of these two groups are presented in Table 4-3. Comparing top terms (nouns and noun phrases) of the 1st and the 3rd stages of Ukrainian conflict, it was observed that word ‘country’ remains for the both cases for the group DayKiev + delfi.lt, while the same goes for the group BBC + RussiaToday in terms of word ‘group’. Top nouns of the 2nd stage of Ukrainian conflict did not have matches for the group DayKiev + delfi.lt neither in the 1st, nor in the 2nd stage of the conflict. However, 3 of the top nouns of the 2nd stage of Ukrainian conflict matched the ones for the group BBC + RussiaToday: ‘protester’ and ‘sniper’ – 3rd stage of the Ukrainian conflict, ‘group’ – 1st as well as 3rd stage of the conflict.

4.2 Results of Sentiment-Based Storyline Analysis

In this section we present results of sentiment-based storyline analysis in terms of BBC narrative regarding 3 stages of Ukrainian conflict (see Table 2-1). Firstly, we discuss emotional shifts representing narrative development in BBC reporting (Section 4.2.1). After that we discuss emotional “profiles” of BBC narratives in terms of 3 stages of Ukrainian conflict (Section 4.2.2).

4.2.1 Sentiment-Based Plot “Arcs”: Emotional shifts as Narrative Development in BBC Coverage of Ukrainian Conflict

In this section we present results of sentiment-based storyline analysis of BBC coverage of Ukrainian conflict. Generalized sentiment values allow generation of plot trajectory curves (or arcs) in terms of polarity scale (positive-neutral-negative). These plot arcs represent emotional shifts in the reporting on 3 stages of Ukrainian conflict (see Table 2-1) separately. Emotional shifts in our study was used to evaluate narrative development in BBC coverage of the conflict.

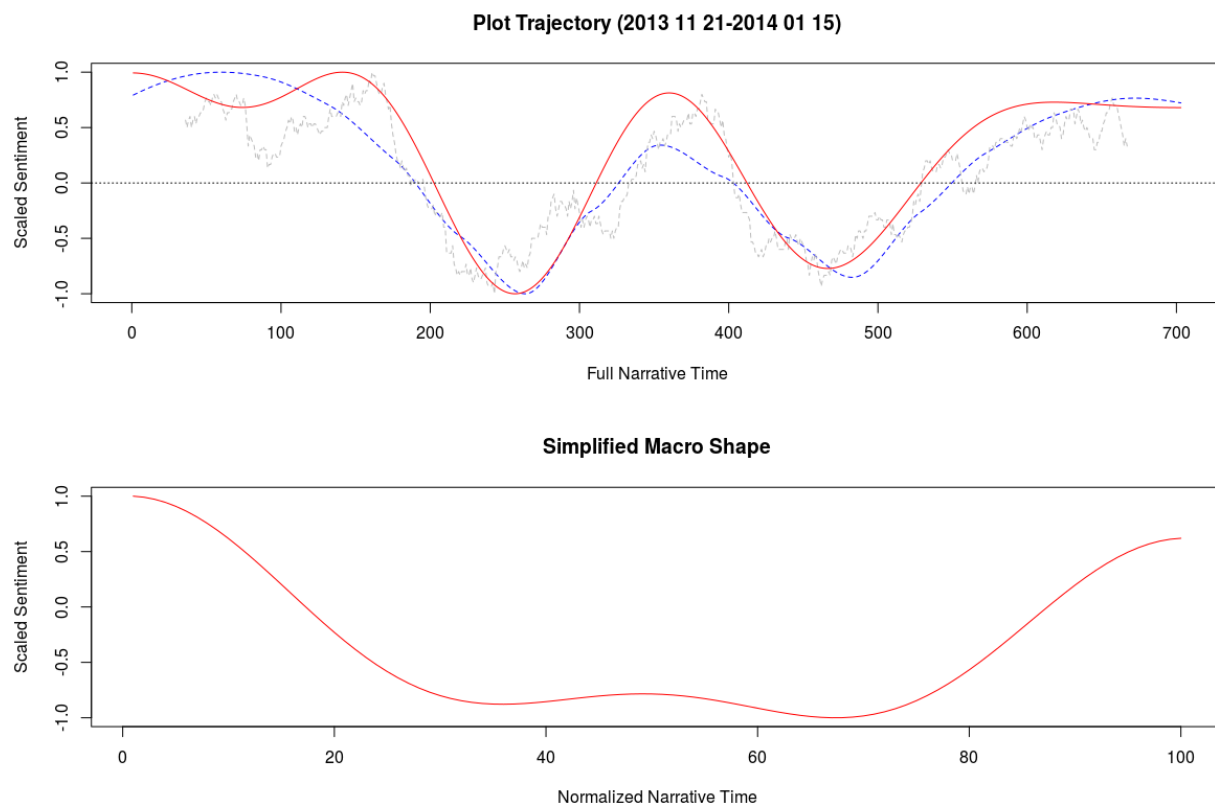


Figure 4-4: Sentiment-based Narrative Trajectory: 1st stage of Ukrainian conflict.

Figure 4-4 shows sentiment-based narrative trajectory for the 1st stage of Ukrainian conflict (Nov 21, 2013 – Jan 15, 2014), Figure 4-5 represents sentiment-based narrative trajectory for the 2nd stage of Ukrainian conflict (Jan 01 – Feb 17, 2014) and Figure 4-6 – narrative trajectory for the 3rd stage of Ukrainian conflict (Feb 18-28, 2014). Upper diagrams represent full narrative time (measured in number of sentences that comprise our dataset for this stage of Ukrainian conflict), while the lower ones depict simplified macro shapes of the narrative, expressed as emotional ups and downs. Both graphs in the Figures 4-4, 4-5 and 4-6 evaluate sentiments across narrative of separate stages of Ukrainian conflict as reported by BBC. Sentiments range between -1 and 1, thus the closer to 1 sentiment value is, the more positive it is.

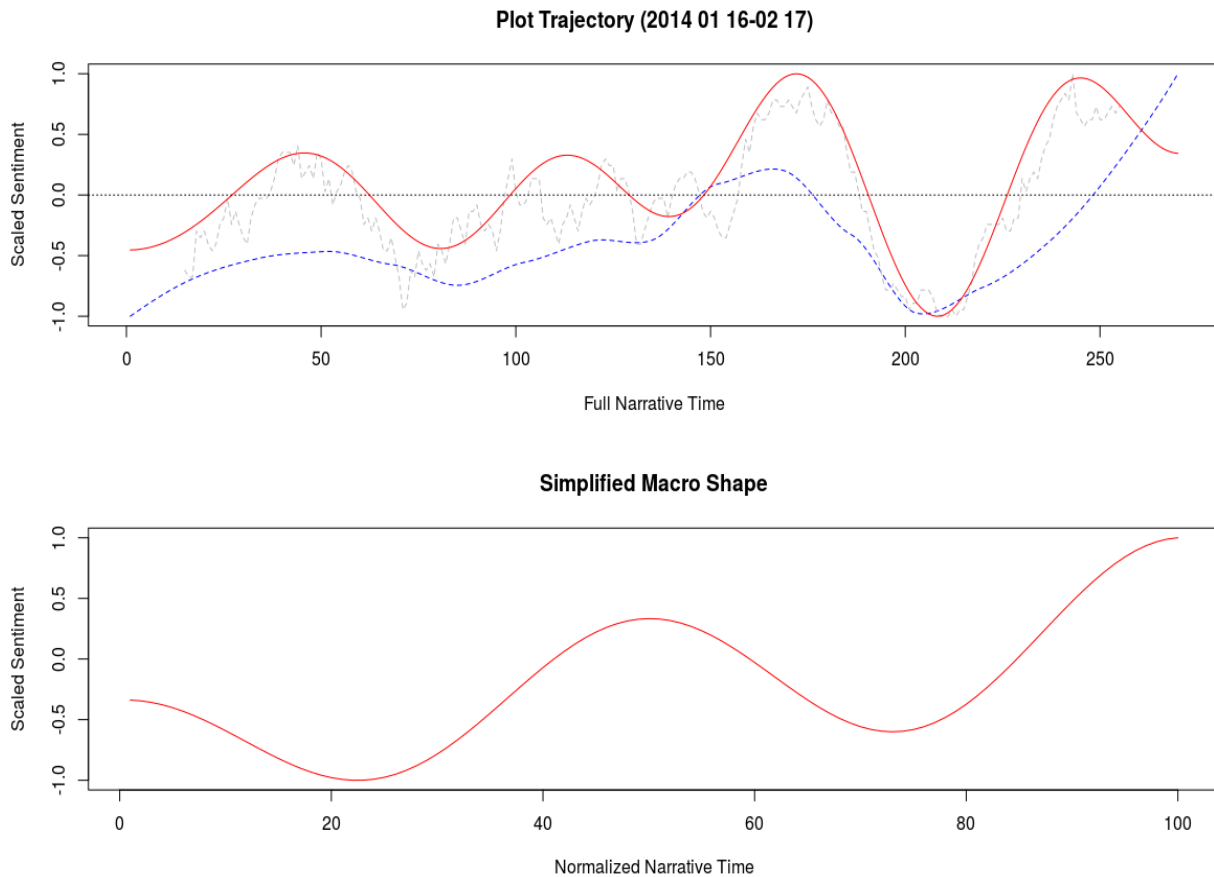


Figure 4-5: Sentiment-based Narrative Trajectory: 2nd stage of Ukrainian conflict.

The narrative time is measured in number of sentences. Plot trajectory (upper diagram) is composed of 3 plots, each of them representing a smoothing method applied on the data. The red line indicates the Discrete Cosine Transformation (DCT) of the narrative, which was used to extract sentiment and sentiment-derived plot arcs from text (for more detail, see Section 3.2). This transformation made it possible to evaluate the overall emotional trends through narrative arcs of BBC reporting about different stages of Ukrainian conflict. The grey dashed line shows the results of applying a moving-average filter on the narrative in order to smooth the short-term fluctuations of sentiments detected in the sentences of textual data. It features extended emotional sense of the stories being told by BBC coverage of separate stages of Ukrainian conflict. Finally, the blue dashed line shows another smooth pattern of the sentiments created using Loess regression [19].

The simplified macro shape plot (lower diagram) is a flattened representation of data previously depicted by 3 curves mentioned above. This macro shape plot was generated by adopting a combination of discrete cosine transformation (DCT) and a low pass filter [20], where the emotional trajectory bears the analogy to functions fluctuating at different frequencies. Simplified macro computed with DCT retained fewer low frequency components, that is, minimized the noise. Also, it used the reverse transform process to normalize the x-axis to 100 units, i.e., the sentiment trajectory was sub-sampled to have 100 elements so that sentiment arcs of BBC news coverage for each stage of Ukrainian conflict could be represented from 0% (beginning of the textual data belonging to certain stage of Ukrainian conflict) to 100% (end of the movie of the textual data belonging to certain stage of Ukrainian conflict).

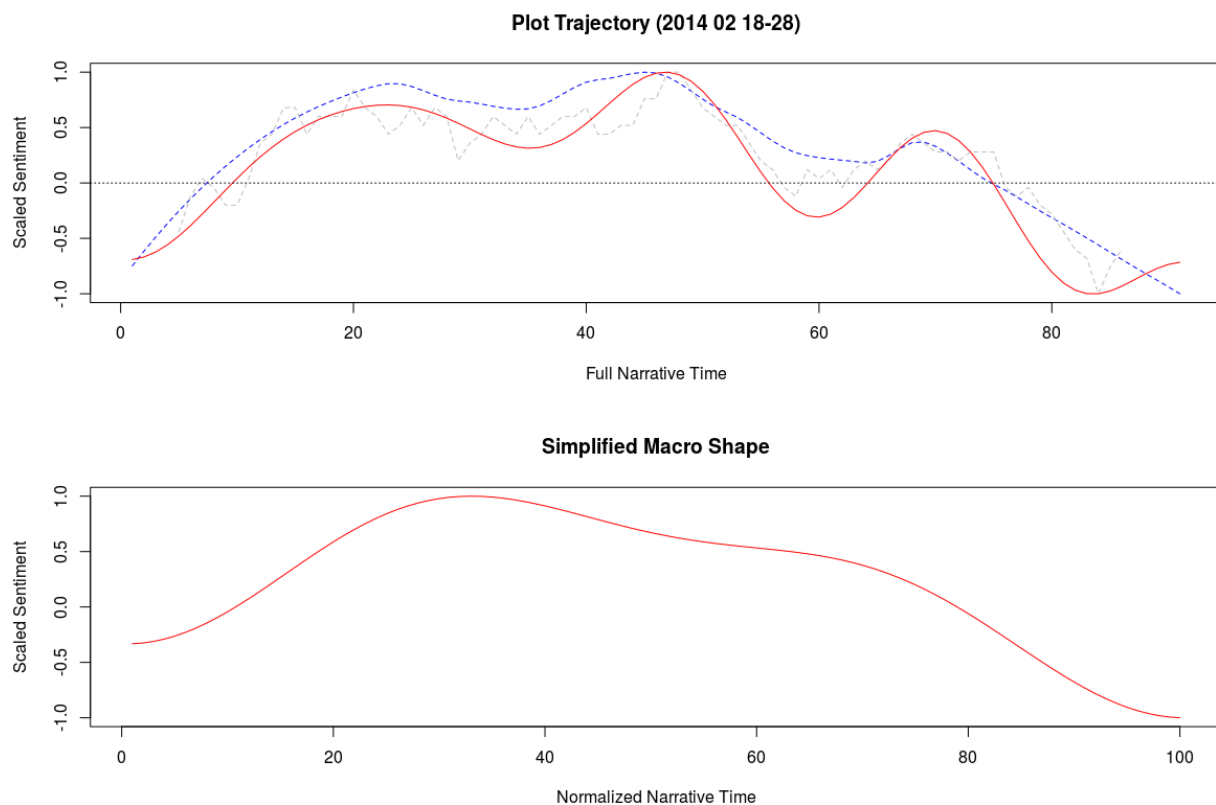


Figure 4-6: Sentiment-based Narrative Trajectory: 3rd stage of Ukrainian conflict.

In Figure 4-4, representing the 1st stage of Ukrainian conflict, macro shape indicates that narrative trajectory starts on the high positive note, then rapidly descends down into negativity and at the end we see the rise of positive sentiment again. In Figure 4-5, which shows the 2nd stage of Ukrainian conflict, macro shape reveals that there several “ups and downs”, the beginning is marked by moderately negative sentiment value, however, the narrative trajectory ends very positively. Finally, in Figure 4-6, representing the 3rd stage of Ukrainian conflict, macro shape shows that narrative trajectory again starts with moderately negative sentiment, then we can see a “peak” of positive sentiment, which gradually deteriorates, and thus narrative trajectory ends with extremely negative sentiment.

Explanations of the narrative trajectories above can be the following ones:

1. 1st stage of Ukrainian conflict: starts with Ukraine being “courted by East and West” (positive sentiment), then Yanukovich postponed the signing of an association agreement with the EU and large-scale protests began (sentiment rapidly descends down into negativity), and this stage of the conflict ended with Maidan “rally” (the rise of positive sentiment).
2. 2nd stage of Ukrainian conflict: starts with adoption of anti-demonstration law (moderately negative sentiment), then first deaths occurred (extremely negative sentiment), after that – some attempts to solve crisis peacefully (positive “peak”), clashes with police (moderately negative descend of sentiment) and the 2nd stage of Ukrainian conflict ended with nullification of the laws directed against the protests (narrative trajectory ends very positively).
3. 3rd stage of Ukrainian conflict: beginnings of the clashes (moderately negative sentiment), then Yanukovich signed peace agreement with opposition, liberation of Tymoshenko (“peak” of positive sentiment) and the finale of this stage of the conflict – occupation of Crimea (sentiment deteriorates, and narrative trajectory ends with extremely negatively).

4.2.2 Emotional “Profiles” of BBC Coverage of Ukrainian Conflict

Package *syuzhet* and NRC Emotion Lexicon allowed to evaluate BBC coverage of 3 stages of Ukrainian conflict in terms of prevalence of 8 emotions – anger, fear, anticipation, trust, surprise, sadness, joy and disgust). Thus we generated emotional “profiles” for each Ukrainian conflict stage we analysed. These emotional “profiles” are presented in Figures 4-7, 4-8 and 4-9.

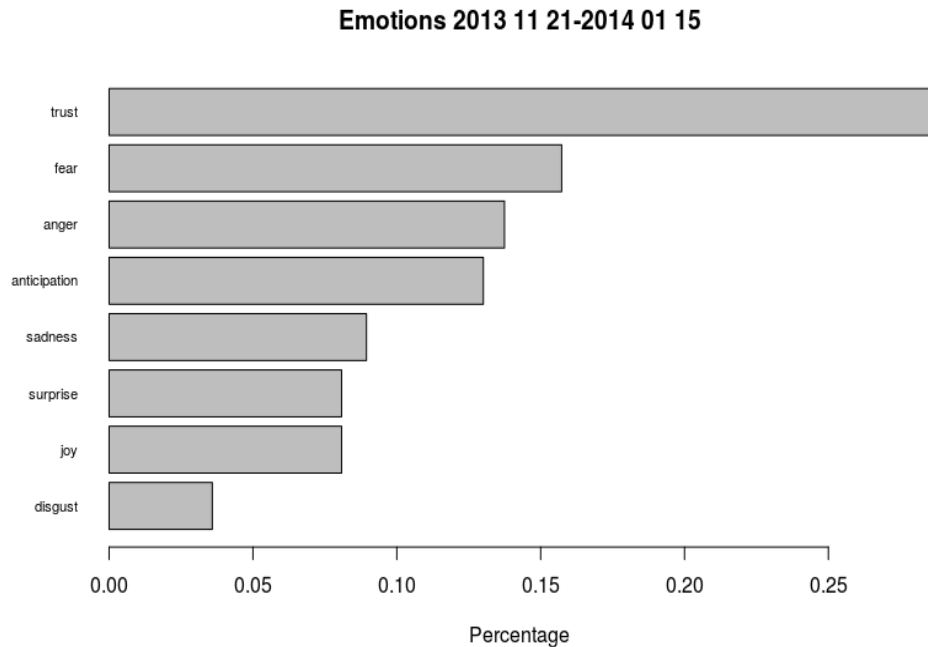


Figure 4-7: Emotional “profile”: 1st stage of Ukrainian conflict.

Figure 4-7 shows emotional “profile” of BBC reporting on the 1st stage of Ukrainian conflict. The most prevalent emotion is ‘trust’, other emotions were less expressed. However, Figure 4-8 reveals that although ‘trust’ is still the most expressed emotion in BBC coverage of 2nd stage of Ukrainian conflict, ‘fear’ is much closer to it than in 1st stage of the conflict. Other emotion, such as ‘anger’, ‘sadness’ and ‘disgust’ was also expressed more than in the 1st stage, while ‘anticipation’ and ‘joy’ – less than in the 1st stage of conflict.

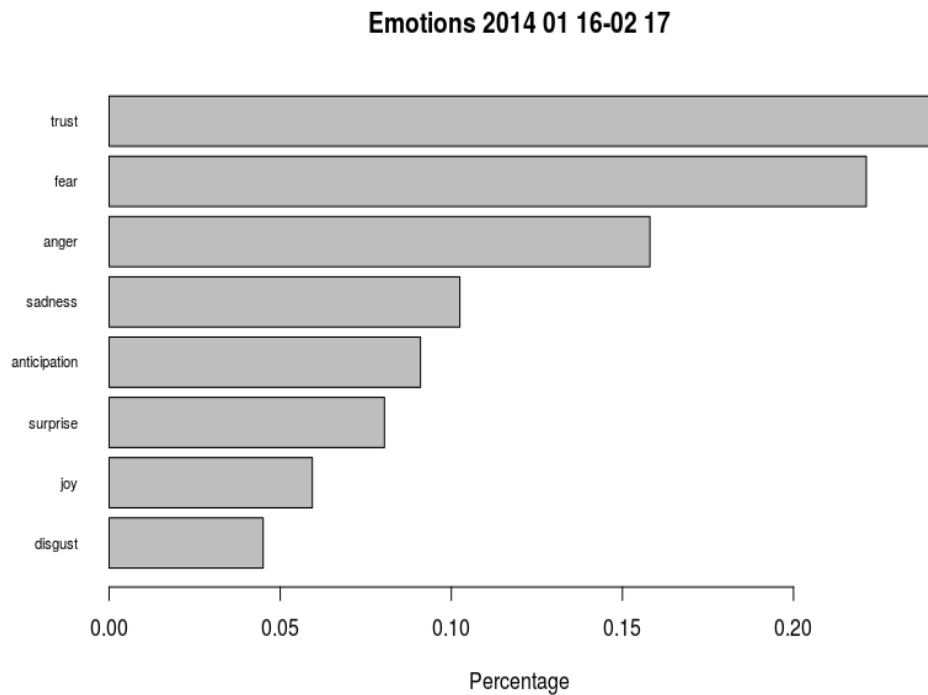


Figure 4-8: Emotional “profile”: 2nd stage of Ukrainian conflict.

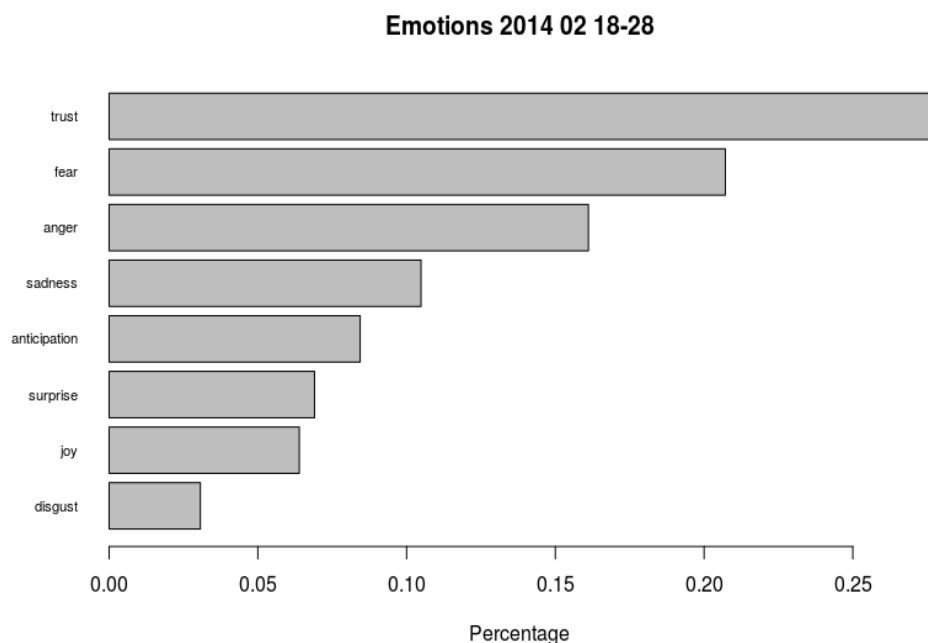


Figure 4-9: Emotional “profile”: 3rd stage of Ukrainian conflict.

During the 3rd stage of Ukrainian conflict BBC reporting still expressed ‘trust’ the most and more than during 2nd stage (it might be assumed that ‘trust’ level more or less returned to the level of reporting about the 1st stage of Ukrainian conflict). The same tendency was observed for ‘disgust’ and ‘joy’ – the former decreased, the latter – increased. ‘Fear’ was expressed on moderately lesser degree, although it did not shrink to the same level as during coverage of the 1st stage of the conflict. ‘Anger’ and ‘sadness’ levels remained

almost the same as during reporting on 2nd stage of Ukrainian conflict and ‘anticipation’ might be considered as expressed the least in comparison to reporting on all 3 stages of Ukrainian conflict.

5.0 CONCLUSIONS

The goal of the research was two-fold: to evaluate possibility to analyse media – population interaction during conflict, and to evaluate possibility to automate the process.

Research results show that it is comparatively easy to track opinion and rhetoric change using Natural Language Processing Tools. Events in the conflict are well reflected in media, and media partially reflects population mood.

While automation was not investigated in details, in most of the cases separate tools that could work with minimal human interference. Of course, several aspects still should be chosen by analysts, namely the conflict itself, media sources (in full-fledged tool just ticking corresponding checkboxes), and, events. We did not investigate automatic or even semi-automatic choice of events. Of course, to make visualizations nicer, human help is valuable as well.

Results:

1. Description of two media rhetoric analysis processes: cooccurrence networks and syuzhet based.
2. Experimental evaluation of cooccurrence networks approach on Ukrainian conflict reflection in 4 media channels: BBC, RussiaToday, DayKiev and Delfi.lt.
3. Experimental evaluation of syuzhet-based approach for Ukrainian conflict dynamics in BBC.

Conclusions:

1. Cooccurrence networks approach reflects change of rhetoric in different media channels. Spatial visualisation of networks shows rhetoric differences and similarities between channels.
2. Syuzhet-based analysis reflects change of opinion over time in media channel. However, sentence-wise analysis does not reflect sufficiently well change over time due to sampling, hence representation can be badly distributed.
3. Both approaches are suitable for automation.

Based on the results, the following future developments are proposed:

1. Automation of both processes (except selection of events and topic).
2. Research of automatic events detection (which is quite popular research area with some interesting results).
3. Detection of narrative itself (it is quite well explored research area as well).
4. Modification of syuzhet-based model, i.e. time-normalization, to better reflect change over time, instead of simple sampling, some kind of weighted aggregation.
5. More advanced sentiment analysis methods could improve precision of syuzhet as well.

References

- [1] J. Elliott, "Using narrative in social research: Qualitative and quantitative approaches," SAGE Publications, 2005.
- [2] W.-Y. S. Chou, Y. Hunt, A. Folkers and E. Augustson, "Cancer survivorship in the age of YouTube and social media: a narrative analysis," *Jrnl. of medical Internet research*, vol. 13, no. 1, p. e7, 2011.
- [3] S. R. Corman, S. W. Ruston and M. Fisk, "A pragmatic framework for studying extremists' use of cultural narrative," In *Proc. of the 2nd Int. Conf. on Cross-Cultural Decision Making: Focus*, 2012.
- [4] M. Holmstrom, "The narrative and social media," vol. 1, no. 1, pp. 118-132, 2016.
- [5] S. W. Ruston, "More than just a story: Narrative insights into comprehension ideology and decision-making," *Modeling Sociocultural Influences on Decision Making*, pp. 57-72, 2016.
- [6] A. Bell, *The Language of News Media*, Blackwell, Oxford, UK, 1991.
- [7] J. Scott, "Social network analysis," *Sociology*, vol. 22, no. 1, pp. 109-127, 1988.
- [8] L. Freeman, "The development of social network analysis (A Study in the Sociology of Science)," vol. 1, Empirical press, 2004.
- [9] D. Paranyushkin, "Identifying the pathways for meaning circulation using text network analysis," *Nodus Labs*, vol. 26, 2011.
- [10] P. Ruiz, C. Plancq and T. Poibeau, "More than word cooccurrence: Exploring support and opposition in international climate negotiations with semantic parsing," In *LREC: The 10th Language Resources and Evaluation Conference, ELRA, May 2016, Portorož, Slovenia*, 2016.
- [11] C. A. Bail, "Combining natural language processing and network analysis to examine how advocacy organizations stimulate conversation on social media," *Proc. of the National Academy of Sciences*, vol. 113, pp. 11823-11828, 2016.
- [12] A. Rule, J.-P. Cointet and P. S. Bearman, "Lexical shifts, substantive changes, and continuity in State of the Union discourse 1790-2014," *Proc., of the Nat. Acad. of Sc.*, vol. 112, pp. 10837-10844, 2015.
- [13] E. Cambria, "Affective computing and sentiment analysis," *IEEE Intelligent Systems*, vol. 31, no. 2, pp. 102-107, 2016.
- [14] J. Archer and M. L. Jockers, "The bestseller code: Anatomy of the blockbuster novel," St. Martin's Press, 2016.
- [15] J. Gao, M. L. Jockers, J. Laudun and T. Tangherlini, "A multiscale theory for the dynamical evolution of sentiment in novels," In *Proc. of the Int. Conf. on Behavioral, Economic and Socio-cultural Computing (BESC)*, 2016.
- [16] S. M. Mohammad and P. D. Turney, *NRC emotion lexicon*, National Research Council, Canada, 2013.
- [17] M. Jockers, *Package 'syuzhet'*, 2017.
- [18] P. De Meo, E. Ferrara, G. Fiumara and A. Provetti, "Generalized Louvain method for community detection in large networks," In *Proc. of the 11th Int. Conf. on Intelligent Systems Design and Applications*, 2011.
- [19] W. G. Jacoby, "Loess: a nonparametric, graphical tool for depicting relationships between variables," *Electoral Studies*, vol. 19, no. 4, pp. 577-613, 2000.
- [20] M. L. Jockers, "Syuzhet: Extract Sentiment and Plot Arcs from Text," 2015.

Transparent conversational agent systems for intelligence analysis

Sam Hepenstal

Defence Science Technology Laboratory (Dstl)
UK

shepenstal@dstl.gov.uk

ABSTRACT

Big data presents a problem to intelligence analysts who need to extract insights from the data and advise decision makers. Traditional methods to explore and filter large datasets are complex and require manual configuration of search parameters or query syntax. A more natural approach to retrieve interesting data could be provided via a conversational agent (CA), where analysts can interact with available information through natural language. This can speed up processing of big datasets; however there are critical flaws in existing CA systems which must be resolved before they can provide good situation awareness. In this paper we discuss two such flaws, system transparency and brittleness. We present our work to date, including the development of an ‘algorithmic transparency framework’, published at IUI 2019. This framework identifies the need for analysts to be able to inspect and verify system goals and constraints, in addition to explaining specific outputs. We also discuss cognitive task analysis interviews with analysts which underpin the framework, published at HFES 2019. Our research has led to the development of a prototype transparent CA system for intelligence analysis. In this paper we define the next steps, including experimenting and improving the prototype with operational intelligence analysts.

1.0 PROBLEM OUTLINE

Intelligent systems for performing intelligence analysis provide a high-level capability which is a combination of multiple cognitive computing capabilities. Braines et. al. define this high-level capability as coalition situational understanding (CSU). Specific cognitive computing capabilities include “human-computer collaboration, knowledge representation and reasoning, multi-agent systems, machine learning, natural language processing, and vision/speech processing.” (Braines et. al. 2019) One example of a system for CSU is a conversational agent (CA) which can perform intelligence analysis. A CA typically comprises various complex processes, including natural language processing of a user’s question and intention, identification of an appropriate action which triggers additional data processing, and formulation of responses. Each of these aspects can be trained and can learn through interactions with an analyst.

Artificial intelligence (AI) based CA technologies are still in their infancy. However, their popularity is increasing (Kinsella 2018, Kinsella 2019) because they provide a more intuitive, natural, and quicker way to access information. Typical uses of CA technologies involve low risk and low consequence scenarios. For example, when someone uses Amazon Alexa to select and play music tracks. It does not matter to the user exactly how Alexa has interpreted their question and reasoned what track is most appropriate, so long as a satisfactory track is played. If Alexa gets the selection wrong, the user will easily recognise the error and can ask a different question. So long as Alexa is not repeatedly incorrect, the user does not require a detailed understanding of the underlying algorithms which process the data. While the potential benefits of using CAs in high risk and high consequence decision making environments are large, there are critical issues of *transparency* and *brittleness* which must first be addressed.

In this report we discuss our work to date, including the development of an ‘Algorithmic Transparency Framework’ which was presented in a paper at IUI 2019. In our framework we define *transparency* as the ease with which a user can (i) explain any results provided by a system, in addition to (ii) being able to

inspect and verify the goals and constraints of the system within context (Hepenstal et al. 2019). *Brittleness* (iii) occurs when the technology fails to cope with the variety of demands that it has to cope with when in use. We look to address these three important aspects of CA design. Here we outline the problem faced regarding CA technologies and their use for intelligence analysis. We propose that algorithmic transparency is a crucial requirement of CA systems and that this, underpinned by research into areas such as human cognition and sensemaking, should run through their design. Our framework has been informed and developed by interacting with operational intelligence analysts, including through Cognitive Task Analysis (CTA) interviews, for which we are presenting a paper at HFES 2019. Other intelligence analysis related aspects of using algorithmic systems, such as trust, have also been explored, and we are part of a panel paper on this topic at HFES 2019. Our research has led to the early development of a prototype CA system. In this report, we also discuss the next steps for our research, including experimentation with operational intelligence analysts. Our aim for this research is to understand how algorithmic transparency can be designed into advanced CA systems, and to demonstrate this in an evaluation of a prototype for intelligence analysis.

2.0 INTRODUCTION

CA technologies such as Google Home, Siri and Amazon Alexa present us with an easy way to access music, films, or plan our day. Many services have incorporated chatbots into existing processes to manage interactions with customers, including to direct them to the right information or department. This saves companies money and can save customers time waiting in a queue. Typical applications for conversational agents tackle concise user tasks for mundane processes which can be translated to a finite set of user intentions. Here the risks of an incorrect or misleading response are low and the resulting consequences limited, particularly given the ease with which a user can validate results against an expected and desired conclusion to their interaction. Traditional CAs have, therefore, not been built with algorithmic transparency in mind. If you ask Google Assistant, for example, why it has provided a particular response it will not be able to tell you and instead responds with humour, such as “*Let’s let mysteries remain mysteries.*” The nature of CA technologies and the ability to interact in a natural human way can actually amplify misunderstandings. The tasks which traditional CAs can perform are constrained to a predefined set of ‘intentions’, which can lead to frustration where users expect the CA to behave like a human.

Typically, when a user asks a question the CA attempts to match their input to an intention. For example, if we develop a CA with Artificial Intelligence Markup Language (AIML), a commonly used approach to develop chat interactions which supports many chatbots platforms and services (Radziwill and Benton 2017). A pattern is described for a task category (intention) together with a template response if the users’ text matches with the pattern. There is a need to define or learn the intentions (and subsequent actions) which the CA can fulfil, and to ensure that they are consistent and distinct so as not to confuse either the pattern matching algorithm or the user. The CA can only respond in accordance with the intention, additionally the intention will consist of processes with their own limitations. The possible intentions and the various constraints are not visible in commercial CA systems, so are not transparent. The predefined nature of intentions and related tasks also creates the critical issue of brittleness in CA systems.

CAs can significantly speed up repetitive information retrieval tasks, when compared to traditional query methods or manual processes, and are therefore also attractive for a wide range of applications which need fast responses in critical situations. The domain of criminal intelligence analysis is one example of an environment requiring critical decision making which could benefit from CA technologies. The volume of data which requires processing by police today is a significant barrier to bringing criminals to justice. In June 2019 Cressida Dick, the Commissioner of the Metropolitan Police, explained that “sifting through vast amounts of phone and computer data is partly to blame (for low solved crime rates) as it slows down investigations”. (<https://www.bbc.co.uk/news/uk-48780585>) This is a key area in which an intelligent CA which can interpret an analysts questions and retrieve the appropriate information, including providing machine reasoning, can have significant impact. There are challenges, however, with reference to high risk

and high consequence decision making environments. Algorithmic transparency is a major flaw in commercial conversational agents, where due to “black box” approaches it is difficult to understand how they have interpreted the user, explored the data, and built a response. Clearly, intelligence analysts need a much greater understanding of the underlying algorithmic processes of a CA than someone selecting music. During analysis the information retrieval tasks are complex and different methods can be applied to achieve the same aim, but depending upon the methods involved and their interpretation an investigation can be guided towards different paths with consequences. Trust in methods to gather information is important for intelligence analysts, who may be required to go to court to explain their analysis and advice. The ethical implication of using artificial intelligence applications is also a key consideration for CAs. In Leslie’s (2019) guide for the responsible design and implementation of AI systems in the public sector, transparency is a requirement in terms of both the product and the processes underpinning AI systems. We look to bridge the gap between theoretical transparency requirements and guidelines and the design of actual operational systems.

To achieve effective teaming between an analyst and a CA the analyst must be able to trust the information retrieved by the CA and the processing involved. A key aspect enabling trust engineering in systems, and addressing ethical concerns, is transparency so that the analyst can predict, interpret and refute any results, acknowledging caveats where they exist (Ezer et. al. 2019). While trust is a crucial element in enabling analysts to team with AI systems, it needs to be handled carefully. In intelligence analysis, where analysts should apply critical thinking and scepticism before accepting hypotheses, the impact of trust is a complex issue where it is damaging if it leads to analytical complacency. If an analyst trusts the results of an algorithm without interpreting the reasoning then their trust could lead to adverse decisions, due to unrecognised bias or deception. Analysts should therefore rightly be wary of trusting algorithmic systems which they cannot understand nor inspect and verify i.e those which are not transparent. Without trust, underpinned by transparency, AI systems will not be used by analysts. One analyst described succinctly that “*an analyst always has to justify (in court) what they have done, and so should the system.*” (Hepenstal et. al. 2019) There is therefore a critical requirement for providing algorithmic transparency in applications for intelligence analysis.

Intelligence analysis requires a fluid system which can adapt to new lines of inquiry and traditional approaches to CA technologies encounter issues of ‘brittleness’. Brittleness occurs when the technology fails to cope with the variety of demands that it has to cope with when in use. Developing a flexible system which can evolve to meet new analyst intentions could help counter issues of brittleness, however such a system also needs to balance algorithmic transparency.

We believe that system transparency is a crucial factor in analyst adoption of AI technologies and propose that by delivering effective transparency of AI reasoning, with a close consideration of human needs, we can design systems which can be used in high risk and high consequence environments. Whilst there is existing work in the areas of human sensemaking, trust and ethics, there is a significant research gap where it comes to translating how these principles can be tangibly captured in the design of algorithms and systems. This research will consider and experiment with specific design requirements and approaches for providing algorithmic transparency for flexible intelligence analysis applications.

3.0 OUR RESEARCH TO DATE

An important focus of our research is to understand what the requirements are for algorithmic transparency of intelligent applications for intelligence analysis. We have considered ‘intelligent applications’ as those which allow for shared human-machine processes when retrieving and analysing information, for example where we use a CA to query a semantic knowledge graph. An analyst will be reasoning about the situation they are faced with and this directs the questions they ask, the CA then interprets their questions and decides upon an appropriate set of processes to provide responses. These processes could include graph traversal methods and clustering algorithms which use understanding and

inferencing from the semantic structure of the underlying knowledge graph data. For example, if an analyst asked a question about “*what vehicles are linked to known offenders in the database?*” The agent could interpret from the question that the analyst is looking for connections between instances of vehicles and instances of people who have had a role as an offender. Whilst it is quicker and easier for an analyst to state questions in natural language, their interpretation is subjective and thus, raises the importance of transparency for CA responses. In the example given, what do we mean by ‘linked’? Different people will have different interpretations and will therefore address answering the question with a variety of approaches, each with its own caveats and purpose. This is also true of a CA and the respective goals and constraints of any process must be clearly revealed so that an analyst can inspect and verify the system.

To date much research in the area of eXplainable Artificial Intelligence (XAI) has focussed upon the underlying model features of a specific classification result, for example explaining the relative importance of different features in an image to its classification (i.e. deep Taylor decomposition). Gilpin et al. (2018), gives the definition that XAI is a combination of interpretability and completeness, where interpretability is linked to explaining the internals of a system and completeness is to describe it as accurately as possible. An explanation of the result of an algorithm, for example the features which are important in providing an image classification, are important to provide understanding to an analyst for why a classification was given. However, this is not the complete picture when we consider the need for analysts to be able to inspect and verify a system of multiple processes. We propose that, in the field of intelligence analysis, the explanation for a result is important, but so is the approach taken to reach it including the related goals and the constraints.

3.1 Algorithmic Transparency Framework: Published IUI ATEC March 2019

We have developed an initial research framework for providing algorithmic transparency, informed by former analysts, which highlights areas that we believe need consideration when designing CA systems.

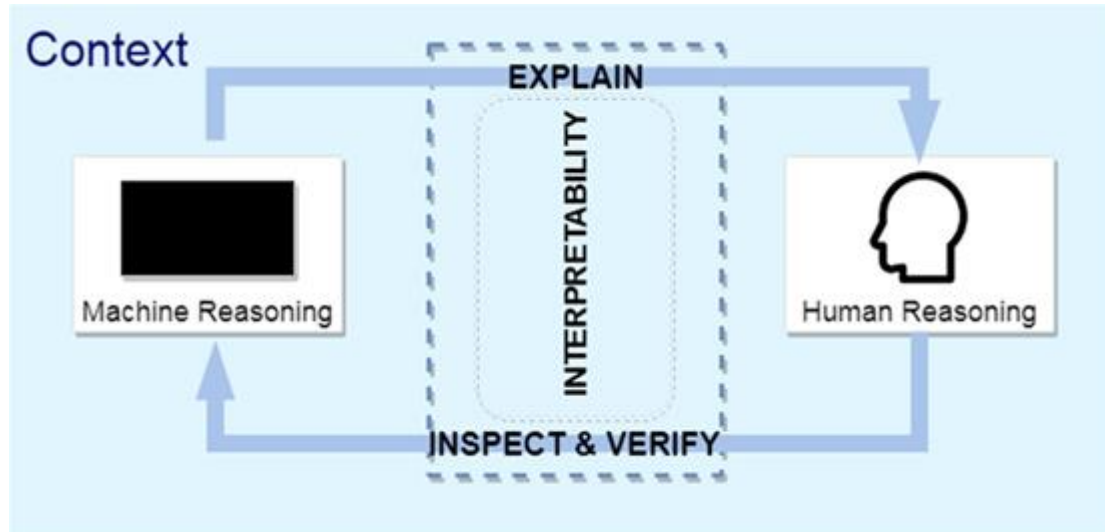


Figure 1: Algorithmic Transparency Framework (Hepenstal et. al. 2019)

Algorithmic Transparency Framework: What the user needs from black box algorithms: (i) explanations of how results from algorithms are arrived at (ii) explanations that are interpretable by the user in a manner that makes sense to them (e.g. the internals of the algorithm, including important features, an indication of accuracy or confidence, and an understanding of the data used and uncertainties, all presented in a manner which enables the user to assess if the results are sensible), (iii) visibility of the functional relationships mapped against the goals and constraints of the system, and (iv) context in which to interpret the explanations. NB: by showing goals and constraints, we include some key elements of

context, e.g. goals include some notion of the priorities and therefore some understanding of the problem, hence the context.

The Algorithmic Transparency Framework shows how transparency relates to all aspects of the application beyond the user interface. This includes modelling the data, for example the classes and properties of the knowledge graph, and processes involved in machine reasoning such as graph traversal algorithms, so they are open to inspection and verification. Transparency cannot simply be provided as a post facto explanation or visualisation of results. Context is an important aspect of the framework and to design appropriate transparency for intelligence analysis applications we must first gain a clear understanding of the context, including specific goals and needs of analysts in different situations.

3.2 Analyst CTA Interviews: Awaiting Publication HFES October 2019

We have developed an initial research framework for providing algorithmic transparency, informed by former analysts, which highlights areas that we believe need consideration when designing CA systems.

To gain an understanding of analyst priorities, needs, and context, and to verify our framework, we have conducted four in-depth interviews with experienced intelligence analysts applying Cognitive Task Analysis (CTA). Three analysts have worked in policing, across various police forces, and the other analyst has a background in defence intelligence. Each analyst has over 3 years of experience working on a range of operational investigations, with a focus upon major crime. We chose CTA because our aim is to understand intentions which underpin how an analyst thinks and reasons. Each interview lasted an hour and applied the Critical Decision Method (CDM) (Klein 1989; Wong 2003) to elicit analyst expertise, cues, goals and decision making on a memorable investigation they were involved with from start to end. The CDM interview technique was used to ensure important information was captured. Of particular interest were the nature and requirements of analyst questions at critical stages in investigations, specifically, their cues, goals, expectations and actions. These stages are typically time-pressured and are therefore prime situations in which CAs could assist analysts. Interviews addressed the analyst's experience, such as the conditions which allowed them to use their prior knowledge and the recognition of situations which a novice analyst may have missed. A timeline of key events was sketched out by the analyst and explored in detail.

Through our interviews we learned key insights into analyst thought processes and tasking which have helped us to consider how CAs can be designed transparently. One interesting finding relates to the scope of an investigation where this is crucial to direct intelligence analysis and enable situation recognition. Scope creates a basis from which expectancies can be drawn. In all interviews it was a key priority for analysts to identify a scope so that investigations could advance. Narrowing scope, however, inevitably leads to constrained investigations and potential bias. AI systems could provide significant benefit if they are able to expand the scope without causing extra burden upon analysts, for example by autonomously checking alternative hypotheses throughout an investigation. The investigation process described in interviews involved repetitive searching for information, where once a search returned a result many additional searches would be triggered. This is a process which could be eased significantly through the introduction of a CA.

Our interviews confirmed the need for an AI system to be open to inspection and verification. We also validated the recognition-primed decision (RPD) model (Klein 1993) as a way to understand how analysts recognise situations in an investigation and then respond. Klein describes “four important aspects of situation assessment (a) understanding the types of goals that can be reasonably accomplished in the situation, (b) increasing the salience of cues that are important within the context of the situation, (c) forming expectations which can serve as a check on the accuracy of the situation assessment (i.e., if the expectancies are violated, it suggests that the situation has been misunderstood), and (d) identifying the typical actions to take.” (Klein 1993) We propose that the factors which feed into the RPD model to aid analyst recognition could also aid a CA's recognition of a situation. Furthermore, if the CA acts upon

components which are designed around the RPD model we propose that this will significantly aid transparency because the CA can present the information it has recognised, and associated goals and constraints, in a way which mirrors analyst recognition. Table 1 presents a generic approach to answer analyst questions using the RPD model, which could help define transparent CA intentions.

Table 1: Consolidated Decision Analysis Table (Hepenstal et. al. 2019)

Hypothesis Scope Assess 5WH

CUES	Inputs for 5WH (person's name, vehicle registration, time span etc...) and relationships where necessary.
Goals	To retrieve summary information, or specific details
Expectancies	Expected event pattern for scope informed by past events with similar scope (experience).
Actions	For information retrieval these include: adjacent information (i.e. who is registered to phone number, or who are their associates), connected information (i.e. what associates linked to a telephone number called by an offender live in a particular location), common connections (i.e. in what locations have both phone numbers been together) amongst others.
Why?	To build on, refute, or confirm scope and associated pattern.
What for?	To advance the investigation

4.0 OUR RESEARCH NEXT STEPS

We propose that algorithmic transparency, underpinned by human factors research, must run throughout the CA system design and cannot be added as an afterthought. We seek to capture design requirements, applying and developing our algorithmic transparency framework, in order to demonstrate the impact of transparency provision upon analyst trust and use of the system. We intend to capture data from users to identify needs for the specific context of a CA for intelligence analysis. We will perform additional scenario based interviews with operational analysts, with a focus on answering the question; what do analysts need to understand about the output, the processes, and the system to have confidence in using a CA? Initially we focus upon the transparency needs for CAs to support information retrieval tasks in investigations. However, we also wish to investigate CAs which can perform additional capabilities, such as prompting and recommending courses of action, or performing elements of investigation autonomously.

There are three specific areas for research which are aligned to important technical aspects of CAs. Firstly, a CA must be able to interpret a user's intentions from their speech or text. Work to date has considered how CAs can dynamically model, learn and recognise analyst intentions, and provide transparency for their thinking and responses (Hepenstal et al. 2019). However, this is yet to be demonstrated and tested in

a working prototype. We will seek to achieve this. Secondly, a CA needs to perform actions once it understands an analyst's intended goals, for example, to trigger an algorithm to retrieve data and follow investigative lines of inquiry. Our proposal is to assess what these possibilities are and how they can be captured, explained, and evolved when an analyst interacts with the CA. Finally, a CA needs to respond to the analyst. In intelligence analysis it is important that responses mitigate bias and provide clear understanding.

4.1 CA Prototype Application

We are first developing a CA system for information retrieval which is underpinned by human factors principles, specifically in the first instance the RPD model. We have applied a novel approach to design possible intentions for the CA which can evolve through interactions with an analyst. In our prototype the tasks which the CA performs are defined by a collection of functional modules, each linked to a specific component of the RPD, which are combined within distinct intention concepts. This approach allows us to effectively design capabilities which can evolve through analyst interaction, and can be inspected according to human factors principles for situation recognition. We have translated human factors research principles into functional components of an application. By developing our CA with the RPD principles at the core of intention design we propose that we can provide transparency so that analysts will be confident to use the application and to trust its responses.

For the next phase of research we intend to run experiments with the CA prototype application and operational intelligence analysts, as they use the system to complete an analysis task requiring repeated information retrieval steps. This will allow a variety of approaches to transparency, including different methods for output explanation and for visibility to inspect and verify system processes, to be tested and compared qualitatively. Of particular interest is the impact of transparency provision upon an analysts trust in a system and the subsequent decisions which are made, including whether the analyst feels comfortable to accept the reasoning and response of the CA and if they are aware of any caveats to the information provided. We plan to experiment with different algorithmic features, or action modules within the RPD structure, for example a deep neural network method to find similarities between instances in the data and a more transparent approach to do the same.

The design of our CA allows the possible tasks it can perform to evolve over time. This is a particularly critical issue for trust and transparency, where an analyst needs to be cognisant of the fact that the CA is learning and may respond differently in the future, even if this may be contrary to the analyst's expectations. While we have chosen the RPD model in the first instance because it suits the recognition nature of the information retrieval tasks described in interviews, we believe that our CA design could make use of alternative models for human cognition. An interesting model we are planning to experiment with is argumentation (Toulmin 1958), where you may task a CA with checking your arguments and looking to find refuting or supporting information. Additionally, we plan to experiment with more advanced CAs, such as those that can conduct investigations with greater autonomy, for example to predict paths of inquiry. Such applications must be designed carefully. It would be damaging if a CA were able to influence the course of an investigation in a way that is not supported by evidence, for example by changing analyst behaviours or questioning strategies. Perhaps with greater transparency analysts will be better able to identify bias, or deception, and challenge a narrow investigation scope more effectively. The impact of bias, however, will likely be subtle and stretch over multiple interactions. We will therefore need to consider how to explain the entirety of the conversation and selected investigation path in the context of alternatives.

5.0 REFERENCES

- [1] Braines, D., Tomsett, R., Preece, A., (2019). Supporting user fusion of AI services through conversational applications
- [2] Ezer, N., Bruni, S., Cai, Y., Hepenstal, S., Miller, C., & Schmorrow, D. (2019). Trust engineering for human-AI teams. Paper presented at HFES 2019,
- [3] Gilpin, L. H., Bau, D., Yuan, B. Z., Bajwa, A., Specter, M., & Kagal, L. (2018). Explaining explanations: An approach to evaluating interpretability of machine learning
- [4] Hepenstal, S., Kodagoda, N., Zhang, L., Paudyal, P., & Wong, B. L. W. (2019). Algorithmic transparency of conversational agents. Paper at IUI ATEC,
- [5] Hepenstal, S., Wong, B. L. W., Zhang, L., & Kodagoda, N. (2019). How analysts think: A preliminary study of human needs and demands for AI- based conversational agents. Paper at HFES 2019,
- [6] Kinsella, B. (2018, 26 Dec). Amazon echo device sales break new records, alexa tops free app downloads for iOS and android, and alexa down in europe on Christmas morning.
<https://voicebot.ai/2018/12/26/amazon-echo-device-sales-break-new-records-alexa-tops-free-app-downloads-for-ios-and-android-and-alexa-down-in-europe-on-christmas-morning/>
- [7] Kinsella, B. (2019, 7 Jan). NPR study says 118 million smart speakers owned by U.S. adults.
<https://voicebot.ai/2019/01/07/npr-study-says-118-million-smart-speakers-owned-by-u-s-adults/>
- [8] Klein, G. A., Calderwood, R., & MacGregor, D. (1989). Critical decision method for eliciting knowledge.
- [9] Klein, G. (1993). A recognition primed decision (RPD) model of rapid decision making.
- [10] Leslie, D. (2019). Understanding artificial intelligence ethics and safety.
- [11] Radziwill, N. M., & Benton, M. C. (2017). Evaluating quality of chatbots and intelligent conversational agents.
- [12] Toulmin, S. E. (1958). The uses of argument, Cambridge University Press.
- [13] Wong, B. L. W. (2003). Critical decision method data analysis, Lawrence Erlbaum Associates.

Determining the Effectiveness of Representing Intrusion Detection System Log Files with Visualization Techniques

Gina Thomas and Kristen Liggett

Collaborative Interfaces and Teaming Branch, Interfaces & Applications Section Air
Force Research Laboratory, 2255 H St, Wright Patterson AFB, OH 45433
USA

gina.thomas.2@us.af.mil; kristen.liggett@us.af.mil

ABSTRACT

After several knowledge elicitation activities with two intrusion detection system (IDS) analysts, two different visualization techniques were selected that best portrayed the most important information relevant to a cyber security task. These two visualizations (customized parallel coordinates and scatter plots) were then investigated for their effectiveness in portraying potential attack patterns to novices. Both visualizations were also subjected to a distortion factor and a complexity factor to determine the robustness of the visualization techniques. A total of 44 volunteers participated in the study. Results revealed a number of significant interactions but overall, performance data showed that scatter plots consistently provided faster response times while parallel coordinates showed more consistent accuracy performance. Lessons learned and future research ideas are discussed.

1.0 INTRODUCTION

According to a report published by a security firm last year, the estimated cost of cybercrime is \$113 billion a year and has victimized approximately 378 million people (White Hats, 2014). Not only are cyber crimes costly, they are becoming more common. According to the Ponemon Institute's annual report on the cost of cyber crime (2013), companies experience approximately two successful cyber attacks per week (infiltration of the company's core network) and the most costly are denial of service attacks, malicious insiders, and web-based attacks.

On the military side, the same trend of increasing attacks is occurring (Leopold, 2013). Superiority in the cyberspace domain is essential to maintaining Air Force mission assurance and enabling new mission capabilities. As such, this domain must be mastered in order to ensure success. The budget more than doubled in 2014 from 2013 and is estimated at \$5.1 billion in 2015, a portion to be invested in science and technology programs to support cyber defense (McCaney, 2014). A common response has been to develop new software to automate processes, but no one has been successful at creating computer systems that are immune to the attacks by adversaries who are capable of adapting their weapons and tactics as new defense technologies come on-line (Jackson, 2012). The cyber domain has the unique characteristic in that, the resources necessary for superiority in the domain are not aircraft and airfields as in the air domain, but highly trained and competent personnel (Bryant, 2013). The Department of Defense expects to add 4000 cyber operators over the next four years with the majority slated to defend the networks that support worldwide operations (Costlow, 2013). The success or

failure of cyber superiority rests on the human factor - in the hands of the cyber operator. Therefore, it is imperative that the systems with which these operators interact are designed for effective human use. Understanding adversaries and their intent is still primarily a human skill that cannot be automated with even the most sophisticated computer software (Aucsmith, 2011). “Cyber situational awareness” has become a popular concept (Barford et. al., 2010); however, there is much debate over whom or what should maintain this awareness – a computer or a human operator. Instead of taking the human further and further out the loop (from in-the-loop to on-the-loop to observing-the-loop), perhaps keeping the operators in-the-loop and designing their systems to better support them is the most efficient way to enhance cyber operations in the future.

Putting the human in charge of defending computer networks requires a change in the way in which cyber traffic data are collected, processed, and presented. Currently, cyber analysts sit diligently monitoring vast amounts of cyber traffic trying to determine if said traffic is suspicious or anomalous (D’Amico, Tesone, Whitley, O’Brien, and Roth, 2008). Specifically, we found during our interactions with experts that analysts are responsible for triaging about 3000 intrusion detection system (IDS) events per hour (represented as plain text rows in a table.) Optimizing the presentation of information that analysts need to perform the job based on human cognitive and perceptual skills and abilities, as well as limitations, should make analysts more effective in this task (Goodall, 2005).

Whatever the attack, the keys to combating it are recognizing it and responding to it appropriately. Recognizing an attack involves understanding the signatures and sources of the malicious traffic (Jackson, 2013). Signatures can mean visual patterns, and visual pattern matching is a skill at which humans excel. Using visualization techniques to portray cyber security data has been studied for a number of years (Conti, 2007; Marty, 2009). Common visualizations for cyber network data include scatter plots, node-link diagrams, histograms, tree maps, and parallel coordinates (see Table 3-1 in Marty, 2009). Each of these visualization techniques may be capable of showing patterns or signatures of attacks, some better than others. A typical IDS log file may contain upwards of 10,000 rows of data for a day’s worth of network traffic. This is especially true if the rule set for the IDS is configured to favor false positives over false negatives, which is often the case due to the potential high cost of missing an attack. If the information in thousands of rows of defensive log files can be presented to analysts in a single visualization containing patterns that might imply an attack, this visualization should help analysts detect and respond to cyber attacks more effectively. Visualizations have been proposed to help with this task; however, the visualizations typically encode all of the information contained in the log file without focusing on only the important elements from the analysts’ perspective to detect anomalies (Harrison and Lu, 2012; Goodall, 2009; Thompson, Rantanen, Yurcik, and Bailey, 2007; Livnat, Agutter, Moon, Erbacher, and Foresti, 2005; Koike and Ohno, 2004).

In this paper, we will discuss how we performed knowledge acquisition activities with cyber defenders to determine the most important dimensions of IDS log files and applied this knowledge to the design of two visualizations for testing. The study reported here focuses on presenting cyber defensive data as from IDS log files graphically in scatter plots and parallel coordinates plots to determine which of these visualizations best aids an analyst in detecting attacks. Scatter plots are most effective at encoding two dimensions of data, as these dimensions can be given a direct spatial encoding. Additional dimensions can be encoded using color, shape, size, texture, etc., but these are less effective than spatial encodings and should be reserved for less important dimensions. Scatter plots are typically used to examine data relationships or to detect clusters and trends in the data. Parallel coordinates (Inselberg & Dimsdale, 1991) depict $n-1$ pair-wise relationships out of a possible n^2 number of pair wise relationships, where n is the number of dimensions. In contrast to scatter plots where data points are depicted as points, in parallel coordinates data points are depicted as line segments that connect the parallel pairs of axes. These two visualization techniques were compared using the same data set. Determining the elements to be portrayed with these two visualization techniques and then designing them to highlight these

element may help analysts more quickly and accurately determine the security of their network. The objective of this research was twofold: 1) to determine if IDS log data could be transformed into visualizations in which network attacks appeared as patterns (visual signatures), and 2) test user human effectiveness (time and accuracy) of the developed visualization candidates for portraying attack information.

1.1 The Human Factors Design Process

Our research is grounded in the human factors design process. Operator challenges in the cyber domain are well-suited for the human factors engineering design/cognitive engineering design process because cyber systems provide operators with a large amount of data and designers want to automate more processes (Roth, Patterson, and Mumaw, 2002). This approach has been successfully implemented by many including Mahoney, Roth, Steinke, Pfautz, Wu, and Farry (2010) to design, develop and evaluate a system supporting cyber situation awareness. Also particular to the cyber security visualization domain, Shiravi, Shiravi, and Ghorbani (2012) acknowledge that “a design process centered on the needs, behaviors, and expectations of security analysts can greatly influence and impact the usability and practicality of such systems.” (p. 1313). A diagram of this process is shown in Figure 1 (Kilgore, Godwin, Hogan, Davis, Pfautz, Woods, Branlat, and Kaufman, 2012). It begins with an analysis phase; effective execution of this phase is **crucial** for the success of the design process as it provides the foundation for the design activities. Operators and/or subject matter experts (SMEs) provide the design team valuable understanding of domain information, such as the tasks and processes by which they accomplish their work. Members of the design team can then determine the cognitive processes that need to be supported when the operators are performing these tasks and incorporate this knowledge into the design to support them in a way that facilitates their work flow.

Once the information from the analysis phase is converted into useful design guidelines, the design team can begin designing, refining, and investigating candidate solutions. The iterative design process leverages users at various stages of the design process to ensure the design concepts are supporting them when conducting their tasks.

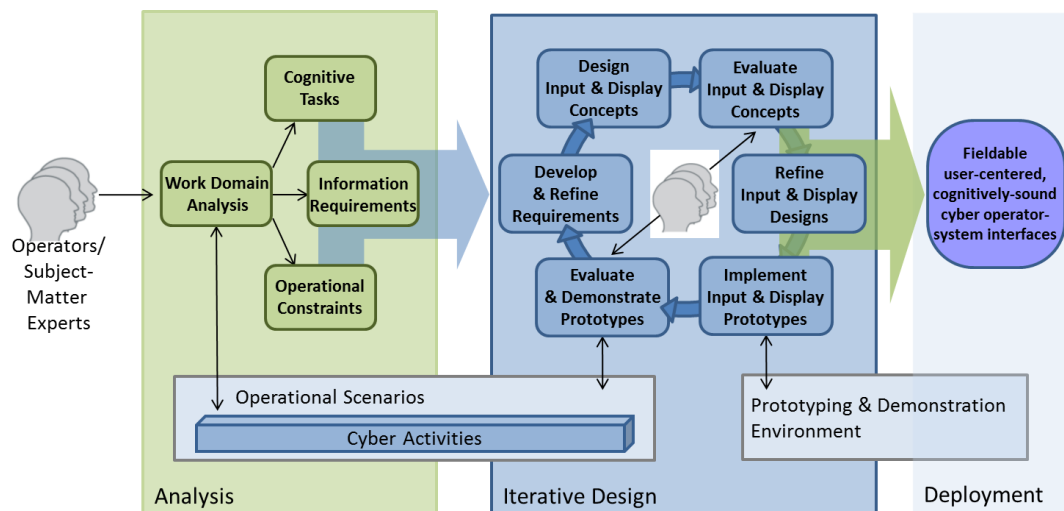


Figure 1: Human Factors Design Process

1.2 Analysis Phase

Cyber analysts have, at their disposal, a large amount of information about the network they are trying to defend. First and foremost, they have an understanding of the machines on the network, each machine's role, and how everything is connected. This information may be in the form of a physical network diagram/document or the analysts' mental model of the network. They also have access to a variety of log files that document all of the traffic on the network. These log files include packet data log files, firewall log files, and IDS log files. Based on the literature and from discussions with IDS analysts for commercial companies, third-party security vendors, and a commercial financial institution, cyber analysts typically monitor events generated by the IDS in near real-time to maintain network security. If data generated by the IDS is deemed anomalous or suspicious by the analyst, they perform further investigation using firewall and packet data log files.

There are three approaches to cognitive task analyses; one of which is using empirical techniques (informal interviews, retrospective analyses of actual incidents, observation of practitioners as they attempt to perform domain tasks, and formal behavioral studies) to analyze how people actually go about performing the task either in the actual task environment or in a simulated task environment (Roth et al., 2002). For this study, the design team conducted a series of knowledge elicitation sessions with two IDS analysts using observations of practitioners performing the task in a simulated environment. Although CTA activities for various defensive cyber analysts have been documented (Mahoney et al., 2010; D'Amico, Whitley, Tesone, O'Brien, and Roth, 2005; D'Amico and Whitley, 2007), none were specifically focused on IDS analysis and contained the level of cognitive requirements necessary to design effective visualizations for their particular tasks. We essentially followed a think-aloud protocol (Rubin, 1994) in which we had the analysts talk through their activities as they reviewed an IDS log file looking for anomalous activity. The IDS log file from the VAST 2011 Mini-Challenge 2 Day 2 (Grinstein, Cook, Havig, Liggett, Nebesh, Whiting, Whitley, and Koneci, 2011) was used for these sessions. Snort was the IDS used for the challenge, and each log entry contains many details. If opened in a plaintext editor such as WordPad, as shown in Figure 2, each log entry takes up to 6 lines and is followed by a blank line before the next entry. The important parts of a log entry are identified in Table 1. Here, the line number refers to the line number of the entry shown in the WordPad example. We parsed the dataset and transformed it into a table representation such that each row represented an IDS event and each column represented one field (Snort rule, text of rule violated, classification of rule, priority of alert, date, time, source/destination IP addresses, etc.). The file contained 22,425 rows corresponding to 7 unique rules firing during Day 2. We provided these data to the analyst in Microsoft Excel. In general, the analysts did a lot of sorting, scrolling, and filtering. Each analyst spent about 45 minutes talking aloud as they tried to determine what, if anything was concerning in the data. At the end of the session, they indicated places where they would investigate further (look at the packet capture data).

```
File Edit View Insert Format Help
[Icons]

[**] [1:100000160:2] COMMUNITY SIP TCP/IP message flooding directed to SIP proxy [**]
[Classification: Attempted Denial of Service] [Priority: 2]
01/27-14:07:55.230358 192.168.3.2:56760 -> 192.168.3.6:5900
TCP TTL:64 TOS:0x0 ID:2512 IpLen:20 DgmLen:52 DF
***A**** Seq: 0xDE93EA5C Ack: 0xA20AA6B2 Win: 0xAC0 TcpLen: 32
TCP Options (3) => NOP NOP TS: 16772556 403107

[**] [1:100000160:2] COMMUNITY SIP TCP/IP message flooding directed to SIP proxy [**]
[Classification: Attempted Denial of Service] [Priority: 2]
01/27-14:08:15.493781 192.168.3.6:5900 -> 192.168.3.2:56760
TCP TTL:64 TOS:0x0 ID:61016 IpLen:20 DgmLen:1341 DF
***AP*** Seq: 0xA25BFD59 Ack: 0xDE93ED54 Win: 0x5B TcpLen: 32
TCP Options (3) => NOP NOP TS: 408172 16774583

[**] [1:100000160:2] COMMUNITY SIP TCP/IP message flooding directed to SIP proxy [**]
[Classification: Attempted Denial of Service] [Priority: 2]
01/27-14:10:23.005903 192.168.3.2:56760 -> 192.168.3.6:5900
TCP TTL:64 TOS:0x0 ID:5361 IpLen:20 DgmLen:52 DF
***A**** Seq: 0xDE93FADA Ack: 0xA2673A56 Win: 0xACB TcpLen: 32
TCP Options (3) => NOP NOP TS: 16787336 440050

[**] [1:100000160:2] COMMUNITY SIP TCP/IP message flooding directed to SIP proxy [**]
[Classification: Attempted Denial of Service] [Priority: 2]
01/27-14:10:24.787077 192.168.3.6:5900 -> 192.168.3.2:56760
TCP TTL:64 TOS:0x0 ID:62900 IpLen:20 DgmLen:88 DF
***AP*** Seq: 0xA27AC01C Ack: 0xDE93FD4A Win: 0x8D TcpLen: 32
TCP Options (3) => NOP NOP TS: 440496 16787514
```

Figure 2: IDS excerpt from Snort opened in WordPad

Table 1: IDS File Content

Line #	Content / Format	Useful Content for this Challenge
1	[**] [<i>Snort Rule File</i>] <i>Text of specific rule violated</i> [**]	Text of specific rule violated
2	[Classification: an optional generalized description of the alert] [Priority of the alert]	Classification (if present) Priority
3	Date/Time MM/DD-HH24:MI:SS.milliseconds Source IP/Port Destination IP/Port	Date/Time Source IP/Port Destination IP/Port
Remaining lines	<i>No content used in the challenge</i>	None

A discussion followed these sessions to ensure that we understood the analysts' procedures and tactics taken during the session. Although there were a total of 7 columns of information available in the Excel file, we discovered that the analysts searching focused primarily on 4 variables – time, source internet protocol (IP) address, destination IP address, and the text of the specific rule violated. They performed sorts on these variables and talked about looking for patterns in the rules being fired or patterns of timing or rules firing, or patterns of source IPs communicating with destination IPs. With that information, we began the visualization design process.

1.3 Visualization Design

The idea that a large amount of IDS alerts could potentially be summarized in a simple visualization (Shiravi et al., 2012) was attractive. Langton and Newey (2010) point out that current cyber security visualization typically focus on static visualizations of specific applications, which is exactly what we propose below, however, they also point out that operators focusing on network monitoring tasks need an overview of activity with the capability to drill down to investigate suspicious activity. Therefore, once our static overview concept is tested and proven, the static representation can be adapted to an interactive visualization allowing which would allow

filtering and a drill down capability to raw data. This methodology supports the visual information-seeking mantra; overview first, zoom and filter, details on demand (Shneiderman, 1996). Since we defined 4 variable of interest, we needed to find visualizations that could support the display of these 4 variables in ways that effectively showed patterns in the data. We started by considering the standard visualization techniques of node-link diagrams, parallel coordinates, scatter plots, and tree maps. We quickly ruled out node-link diagrams because there was no obvious way of representing time, which was a critical variable. We mocked up parallel coordinates with 4 axes (one for each variable) and scatter plots using x position, y position, color of marker, and shape of marker to represent the 4 variables. We were also able to create a variety of nested visualizations embedded in tree maps by partitioning the dataset according to unique rules, sources, or destinations, but because this approach afforded a large number of options (too many to adequately test) we decided to pursue this approach in a follow-up study.

For the parallel coordinates, we went through several iterations of changing axis order, selecting colors, grouping variables into categories versus plotting the actual variables themselves, etc. We considered grouping variables because studies have shown that parallel coordinates suffer issues of occlusion when applied to large data sets (Van Wijk and Van Selow, 1999; Fua, Ward, and Rundensteiner, 1999). The continuous time variable was grouped into 3 categories (morning, work day, and evening); the source IP variable as well as the destination IP variable were organized into 6 categories representing the types sub-network (data center, DC/DNS, office workstations, external web, internet, and unknown); the rule variable was organized into 4 sections based on the rule category (not the individual rule itself). We showed the different versions to our SMEs to get feedback about the variations of visualizations. We settled on a version for which known attacks from the ground truth created discernible patterns. In arranging the rules in this hierarchical way, we discovered that patterns were specific to the rule category. Figure 3 shows a representation of a parallel coordinates visualization and highlights activity indicative of a port scan.

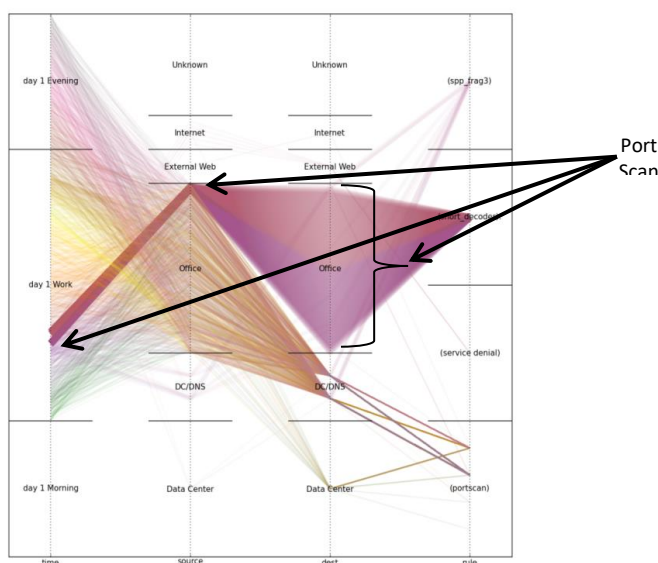


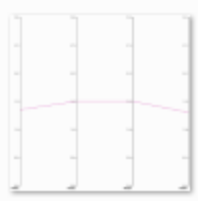
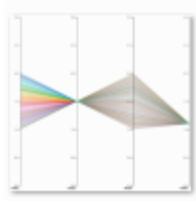
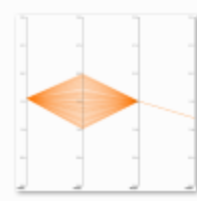
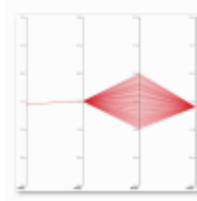
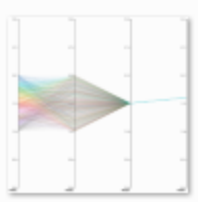
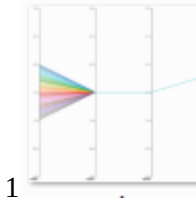
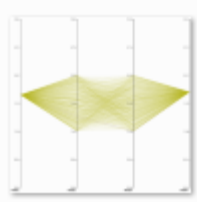
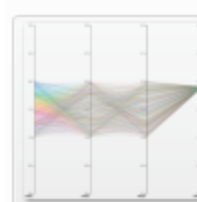
Figure 3: Parallel coordinates representation of IDS log file for Day 1 of VAST Challenge 2011 data. Focusing on the first three axes, a port scan pattern looks like a sideways broom with the handle to the left.

EFFECTIVENESS OF IDS VISUALIZATIONS

To help evaluate our hypothesis that specific attacks might have unique visual signatures when presented in parallel coordinates, we worked with the SMEs to list the most common attacks, then thought about how time, source, and destination would “look” on a parallel coordinates plot for each attack. For instance, a denial of service attack involves a large number of sources communicating with one destination in a very short period of time. If you plot that type of attack on the first three axes of a parallel coordinates visualization shown in Figure 3, it would look like a diamond (a single point on the time axis, many points on the source axis, and a single point on the destination axis). The fourth axis (rule category) always fell into one category because the various attacks only triggered a single IDS rule category. Generalizing these rules to a “one” or “many” indication for each of the 3 primary axes (time, source, destination) allowed us to develop 8 distinct visualization patterns to use as test stimuli.

Although some of the 8 patterns created are not indicative of a known attack, the patterns allow the user to readily sort and group by rule and to recognize things that are more likely to be an attack versus things that are more likely to be noise. Additionally, since each axis has meaning, defenders can derive meaning from the pattern that allows them to make an assessment about which rule violations are more likely to require follow up. For example, a one-many-one relationship of violations of a particular rule tells the user that at a single time (very short time period), multiple IP addresses are pinging a single computer. If, for instance, that computer is a web server, such a pattern might indicate that someone is attempting a denial of service attack. Table 2 shows the 8 parallel coordinates stimulus patterns.

Table 2: Eight Attack Stimuli in Parallel Coordinates

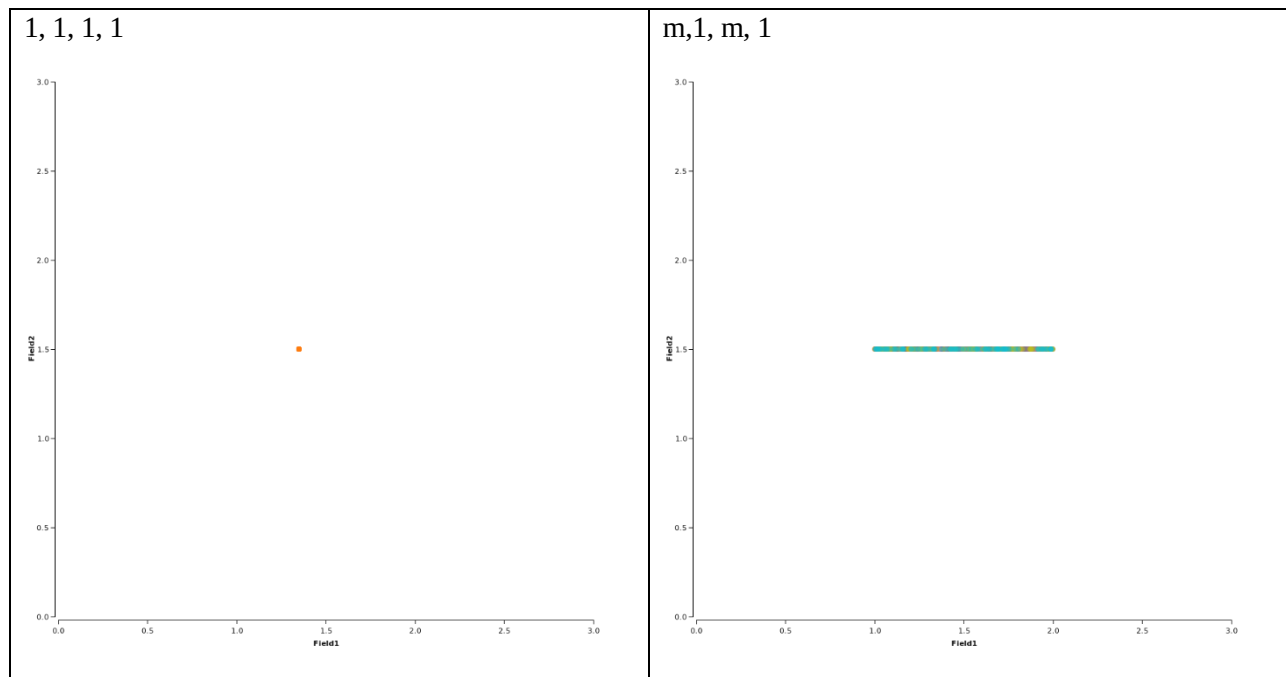
1, 1, 1, 1 	m,1, m, 1 	1, m, 1, 1 	1, 1, m, 1 
m, m, 1, 1 	m, 1, 1, 1 	1, m, m, 1 	m, m, m, 1 

Note: Since the last axis always only contains one point, the patterns may be fully described by first three axes and ignoring the rightmost portion. The patterns are described using commonplace language such as brooms, bowties, houses, etc. For instance, 1,1,1,(1) is a straight line, m,1,m,(1) is a bowtie, 1,m,1,(1) is a diamond, 1,1,m,(1) is a broom with the handle to the left, m,m,1,(1) is a side-way house with the roof to the right, m,1,1,(1) is a broom with the handle to the right, 1,m,m,(1) is a house with the roof to the left, and m,m,m,(1) is a rectangle.

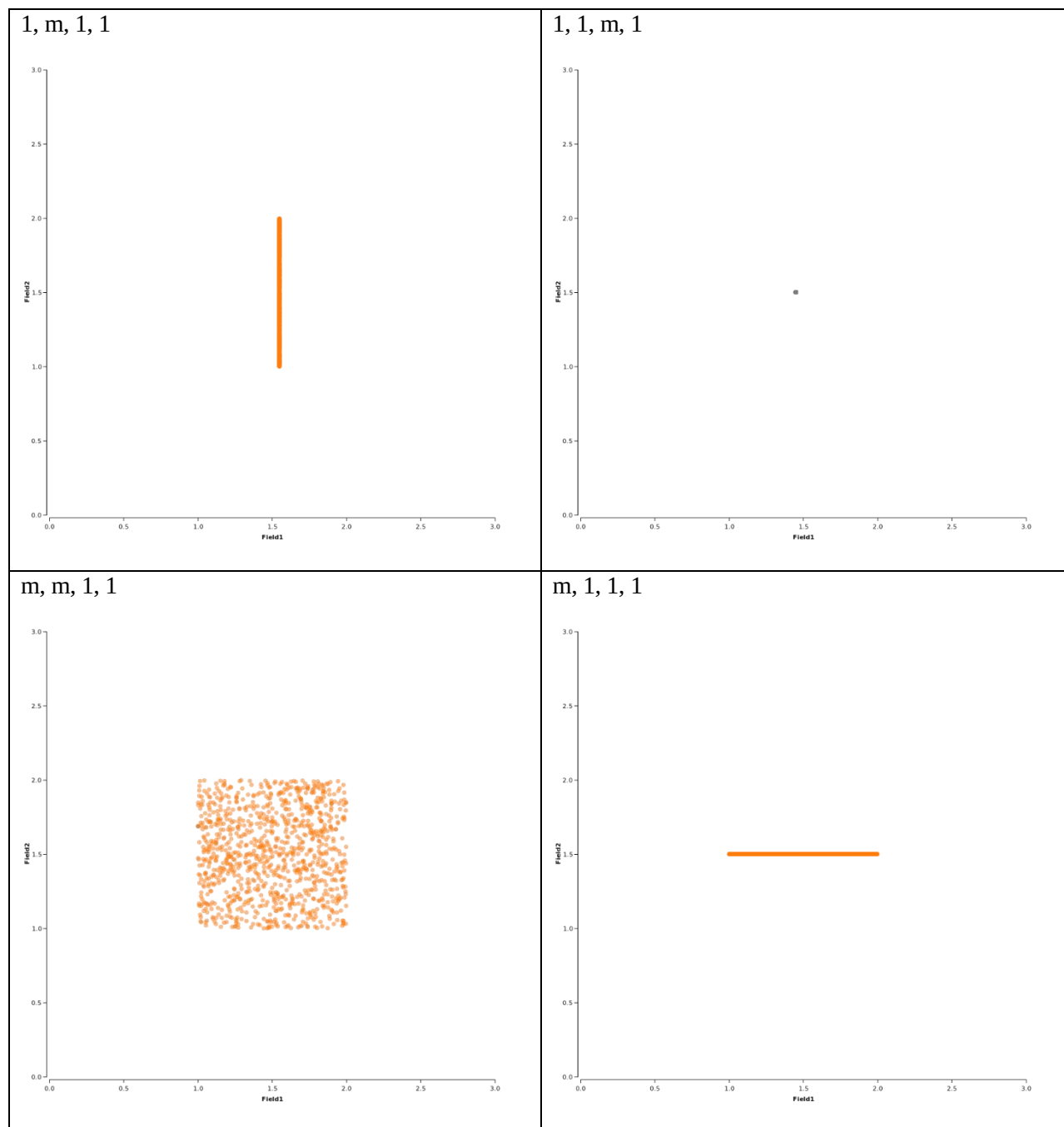
For the scatter plots, we again mocked up many variations of variable assignments among the 4 representational dimensions (x, y, color, and shape). Since the rule category for each type of attack was always the same for a

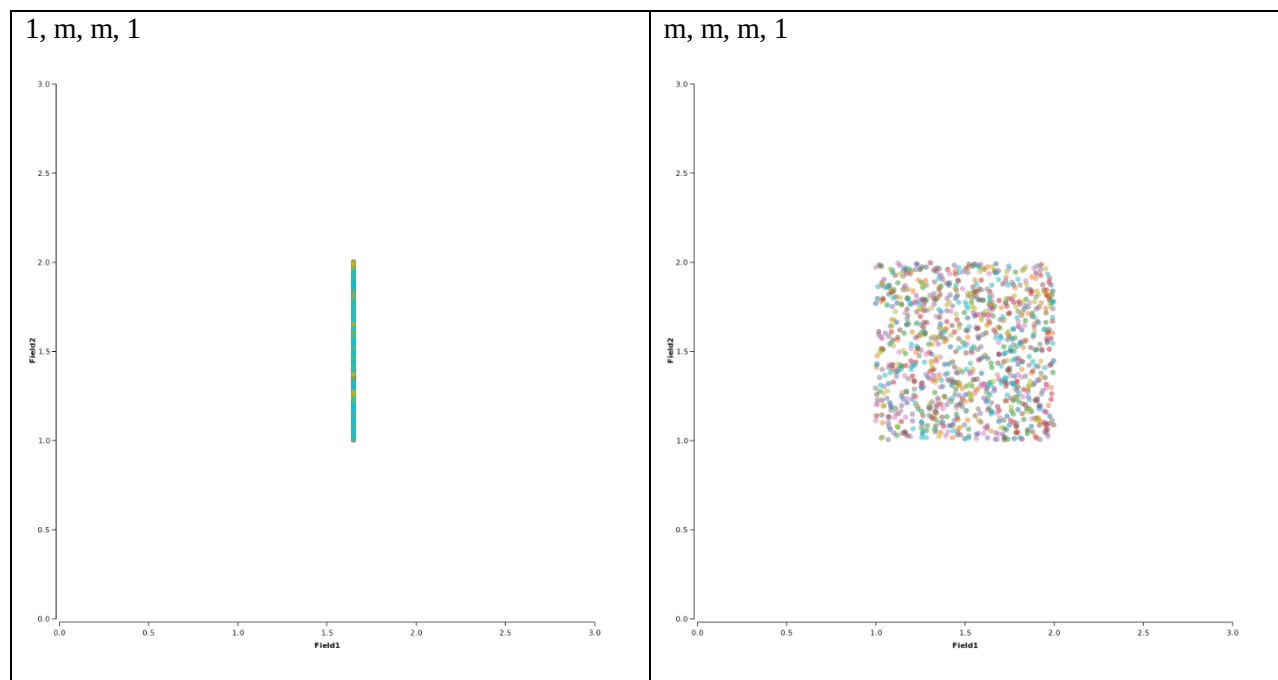
single attack, we set rule category to shape, which is the most difficult of the dimensions to discriminate (each of the 8 patterns were always drawn using the same shaped marker). Time is represented on the x axis, source IP is represented on the y axis, and destination IP is represented by color of marker. Table 3 shows the 8 scatter plot stimuli. The pattern in the 1,m,1,1 block, as in the prior example, might represent a denial of service attack - at 1 time (1 x value), many sources (many y values) sent traffic to one destination (1 color), and only 1 category of rules fired (1 shape of point markers).

Table 3: Eight Attack Stimuli in Scatter Plots



EFFECTIVENESS OF IDS VISUALIZATIONS





Note: Patterns here are 1,1,1,(1) is a single-colored marker in single location, m,1,m,(1) is multi-colored markers in a horizontal line, 1,m,1,(1) is single-colored markers in a vertical line, 1,1,m,(1) is multi-colored markers in single location [appears grey], m,m,1,(1) is single-colored markers in a square, m,1,1,(1) is single-colored markers in horizontal line, 1,m,m,(1) is multi-colored markers in a vertical line, and m,m,m,(1) is multi-colored markers in a square.

2.0 THE RESEARCH STUDY

With 8 distinct attack patterns and 2 visualization techniques, we had the basis for a study. The primary objective of the study was to compare the effectiveness of parallel coordinates and scatter plots as a means of portraying the data in IDS log files. More specifically, we were interested in whether these techniques would be effective in portraying attack patterns when the attacks were embedded in noise due to normal activity occurring at the same time as an attack, and when the patterns were distorted due to particular axis values representing different variable values. In other words, would defenders be able to pick out the attack patterns even if there were numerous attacks and/or among rule violations that were not indicative of an attack, as well as when the patterns were distorted by the position of the values along each axis? We, therefore, considered 2 factors in addition to the visualization type to be varied in this study. These factors included distortion of the pattern, and the amount of visual density each pattern contained both individually and when additional patterns were added to them (noise). Figure 4 shows both scatter plots and parallel coordinates with distortion and noise.

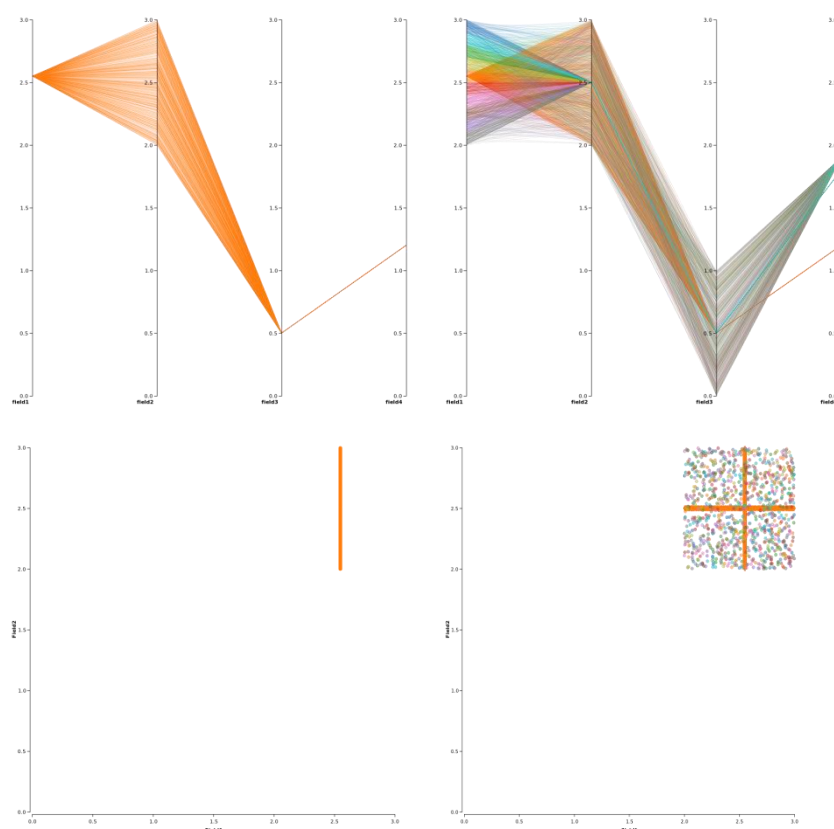


Figure 4: Parallel coordinates and scatter plots with distortion and noise. Top visualizations are parallel coordinates; bottom visualizations are scatter plots. Top left is the 1,m,1,1 (diamond) attack pattern with distortion level 2; top right is the 1,m,1,1 (diamond) attack pattern with noise patterns m,1,1,1 (broom with handle to right) and m,m,m,1 (rectangle); bottom left is 1,m,1,1 (single-colored vertical line) attack pattern with distortion level 2; bottom right is the 1,m,1,1 (single-colored vertical line) attack pattern with additional noise patterns m,1,1,1 (single-colored horizontal line) and m,m,m,1 (multi-colored square).

2.1 Method

2.1.1 Participants

A total of 44 participants were recruited from Wright-Patterson Air Force Base and Wright State University. Wright-Patterson AFB participants were recruited via e-mail and flyers. Wright State Introductory Psychology students were recruited by way of the university's research participation system (a web-based system developed by Sona Systems and used by the Department of Psychology for students to sign up to participate in approved experiments). These students are required to participate in a minimum of 8 hours of research as part of the Introductory Psychology course and receive extra credit for further participation.

2.1.2 Stimuli

Stimuli for the study consisted of various representations of data in both parallel coordinates format and scatter plot format. To make the stimuli less cluttered and eliminate the instinct of people reading the legends, we removed all textual information from the stimuli (as shown in Tables 2 and 3 and Figure 4). The objective of the

task was to have participants determine whether there was an attack pattern in each of the visual stimuli presented. The study employed 5 independent variables; visualization type (parallel coordinates vs. scatter plots), level of distortion (5 levels), level of complexity (a discrete ordinal variable determined by the sum of the visual density of each attack pattern and the pattern or patterns of noise displayed with it; range in this experiment is from 5 to 18), presentation order or “session” (parallel coordinates set first vs. scatter set first), and attack pattern set (which 4 of the 8 patterns were designated as attacks; 2 categories). In further explanation of this last variable there were 8 basic patterns; 4 were defined as “attack” and 4 were defined as “noise” for each of 2 sets. In order to ensure that results are not restricted to particular patterns of attack, we reversed the definitions (patterns that were considered noise in one case were defined as attacks in another and vice versa). The attack pattern sets were selected such that in one set, the overall complexity of the noise was greater than that of the attacks (simple attacks in complex noise) and in the other, complexity was greater for the attacks than the noise (complex attacks in simple noise). For instance, simple attacks were (refer to Tables 2 and 3) 1,1,1,1; m,1,m,1; 1,m,1,1; and 1,1,m,1; and complex noise for those attacks were m,m,1,1; m,1,1,1; 1,m,m,1; and m,m,m,1. This situation was reversed for the other attack pattern set (complex attack patterns were m,m,1,1; m,1,1,1; 1,m,m,1; m,m,m,1 and simple noise patterns were 1,1,1,1; m,1,m,1; 1,m,1,1; 1,1,m,1;). Since it would be difficult for participants to consider a particular pattern as an attack in one situation and as noise in another situation, we treated this variable as a between-groups variable (half of the subjects were told patterns 1-4 were attacks and patterns 5-8 were noise; the other half of the subjects were told that patterns 1-4 are noise and patterns 5-8 were attacks).

2.1.3 Experimental Design

The study utilized a mixed design; attack pattern set was varied between groups and visualization type, level of distortion, level of complexity, and session were varied within-subjects. Each block of trials contained 2 sets of 6 noise only pattern combinations by 5 levels of distortion plus 40 combinations of attack plus noise patterns by 5 levels of distortion, yielding 60 noise only and 200 attack plus noise trials for a total of 260 trials per participant. The dependent measures collected included time to complete each trial and the accuracy of the response. Each presentation was displayed for a maximum of 10 seconds, and participants were instructed to respond as quickly and accurately as possible. Each block took approximately 45 minutes on average. Assignment of the session and attack pattern sets were systematically varied (the 4 possible combinations were alternated) between participants, and trials were randomly presented within blocks.

2.1.4 Hypotheses

Overall, we were uncertain as to which visualization type would provide the best performance. The parallel coordinates had more visual content due to the inherent layout of the 4 axes, and every variable was presented spatially. However, the scatter plot patterns were more simplistic (single location markers, lines, and squares). Also, the scatter plots seemed to be more robust to distortion as changing axis values in the scatter plots simply relocated the stimuli in the scatter plot space (refer to Figure 4). Therefore, we anticipated an interaction between visualization type and distortion with scatter plots being less affected by distortion than parallel coordinates. In terms of complexity, the parallel coordinates were effected in a uniform way as increasing complexity resulted in overlaying patterns on each other, whereas for the scatter plots, since some of the patterns were single markers, increasing complexity might “hide” the single location patterns. Therefore, we hypothesized that the scatter plots would be more negatively affected as complexity increased. We also anticipated an interaction between complexity and attack pattern set as Attack Pattern Set A contained simpler attack patterns and Attack Pattern Set B contained more complex attack patterns, which meant that higher levels of complexity in B did not decrease the “signal to noise ratio” nearly as much as with A. Therefore, it was hypothesized that Attack Pattern Set B would be unaffected by higher complexity levels but Attack Pattern Set A

would be affected by high complexity levels. We expected to see an effect for session (essentially a training effect), so we counterbalanced presentation order to control for systematic learning effects.

2.1.5 Procedure

Upon arrival, the research objectives and procedures were explained, and participants were asked to read and sign an informed consent form and were assigned a participant number. Participants were then administered a color vision test (Ishihara plates). Which visualization type (either parallel coordinates representation or scatter plot representation) was presented first and which set of attack patterns that they were assigned was dependent on their participant number. They were shown pictures that contained simple patterns representative of four cyber attacks as well as examples of distorted versions of those attacks. Once they were comfortable that they knew what the attacks looked like, they began a training block in which sixteen attack and noise patterns were presented at various levels of distortion, and they had to determine whether each was an attack or noise pattern. Feedback was provided. Training blocks were repeated until they were able to identify all attacks and noise with 100% accuracy, at which time data collection proceeded.

During each block of data collection, participants were simply asked whether an attack was present in the stimuli. Participants were informed that there was not a 50% chance that an attack was present so they should not make an effort to balance their responses. Upon completion of the first block of data collection, participants were offered a break. Then, the process of training and data collection was repeated for the other visualization type. Another break was offered, and then participants were asked to complete a brief questionnaire to record any challenges, observations, or comments they might have.

2.2 Results

2.2.1 Main Effects

A total of 44 participants were recruited from Wright-Patterson Air Force Base and Wright State University. Wright-Patterson AFB participants

2.2.1.1 Visualization type

With regard to accuracy, there was a small but statistically significant main effect of visualization type, $F(1,519) = 5.70, p < 0.05$. Overall, participants were slightly more accurate when determining whether attacks were present in scatter plot ($M = 75.4\%$) than in parallel coordinates ($M = 74.8\%$). There was also an effect of visualization type on response time, $F(1,519) = 2163.21, p < 0.05$. On average it took 2.84 seconds for participants to decide whether an attack was present in parallel coordinates visualizations and about 1.71 seconds to make the determination when viewing scatter plots. Since time varied negatively with accuracy, there was not a speed accuracy trade-off that needs to be considered. In general, participants who were more accurate tended to have shorter response times. Visualization type also interacted with several other variables; those interactions will be discussed below.

2.2.1.2 Other variables

There were main effects of distortion, both for response time, $F(4,519) = 19.90, p < 0.05$ and accuracy, $F(4,519) = 10.64, p < 0.05$ as well as on complexity for both response time, $F(13,519) = 21.18, p < 0.05$, and accuracy, $F(13,519) = 76.76, p < 0.05$. There also were main effects of attack pattern set on response time, $F(1,519) = 184.36, p < 0.05$, as well as on accuracy, $F(1,519) = 522.16, p < 0.05$. Finally, there were significant main

effects of session for response time, $F(1,519) = 233.87$, $p < 0.05$, and accuracy, $F(1,519) = 4.57$, $p < 0.05$.

2.2.2 Two-Way Interactions

2.2.2.1 Visualization type by distortion

The effects of distortion significantly interacted with visualization type for both response time, $F(4,519) = 29.08$, $p < 0.05$ and accuracy, $F(4,519) = 15.20$, $p < 0.05$. Response times were consistently better with scatter plots than with parallel coordinates; this effect increased at higher levels of distortion. Also, parallel coordinates resulted in higher accuracies than scatter plots for low levels of distortion, but parallel coordinates accuracies fell below those of scatter plots at the highest level of distortion (Figure 5).

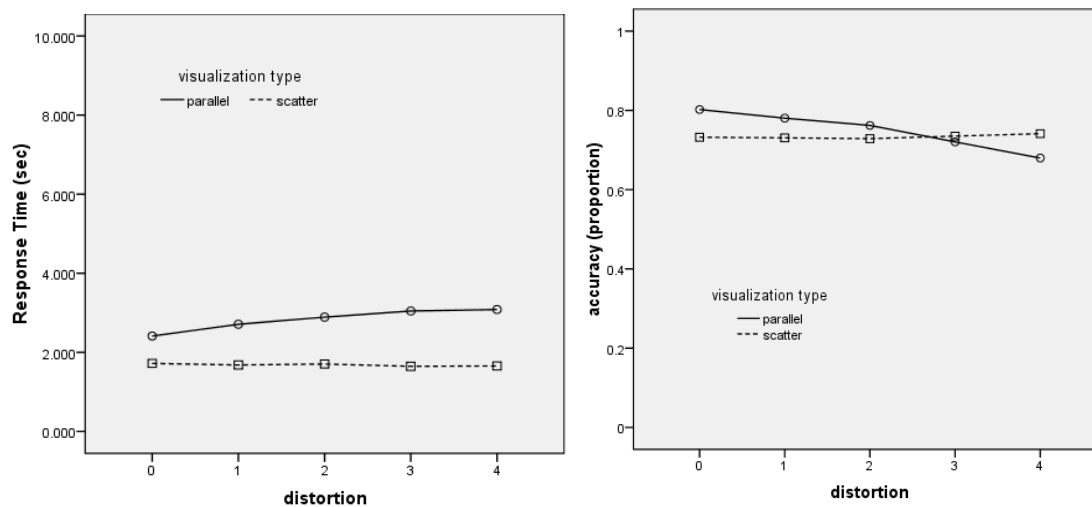


Figure 5: Interaction between visualization type and distortion for response time and accuracy.

2.2.2.2 Visualization type by complexity

Complexity interacted with visualization type for both response time, $F(13,519) = 12.77$, $p < 0.05$ and accuracy, $F(13,519) = 42.35$, $p < 0.05$. In general, complexity had a greater effect on response times of parallel coordinates than those of scatter plots. However, scatter plots showed much more variation in accuracy across levels of complexity (Figure 6).

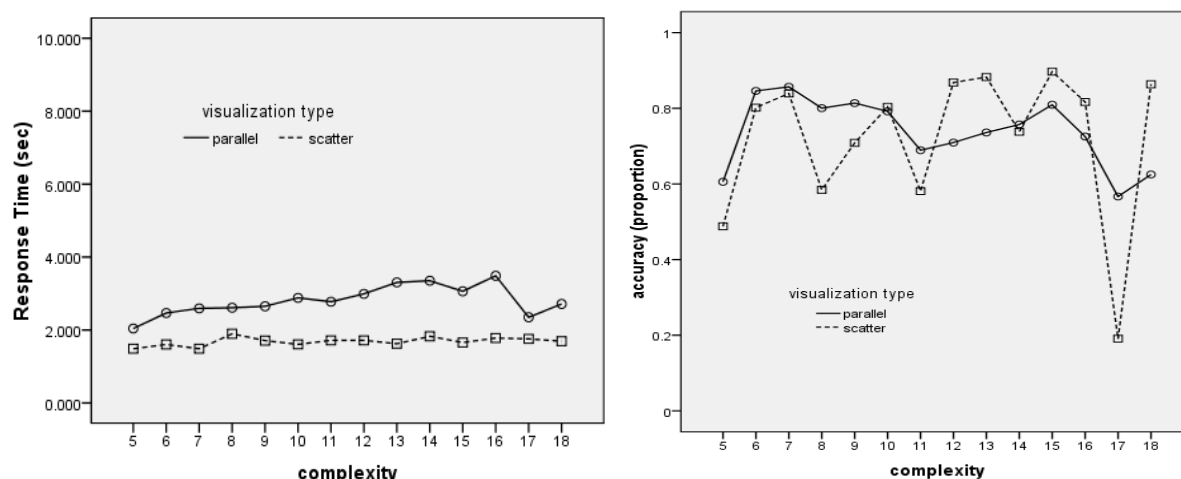


Figure 6: Interaction between visualization type and complexity for response time and accuracy.

2.2.2.3 Visualization type by attack pattern set

There was a significant two-way interaction between visualization type and attack pattern set for response time, $F(1,519) = 311.32, p < 0.05$, and accuracy $F(1,519) = 8.49, p < 0.05$. In terms of response time, the attack pattern set had a greater effect on parallel coordinates than on scatter plot performance. Specifically, it took longer for participants to detect complex attacks in simple noise (Condition B) than detecting simple attacks in complex noise (Condition A) when using the parallel coordinates but attack pattern set had no effect on response time performance when using the scatter plots. However, for the condition where a simple attack was embedded in more complex noise (Condition A), performance accuracy was higher with parallel coordinates, which was generally not true for Condition B (Figure 7).

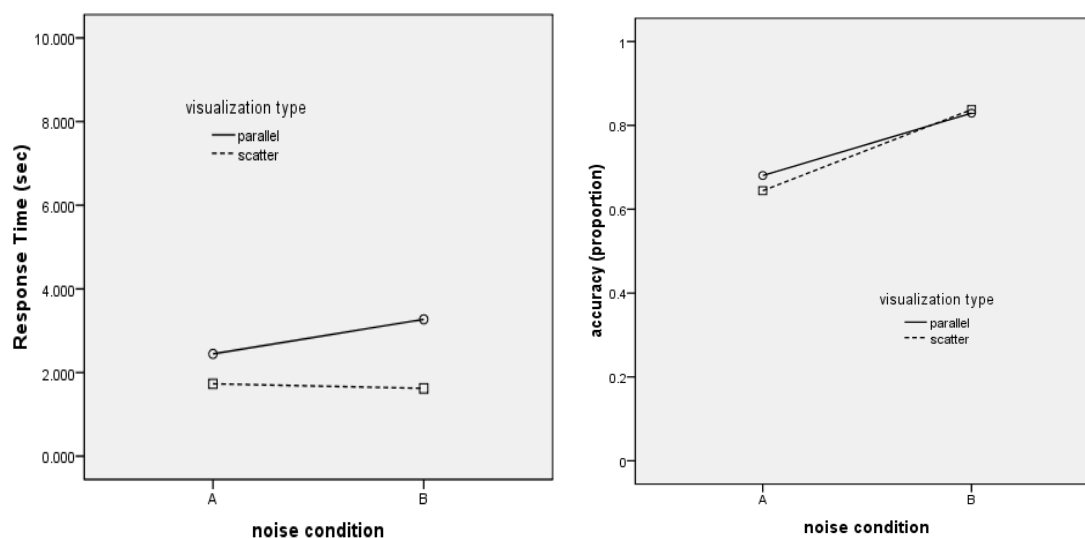


Figure 7: Interaction between visualization type with attack pattern set (noise condition) for response

time and accuracy. Condition A is simple attacks in complex noise and condition B is complex attacks in simple noise.

2.2.2.4 Visualization type by attack session

There was a significant interaction between visualization type and session in terms of both response time, $F(1,519) = 301.38, p < 0.05$, and accuracy, $F(1,519) = 27.68, p < 0.05$. Response time differences were less for Session 2 than for Session 1, in general, but the difference was chiefly from lower response times with parallel coordinates for Session 2; session did not make a difference in reaction times with scatter plots (Figure 8). In terms of accuracy, Session 1 participants were more accurate with parallel coordinates, but were slightly more accurate with scatter plots when comparing Session 2 accuracies.

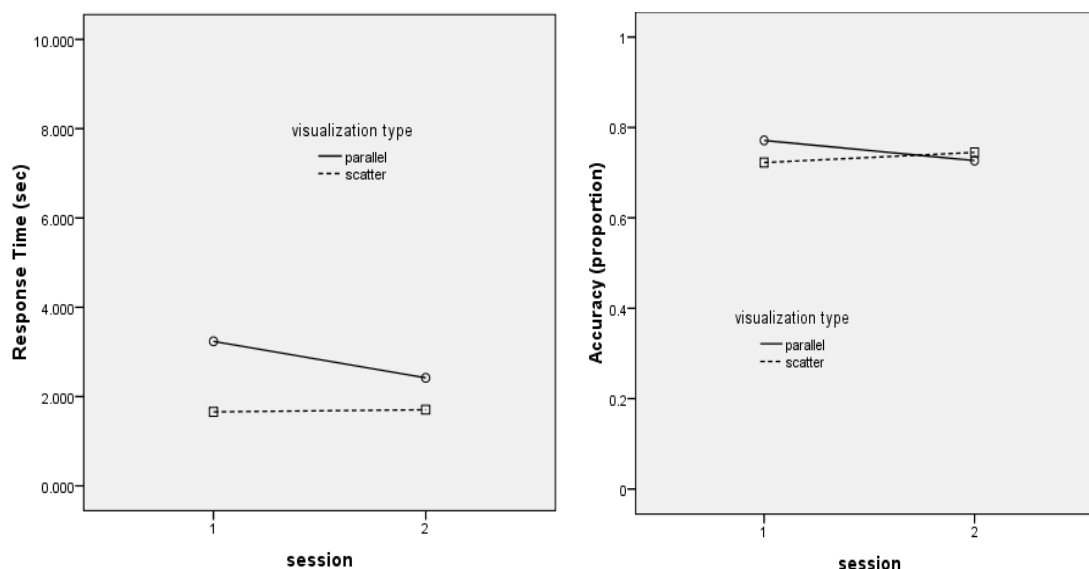


Figure 8: Interaction between visualization type and session (presentation order) for response time and accuracy.

2.2.2 Higher Order Interactions

2.2.2.1 Visualization type by complexity by attack pattern set

There were significant three-way interaction between visualization type, complexity, and attack pattern set for both response time, $F(11,519) = 12.34, p < 0.05$, and accuracy, $F(11,519) = 19.96, p < 0.05$. The effects of complexity on response times for parallel coordinates were less for simple attacks in complex noise than vice versa (Figure 9). What this effect generally means is that as more noise was added, complex attack patterns

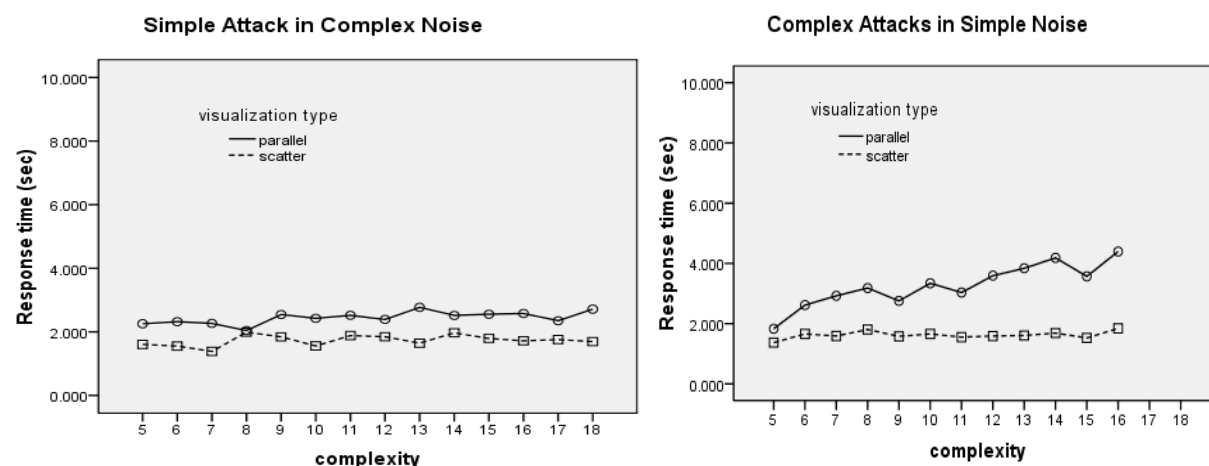


Figure 9: Differential interactions of visualization type and complexity for the two attack pattern sets with regard to response time.

were more difficult to find but simple attacks were not. In terms of accuracy, there was a high level of variability for the scatter plots when simple attacks were presented in complex noise versus when complex attacks were presented in simple noise (Figure 10).

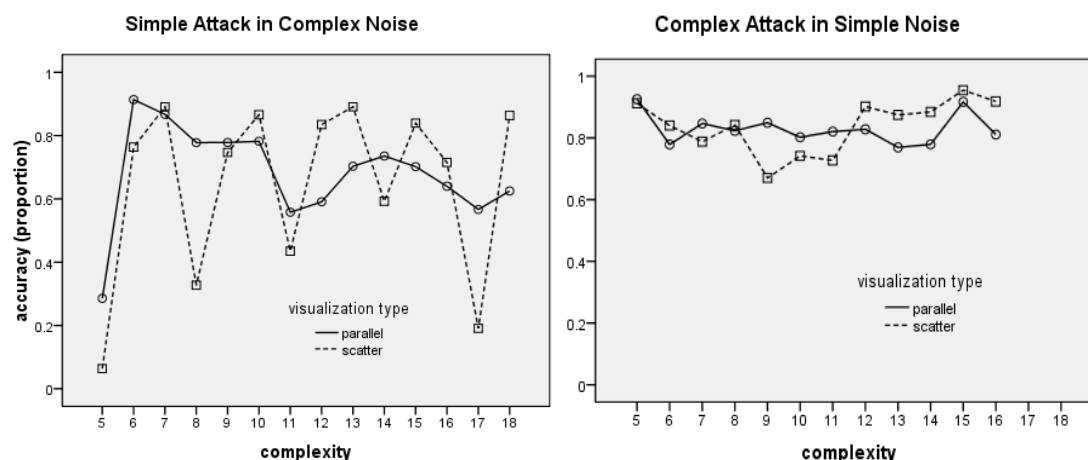


Figure 10: Differential interactions of visualization type and complexity for the two attack pattern sets with regard to accuracy.

2.2.2.2 Visualization type by attack pattern set by session

Analyses revealed a three-way interaction among visualization type, attack pattern set, and session on response time, $F(1,519) = 15.65, p < 0.05$. The differences in response time for the parallel coordinates visualization by session were smaller for simple attacks in complex noise than for complex attacks in simple noise, but did not vary for scatter plots (Figure 11).

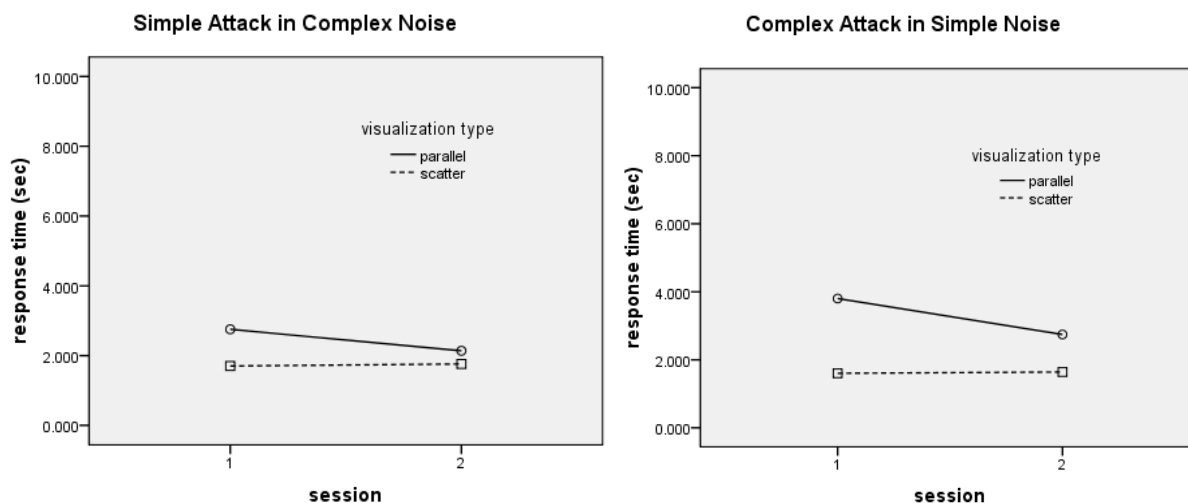


Figure 11: Differential interactions of visualization type and session for the two attack pattern sets with regard to response time.

2.3 Discussion

In terms of the main effects for visualization type, participants performed more quickly and accurately with the scatter plots than with parallel coordinates. However, these effects, although statistically significant, were fairly small from a practical standpoint (scatter plots were 0.6% more accurate and 1.14 seconds faster than parallel coordinates). Questionnaire data revealed that 28 of the 44 participants (64%) thought they performed better with the scatter plots; 13 of the 44 (29%) thought they performed better with the parallel coordinates; and the remaining 3 participants (7%) didn't think they performed differently regardless of the visualization type. Interestingly, during the course of the experiment, it became fairly clear that participants understood the one-to-many relationships better as they were depicted in parallel coordinates than in scatter plots. However, this artifact did not manifest itself in a significant performance advantage for the parallel coordinates.

In terms of the effects of distortion, the significant interaction between visualization type and distortion showed that scatter plots allowed for better performance in terms of response time for all levels of distortion and better performance in terms of accuracy for the higher levels of distortion, which is in line with our hypothesis that scatter plots would be less affected by distortion. Again, distortion manifested itself differently in the two visualization types (refer to Figure 4) resulting in a lesser effect on performance (faster response times and higher accuracies for higher levels of distortion) for the scatter plots (refer to Figure 5).

Since the variables involved in the remaining two-way interactions were also present in the three-way interactions, we will proceed directly to the discussion of the three-way interactions. Regarding the three-way interaction among visualization type, complexity, and attack pattern set, parallel coordinates response times were slightly more affected by complexity than the scatter plots; however this effect differed by attack pattern set. Oddly, for simple attacks in complex noise, response times did not vary much by complexity for either visualization type (refer to Figure 9). When one considers some of the simple attack patterns in the scatter plot condition and how the addition of multi-colored lines and squares to single-location attack patterns may lead to confusion, these results are surprising and contrary to our hypotheses. For instance [Refer to Table 3 for the

following example], one of the simple noise patterns is the 1,1,1,1 (single-colored marker in a single location for the scatter plot condition). To increase overall complexity of the stimuli, complex noise had to be added to this simple attack pattern. This addition could take the form of the m,m,m,1 (multi-colored markers in a square) noise pattern. It seems that it would be challenging to locate a single-colored single location marker in a field of multi-colored markers in a square, but perhaps the fact that the shapes of the markers were different for the two patterns, made that task reasonable. Finding a singular circular marker in a field of diamonds was not as challenging as it sounds. Also, consider adding the complex noise pattern m,m,m,1 (multi-colored markers in square shape) to the simple attack pattern m,1,1,1 (single-colored marker in a horizontal line). As in the previous example, the shapes of the markers for the two patterns would be different, but if the line was drawn first, then the square, there might be some overlapping of multi-colored markers from the square on the single-colored markers of the horizontal line, making the line look multi-colored. Again, the different marker shapes might have been a strong cue reducing the uncertainty of this situation. While this might explain the consistent time variable, this theory is challenged when looking at the accuracy data of the same 3-way interaction.

Figure 10 shows a high level of variability in accuracy with regard to the scatter plot data when looking at simple attacks in complex noise. Specifically, complexity levels 5, 8, 11, and 18 produce lower accuracies than other complexity levels. When looking at the specific stimuli that generate the levels of complexity stated above for the scatter plots, the attack patterns that participants were to identify in these cases were 1,1,1,1 and 1,1,m,1 (the single location marker, both single-colored or multi-colored). Note, the other two simple attack stimuli are m,1,m,1 and 1,m,1,1 (lines – which contained more visual content than the single location markers). Therefore, this effect may be more related to the specific patterns that participants were looking for versus the level of complexity variable we calculated for each. So to summarize the meaning of this interaction in terms of the scatter plots, participants responded fairly consistently regardless of attack pattern set, but their accuracy suffered when they were looking for 2 particular simple attack patterns.

With the parallel coordinates, response times for complex attacks in simple noise increased with complexity such that they were generally higher than with simple attacks in complex noise (refer to Figure 9). Accuracies, although not always greater, were more stable for parallel coordinates, particularly for simple attacks in complex noise (refer to Figure 10). The fact that, in the parallel coordinates format, the visual stimuli of the attack pattern with additional noise patterns that are required to increase the complexity score resulted in a total overlapping of the patterns on each other, made the task of finding complex patterns in simple noise more difficult (refer to Figure 4). For instance [Refer to Table 2 for the following example], think about the situation where the attack pattern is the m,m,m,1 (rectangle) combined with any of the simpler noise patterns, such as m, 1, 1, 1 (broom with handle to right). Participants many have been spending too much time trying to identify the individual patterns present in the stimuli before making a choice. For instance, in the above example, they may have spent more time trying to determine if the many lines on the first axis belonged to the many colored lines of the rectangle or the many lines of the broom before responding. The complete overlapping of these patterns may have caused participants a bigger challenge then deciphering the complexity of the scatter plots. Also consider a visual stimuli that might include both brooms (1,1,m,1 and m,1,1,1). Although the 1,1,m,1 pattern would be a single color and the other m,1,1,1 pattern would be multi-colored, at a quick glance, one might mistake the two brooms for the bowtie. The color coding for the parallel coordinates in this experiment was redundant with the first axis (time). It occurred to us during the course of the experiment that the patterns would be more separable if the color had instead been redundant with the fourth axis (rule category) thereby making each pattern one color similar to the ones in Table 2 that have a 1 in the first axis (1,1,1,1 [line]; 1,m,1,1 [diamond]; 1,1,m,1 [broom with handle to the left]; and 1,m,m,1 [house with roof to left]). So in summary, the three-way interaction between visualization type, complexity, and attack pattern set challenged our hypotheses and led us to some additional designs to test in future research.

In terms of the three-way interaction among visualization type, attack pattern set, and session, participants tended to perform better in the second session than in the first (refer to Figure 11). Although we did not predict an interaction, it is not particularly surprising that the effects of session were greater for parallel coordinates. It appears that participant tried to go faster in the second session when the second session was with parallel coordinates. Perhaps, this tendency is due to the fact that they had already experienced faster response times in the first session with scatter plots and this influenced the pace they set for themselves in the second session. Response times were generally closer and differed slightly less by session for simple attacks in complex noise than for complex attacks in simple noise. The fact that when parallel coordinates were presented in the first session resulted in longer response times is supported by the training data. Participants were required to go through the training session for each visualization type until they reached 100% accuracy on all 16 trials. Training data showed that the 44 participants ran through 112 training trails before reaching 100% accuracy for the parallel coordinates versus 88 training trials before reaching 100% accuracy for the scatter plots. Therefore, it took participants 2.5 practice runs on average versus 2 practice runs to reach 100% accuracy for the parallel coordinates as compared to the scatter plots.

3.0 CONCLUSIONS

This research shows that, based on knowledge elicitation activities with experts, it is possible to represent IDS log data in static visual representations which can portray visual signatures of attack patterns. The subsequent experiment showed that this is a reasonable approach to represent IDS attack data; that participants can quickly and accurately identify these visual signatures. While overall performance showed that scatter plots consistently provided the faster performance times, parallel coordinates showed more consistent accuracy performance. Based on the complex interactions revealed in this study, ideas for optimizing both visualization techniques should be pursued. Also, future research should investigate additional visualization techniques to determine how best to portray complex, abstract cyber information to operators in a meaningful way.

4.0 REFERENCES

- Aucsmith, D. W. (2011). Rethinking cyber defense. *High Frontier*, 7(3), 35-37.
- Barford, P., Dacier, M., Dietterich, T. G., Fredrikson, M., Griffin, J., Jajodia, S., Jha, S., Li, J., Liu, P., Ning, P., Ou, X., Song, D., Strater, L., Swarup, V., Tadda, G., Wang, C., & Yen, J. (2010). Cyber SA: Situational awareness for cyber defense. In S. Jajodia, P. Liu, V. Swarup, & C. Wang (Eds.) *Cyber Situational Awareness* (pp. 3-13). US: Springer.
- Bryant, W. D. (2013). Cyberspace Superiority A Conceptual Model. *Air & Space Power Journal*, 27(6), 25-44.
- Conti, G. (2007). *Security Data Visualizations: Graphical Techniques for Network Analysis*. San Francisco: No Starch Press.
- Costlow, T. (2013). Commercial tools playing expanded role in cyberops. *Defense Systems*, 7(6), 12-14.
- D'Amico, A., Tesone, D., Whitley, K., O'Brien, B., & Roth, E. (2008). *Understanding the Cyber Defender: A Cognitive Task Analysis of Information Assurance Analysts* (Report No. CSA-CTA-1). Northport, NY: Secure Decisions.
- D'Amico, A., & Whitley, K. (2007). The real work of computer network defense analysts. *Proceedings of the 4th International Workshop on Visualization for Cyber Security* (pp. 19-37).
- D'Amico, A., Whitley, K., Tesone, D., O'Brien, B., & Roth, E. (2005). Achieving cyber defense situational awareness: A cognitive task analysis of information assurance analysts. *Proceedings of the Human*

- Factors and Ergonomics Society 49th Annual Meeting* (pp. 229-233).
- Fua, Y. H., Ward, M. O., & Rundensteiner, E. A. (1999). Hierarchical parallel coordinates for exploration of large datasets. *Proceedings of the IEEE 1999 Conference on Visualization* (pp. 43-50).
- Goodall, J. R. (2005). User requirements and design of a visualization for intrusion detection analysis. *Proceedings of the 2005 IEEE Workshop on Information Assurance and Security* (pp. 394-401).
- Goodall, J. R. (2009). Visualization is better! A comparative evaluation. *Proceedings of the 6th International Workshop on Visualization for Cyber Security*. (pp. 57-68).
- Grinstein, G., Cook, K., Havig, P., Liggett, K., Nebesh, B., Whiting, M., Whitley, K., & Koneci, S. (2011). VAST 2011 challenge: Cyber security and epidemic. *IEEE VAST 2011*. (pp. 299-301)
- Harrison, L., & Lu, A. (2012, November). The future of security visualization: Lessons from network visualization. *IEEE Network*, 26(6), 6-11.
- Inselberg, A., & Dimsdale, B. (1991). Parallel coordinates. In A. Klinger (Ed.) *Human-Machine Interactive Systems* (pp. 199-233). US: Springer.
- Jackson, W. (2012). How to design a network that fights for itself. *Government Computer News*, 31(12), 18-21.
- Jackson, W. (2013). Surviving your next denial-of-service disaster. *Government Computer News*, 32(2), 24-27.
- Kilgore, R., Godwin, A., Hogan, C., Davis, A., Pfautz, J., Woods, D. D., Branlat, M., Kaufman, H. (2012). Tangible trustworthiness for mixed-initiative network defense (Final Report from SBIR Phase I). Cambridge, MA: Charles River Analytics.
- Koike, H. & Ohno, K. (2004). SnortView: Visualization system of snort logs. *Proceedings of the 2004 ACM Workshop on Visualization and Data Mining for Computer Security* (pp. 143-147).
- Langton, J. T., & Newey, B. (2010). Evaluation of current visualization tools for cyber security. *SPIE Defense, Security, and Sensing* (pp. 10-20).
- Leopold, G. (2013). Dempsey says cyberattacks are the new normal. *Defense Systems*, 7(5), 28.
- Livnat, Y., Agutter, J., Moon, S., Erbacher, R. F., & Foresti, S. (2005). A visualization paradigm for network intrusion detection. *Proceedings of the 2005 IEEE Workshop on Information Assurance and Security* (pp. 92-99).
- Mahoney, S., Roth, E., Steinke, K., Pfautz, J., Wu, C., & Farry, M. (2010). A cognitive task analysis for cyber situational awareness. *Proceedings of the Human Factors and Ergonomics Society 54th Annual Meeting* (pp. 279-283).
- Marty, R. (2009). *Applied Security Visualization*. Boston: Addison-Wesley.
- McCaney, K. (2014). 2015 budget proposal reflects the impact of cyber operations. *Defense Systems*, 8(2), 28.
- Ponemon Institute. (2013). *2013 Cost of Cyber Crime Study: United States*, Sponsored by HP Enterprise Security. Retrieved from http://media.scmagazine.com/documents/54/2013_us_ccc_report_final_6-1_13455.pdf
- Roth, E. M., Patterson, E. S., & Mumaw, R. J. (2002). Cognitive engineering. In J. J. Marciniak (Ed.) *Encyclopedia of Software Engineering* (2nd ed., Vol. 1, pp. 163-179). New York: Wiley.
- Rubin, J. (1994). *Handbook of Usability Testing: How to Plan, Design, and Conduct Effective Tests*. New York: Wiley.
- Shiravi, H., Shiravi, A., & Ghorbani, A. A. (2012). A survey of visualization systems for network security. *IEEE Transactions on Visualization and Computer Graphics*, 18(8), 1313-1329.
- Shneiderman, B. (1996). The eyes have it: A task by data type taxonomy for information visualization.

Proceedings of the IEEE Symposium on Visual Languages (pp. 336-343).

Thompson, R. S., Rantanen, E. M., Yurcik, W., & Bailey, B. P. (2007). Command line or pretty lines?: Comparing textual and visual interfaces for intrusion detection. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (pp. 1205).

Van Wijk, J. J., & Van Selow, E. R. (1999). Cluster and calendar based visualization of time series data. *IEEE Information Visualization* (pp. 4-9).

White hats to the rescue: Law-abiding hackers are helping businesses to fight off the bad guys. (2014). Economist. Retrieved from <http://www.economist.com/news/business/21596984-law-abiding-hackers-are-helping-businesses-fight-bad-guys-white-hats-rescue>





An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

Michael Boardman and Fiona Butcher

Dstl Porton Down
Building 5 - Room G02 - ISAT E
Porton Down
SP4 0JQ
UNITED KINGDOM

mjboardman@dstl.gov.uk

ABSTRACT

Future operating environments are likely to require the effective integration of humans and artificial intelligence (AI) enabled systems within decision-making processes. As these AI-based systems are adopted it will become increasingly important to ensure that appropriate human control is maintained for big data analytics and decision support systems. Meaningful human control (MHC) can be described as the ability to make timely, informed choices to influence AI-based systems that enable the best possible operational outcomes. There are a number of factors contributing to MHC including freedom of choice and sufficient human understanding of the situation and system. Therefore, to ensure appropriate application of MHC future agile command and control system designs will need to give due consideration to human judgement and decision making requirements. Over the last year the NATO HFM-ET-178 has explored the broad range of human factors issues associated with MHC and has agreed a subset of these to investigate in a follow-on NATO workshop planned for 2020. Although led by the HFM panel, future activities will be intensely cross-panel to produce comprehensive solutions. This paper discusses the importance of maintaining human control within military, AI enabled systems. It is based on work conducted over a 12 month period by members of the NATO HFM178 Exploratory Team with the intent of encouraging future linkages across NATO panels and activities.

1.0 INTRODUCTION

This paper discusses the issues involved in maintaining appropriate human control within military, Robotic, Intelligent and Autonomous (RIA) Systems and other technologies utilising Artificial Intelligence (AI). It is based on work conducted over a 12 month period by members of the NATO HFM178 Exploratory Team [1] with the intent of encouraging future linkages across NATO panels and activities.

The Exploratory Team included chairs: Mark Draper (USAF USA) and Jurriaan van Diggelen (TNO Netherlands); team members: Marcel Baltzer (Fraunhofer Germany), Michael Boardman (Dstl UK), Fiona Butcher (Dstl UK), Robert Shively (NASA USA) and Rogier Woltjer (FOI Sweden) and mentors: Frank Flemisch and Adelbert Bronkhorst.

The use of Robotic, Intelligent and Autonomous Systems within both military and civil applications is becoming increasingly common and it likely that this trend will accelerate as the technologies mature and become more affordable. These systems could be physically embodied as in the case of autonomous vehicles or workplace robots, or exist only as software agents that autonomously process data at speeds, levels of accuracy or for durations that humans are not capable of, provide support to decision makers, or even make decisions in their own right. While these technologies offer potentially significant benefits the use of advanced automation, autonomous systems, AI and Machine Learning raises many complex socio-technical, legal, moral and ethical questions. These concerns are often addressed through the inclusion of human

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

oversight and control of the system. Human control can be achieved through a variety of means ranging from supervising the system and taking over complete control when required through to exercising control as part of a Human Machine Team (HMT).

The demands of the future operating environment in terms of operational tempo and increasing information availability and a desire to reduce the risk to personnel is likely to lead to the more widespread adoption of sophisticated automation, autonomous, and AI based systems within defence applications. However, the physical, informational and political complexity of the future battlespace will require human judgement, adaptability, flexibility and active involvement in decisions that are outside the capabilities of the technology in order to deliver operationally effective, legal and resilient capabilities. Therefore the effective integration of humans and artificial intelligence (AI) enabled systems within data analysis, planning, decision-making and delivery of effects will become increasingly important.

The use of AI, machine learning and robotics is more advanced in many civilian systems than in comparable military applications; as such there are significant lessons that can be learnt from their development, use and evolving legislation. For example in the development of autonomous cars, “controllability” proved to be a critical consideration combining complex technical, ergonomic, legal, moral and organisational factors. Some of these lessons might well be transferred to the military domain; however, there are significant differences between military and civil applications and requirements. While the environment and scenarios of use may be relatively predictable in the civil domain, RIA systems operating in a military context must: withstand adverse and rapidly changing conditions, accommodate imperfect and potentially erroneous data sources, protect against deliberate adversarial actions, deal with ethical implications surrounding the use of force, and remain within bounds set by international humanitarian law (IHL) and rules of engagement. Moreover, military systems must be able to function safely and effectively under a wide range of highly dynamic environments and use cases that are hard to predict or anticipate during the design phase. They must also be resilient to failure and to complex, uncertain and unpredictable events and situations where the dynamics of the military domain necessitate complex judgements regarding acceptable actions based on rules of engagement, international law and judgements over legality, proportionality and risk. Because of this the maintenance of Human Control through a combination of specification, design, training, operating procedures, and assurance processes is seen as critical in many, if not all military systems.

This paper examines:

- The nature of human control within systems utilising AI or Machine Learning technology
- The factors that can contribute to or impede human control in various future military applications of AI-based systems.
- The complexity of ensuring that human control and accountability is delivered and maintained within future systems and Systems of Systems

The paper concludes by highlighting gaps identified for further research and is intended to initiate further discussions and debates in this critical area for defence.

2.0 BACKGROUND

The initial objective of the NATO Exploratory Team was to understand what was meant by *Meaningful Human Control (MHC)* within systems utilising AI or Machine Learning technology and develop a shared

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

working definition of what it is.

To guide the development of this understanding MHC was examined from multiple perspectives, these included:

- Authority, responsibility and accountability in the chain of command;
- The type of application in which AI technologies might be used;
- The type of interaction between humans and AI-based systems, such as: human oversight and intervention (e.g., ‘human-on/over-the-loop’), the use of AI-based systems as support tools, Human Machine Teaming (HMT).

Examining MHC from these perspectives highlighted to the NATO Exploratory Team the broad range of factors and complexities of different application areas that need to be taken into account when providing a definition. It also identified the complex legal and political environment associated with the use of AI and Autonomy within systems that are capable of delivering lethal effects.

The concept of MHC has recently gained attention as an important concept during the 2016 expert meetings organised by the UN Convention on “Certain Conventional Weapons” (CCW) and the use of Lethal Autonomous Weapon Systems (LAWS). While the concept of MHC has become linked to autonomous weapons it can be applied much more generally to a broader set of applications [12], including AI-based military applications for planning, decision support and Intelligence, Surveillance, Target, Acquisition & Reconnaissance (ISTAR). This is because these applications may be used within the processes that result in the decision to take actions that may result in the loss of life and/or have legal or moral considerations that require human involvement and therefore is essential maintaining human accountability. Human control can also be seen as a critical component of safeguarding the performance and resilience of such systems as the adaptability and flexibility provided by active human involvement can address situations outside of the capabilities of the technology alone.

Because of the role that MHC can play in delivering effective systems and also the wide range of system types that contribute to the intelligence collection, decision-making, planning and situational understanding that precedes the use of lethal and non-lethal effects it was decided that the ET should examine a wider range of systems and not just those delivering lethal effect.

3.0 THE NATURE OF MEANINGFUL HUMAN CONTROL

In order to develop a definition of MHC it was necessary to explore the components that comprise it. It was recognised that many of the issues being raised in relation to human control within Autonomous and AI enabled systems are not new, but share many parallels with those associated with automated systems. For example the following design principles for automated command and control systems were published by NASA in 1995 [2]:

- To command effectively, the human operator must be involved.
- To be involved, the human operator must be informed.
- The human operator must be able to monitor automated systems.

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

- The automated system must also be able to monitor the human operator.
- Automation systems must be predictable.
- Each of the elements of the system must have knowledge of the other's intent.

This identifies some key themes which feature in many recent papers and guidance documents relating to autonomy [3, 4, 5], such as: human awareness of system state and ability to predict system behaviour, the importance of the user maintaining situational awareness, the need for two-way understanding between human and machine, and shared understanding of each other's objectives. However, there are some concepts that don't feature in this NASA guidance such as *trust* and the concept of the machine acting as a member of a *team*¹ rather than as a *tool*. These two aspects are particularly important in understanding how human control is enabled and how it can be lost.

There are various automation and autonomy taxonomies [6, 7, 8], which provide various insights into the locus and nature of control within automation and autonomous systems, one such example are the levels of automation of decision and action selection [9]; see Table 1.

Table 1 - Levels of automation of Decision and Action Selection [9]

Low
1 The computer offers no assistance; human must take all decisions and actions.
2 The computer offers a complete set of decision/action alternatives from which the human selects and executes.
3 The computer narrows the selection down to a few from which the human selects and executes.
4 The computer suggests one alternative that the human executes
5 The computer suggests one alternative executes it if the human approves.
6 The computer suggests one alternative and allows the human a restricted veto time before automatic execution.
7 The computer executes automatically, then necessarily informs the human.
8 The computer executes automatically and informs the human only if asked.
9 The computer executes and informs the human only if it, the computer, decides to.
10 The computer decides everything, acts autonomously, ignores the human.
High

Whereas these Levels of Automation have proven useful as a broad characterization of options in many applications (such as autonomous vehicles), we consider them as too simplistic for the MHC debate.

¹ The notion of the machine as being referred to as a member of a team is not intended to anthropomorphise the technology, but convey the types of interaction and collaborative activity between human and machine.

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

Based on a review of the available literature and consideration of a range of potential applications of AI in military systems it was identified that human control over a system is *complex* and *rarely discrete* (simply present or absent), rather it is *dynamic* and *multi-dimensional* and the level of control required is highly situation dependent, and that it is not bound to one particular moment in time. A proposed set of dimensions or characteristics that contribute to human control was developed by the ET. These are described in Table 2.

Table 2 - Proposed dimensions/characteristics of human control

Dimension of Human Control	Example
The human has freedom of choice	This dimension captures the degree to which the human user can choose from all of the possible courses of action available. At one extreme the human has complete, unconstrained freedom to choose any course of action, including detrimental or illegal ones with no influence or assistance from the system. At the other extreme the machine has constrained the human to a single course of action which they have no freedom to deviate from. The constraining of human freedom of choice can be as a result of: • Direct effects – The system does not make a particular course of action or element of system functionality available to the user. • Indirect effects – Information is inaccessible or incomprehensible by the user preventing an informed, free choice. • Cultural effects - The organisational culture is such that selecting a course of action other than one recommended by the system carries potential blame or users feel that poor outcomes can be blamed on AI recommendations. • Workload effects – high workload prevents users from processing information to make a free choice and has to rely on the system to make the decision for them in totality or in part. It should be noted that not all decisions and actions require complete freedom of choice; in fact it may not be desirable in some situations due to the speed at which decisions must be made or where workload means that humans are unable to process all of the information available to them and consider all of the potential courses of action or options.

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

The human has ability to impact the behaviour of the system	This dimension captures the extent to which the user is provided with the functionality to change the behavior of the system. This could be in real time, or in advance through the setting of bounds or constraining allowable actions and behaviors. At one extreme the user has complete control over system behavior e.g. manual control over a UAV, ability to override an autonomous system. At the other extreme there is no functionality built into the system to allow the user to change the behavior of the system.
The human has time to decide to engage with the system and alter its behaviour.	This dimension captures the temporal aspect of user interactions with the system i.e. does the system allow the user sufficient time to process information, make decisions and impact on its behavior if required. At one extreme no time constraints are placed on the user to make decisions or act on the system. At the other extreme the system processes information and executes courses of actions at a speed beyond the capabilities of the user. This has been referred to in some quarters as machine speed warfare. There are situations where it is not feasible or desirable for the human to be “in-the-loop” e.g. defensive systems, so they have to be involved “before-the-loop” e.g. by setting constraints on action, to ensure that human control is maintained
The human has sufficient situation understanding	This dimension describes the extent to which the human has sufficiently accurate situational understanding to make an informed choice. At one extreme the user is presented with a completely accurate situational representation of the real world. At the other extreme the user is provided with no situational information at all.

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

The human has sufficient system understanding	This dimension captures the degree to which a human has a sufficient understanding of the system state, in order to understand the provenance, quality and accuracy of the information and the rationale of the decisions and recommendations made. At one extreme the user has a complete and accurate understanding of the system state, capabilities, the information that has informed the actions that the system takes and/or its behavior, and how the system uses information to determine its actions. At the other extreme the user has no understanding of the system state or how it is making decisions.
The human is capable to predict the behaviour of the system and the effects of the environment (physical and information)	This dimension captures the extent to which the user is able to predict how the system will behave in different circumstances. At one extreme the user has a complete and accurate understanding of how the system will behave in response to any given different inputs and/or conditions. At the other extreme the user has no understanding of how the system will behave in any situation.

3.1 Factors that could influence human control of Defence AI-based Systems

As a result of a review of the available literature and applying it to the defence context the ET determined that there were multiple dimensions and concepts wrapped up within the notion of *Meaningful Human Control*. As a result the ET felt that it was more helpful to consider Human Control of AI enabled systems more broadly under the notion of *Appropriate Human Control*. This recognises that the degree and nature of human control is dynamic, highly context dependent and made up of multiple factors which are more or less important depending on the situation and nature of the system and potential benefits and consequences provided by the system.

As a result the ET proposes two complimentary and interlinked components of *Human Control* within AI enabled and Autonomous systems; these are: *Meaningful Human Control* and *Effective Human Control* (see Figure 1).

Meaningful Human Control (MHC) encompasses the following elements:

- **Legal** – MHC is required to prevent violation of International Humanitarian Law and maintenance of accountability of military action (in both physical and information domains).
- **Moral / Ethical** – MHC provides the moral and ethical dimension to military decision making and military action (in both physical and information domains).

Effective Human Control (EHC) is required because humans and machines working towards a common goal can be better than either working in isolation. Effective human control within human machine teams can be an:

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

- **Enabler of Performance/effectiveness** – EHC (human involvement in decision making) is an enabler of improved operational outcomes
- **Enabler of Risk Reduction** – EHC (human involvement in decision making) reduces risk of undesirable outcomes and is a critical enabler of system resilience.

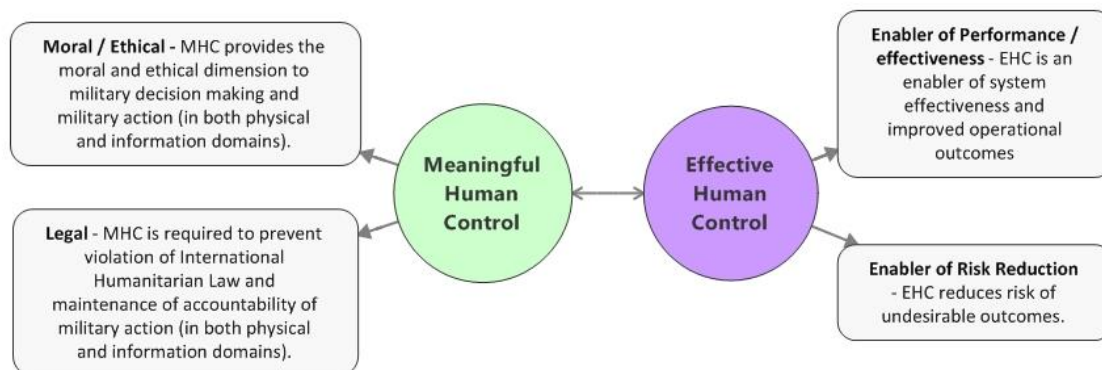


Figure 1 Meaningful and Effective Human Control

These two aspects need to be considered together as a focus on one at the expense of the other may have negative outcomes. For example a focus on effectiveness and optimising system performance could inadvertently impact on meaningful control and vice versa excessive involvement of users in some decision types could lead to delays, biases or errors that could impact on operational effectiveness.

3.2 A Description of Meaningful Human Control

As a result of the investigations conducted and in order to avoid potential political sensitivities associated with the evolving LAWS debate it was decided that a working description, rather than a formal definition would be more helpful in order to support further study.

Based on the factors identified the following description of HC was agreed upon by the members of the ET:

'Humans have the ability to make informed choices in sufficient time to influence AI-based systems in order to enable a desired effect or to prevent an undesired immediate or future effect on the environment.'

3.3 Achieving Meaningful and Effective Human Control

The ET also considered how systems might be designed to deliver an appropriate level of Human Control. This included consideration of both the processes and stakeholders involved in the fielding of systems that enable MHC and EHC so that suitable guidance and support can be developed.

In order to field systems that deliver Effective and Meaningful Human Control there are a wide range of activities that will need to contribute, including:

- National or Organisational Policy
- Systems specification
- Systems design

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

- Systems V&V
- Training of Users
- Training of AI and ML
- Systems of Systems (SoS) integration
- C2 process development
- Interoperability
- Operational Use
- After action review and lessons learned

The span of these activities will require the involvement of multiple stakeholder groups to ensure that appropriate human control is delivered by fielded systems. See Table 3.

Table 3 – Stakeholder groups that contribute to ensuring appropriate human control in fielded systems

Stakeholder Groups	Role in Delivering Appropriate Human Control in Future Systems
Policy Makers	• Set policy, standards and doctrine to facilitate EHC and MHC across NATO systems.
R&D / Scientific Community	• Identify and fill knowledge gaps in the field of EHC and MHC. • Conduct research to provide evidence to support the specification and design of systems that support MHC and EHC.
System Acquirers	• Specify, contract and accept AI enabled systems that support EHC and MHC.
System Designers	• Conduct and analysis, use Human Centred Design approaches such as those described in ISO 9241 210 [0], and apply best practice to systems design and testing to develop systems that support the delivery of EHC and MHC.
Organisational Users	• Integrate AI enabled systems within their wider system of systems and organisational structures such that they enable EHC and MHC.
End users	• Train, support and employ human machine teams that deliver EHC and MHC. • Develop operating procedures and practices to deliver EHC and MHC.

These activities and stakeholders conducting them will need to be supported by evidence based guidance, tools, standards and policies including:

- Enablers and Tools:

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

- Common language and terms.
- Models of MHC and EHC.
- Tools to assess and measure MHC and EHC within Systems / SoS and organisations.
- EHC risk analysis tools.
- Evidence base for MHC and EHC and associated research literature.
- HC Use Cases and examples.
- Guidelines and Standards:
 - Human Systems Integration (HSI) and acquisition approaches for MHC and EHC.
 - MHC and EHC Best Practice/Exemplars.
 - Human Machine Teaming Guidance.

3.4 Use of Human Control Dimensions to Support the Delivery of Appropriate Human Control within Fielded Systems

An initial step in this process was to identify whether the dimensions identified align to proposed approaches to assuring human control within Autonomous and AI enabled systems. A mapping of the ET dimensions of human control to the iPRAW requirements for human control [11] was conducted to determine the extent to which these aligned and how the two approaches might support the design of these types of system, see Table 4.

Table 4: Application proposed dimensions of ET developed Dimension of human control to iPRAW's Requirements for Human Control in the Use of Force [11].

	Situational Understanding	Intervention
	iPRAW Design of systems that allows human commanders the ability to monitor information about environment and system	iPRAW Design of systems with modes of operation that allow human intervention and require their input in specific steps of the targeting cycle based on their situational understanding
Control by Design (Technical Control)	Application of Dimensions of Human Control • The system design allows the Human to develop sufficiently accurate situation, and system awareness/understanding to identify risks to violating IHL and/or unacceptable moral, ethical or operational outcomes. • The system design allows the Human to predict the behavior of the system and its effects on the environment (physical and information).	Application of Dimensions of Human Control • The system design allows the Human to impact on the behaviour of the system in time to prevent an undesirable act (violating IHL and/or unacceptable moral, ethical or operational outcomes).

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

Control in Use (Operational Control)	iPRAW	iPRAW
	Appropriate monitoring of the system and the operational environment	Authority and accountability of human operators, teammates and commanders; abide by IHL
	Application of Dimensions of Human Control - The Human has sufficiently accurate situation, and system awareness/understanding to identify risks to violating IHL and/or unacceptable moral, ethical or operational outcomes. - The Human is able to predict the behaviour of the system and its effects on the environment (physical and information).	Application of Dimensions of Human Control - The Human is able to exercise freedom of choice and has the ability to affect system behaviour during use to ensure that accountability and adherence to IHL are maintained. - Training with systems allows users to understand and predict system behaviours across different situations in order to avoid undesirable outcomes or failures to comply with IHL. - Organisational culture does not indirectly impact on freedom of choice and willingness to question system behaviours and actions.

3.5 Complexity of Human Control

Considering the nature of human control within single user - single systems can be complex in its own right, but this complexity increases significantly when wider systems of systems aspects are considered. Interactions between multiple AI enabled systems, Humans, and Human Machine Teams are likely to lead to complex and unanticipated emergent behaviours and systems properties, see Figure 2. This complexity increases the potential for Human Control to be lost with associated detriments in system performance and with risks of accountability, legal and moral issues arising. There is also a risk that Human Control at an individual system level is lost at a system of systems level.

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

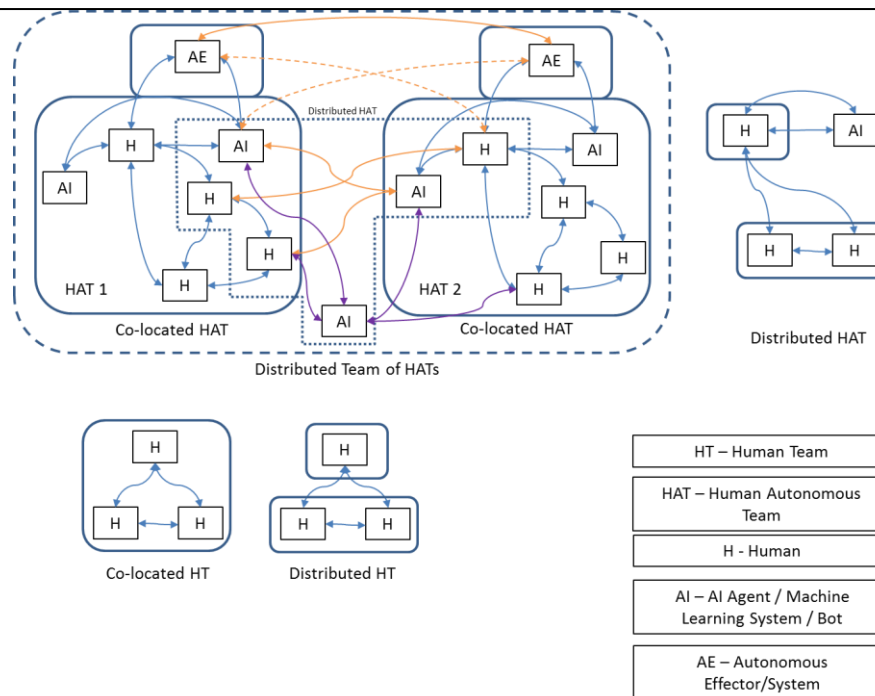


Figure 2 Complexity of multiple interacting elements in achieving, maintaining human control.

An added complexity is that systems may vary in functionality and potential impact on operational outcomes depending on system state, warfighting posture, rules of engagement, and wider the environmental context (e.g. where communications are constrained or denied). The breadth of potential system behaviours across all of these system behaviours these systems states needs to be understood and considered both within design and planning for deployment to ensure that Appropriate Human Control is maintained and human accountability is not lost within this complexity.

3.6 What are the Challenges and Complexities for NATO Regarding Human Control of AI Enabled Systems?

Bringing together the exploratory work conducted the ET identified a number of gaps in knowledge and complexities surrounding human control which require further investigation and research. Including the following:

- Further exploration and validation of the dimensions of human control.
- Methods of establishing the appropriate level of human control required by a system within a specific context and what factors affect this, such as as: the speed/tempo required to achieve effect; confidence (trust) in the output/behaviour of the AI Enabled System; potential consequence of action; the level of risk that is considered acceptable; potential operational benefits.
- How human control can be measured/evaluated.
- Whether appropriate human control can be guaranteed at design time, or is it something that must be continuously managed after deployment.
- Methods for delivering human control in complex systems of systems.
- Methods for delivering human control in multi-national or multi-agency organisations.
- Methods for tracing human accountability through complex systems of systems and identifying accountability gaps.

To address some of these gaps the ET recommended that a new RTG should be established together with a

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

dedicated workshop to further investigate the most pressing influencing factors regarding HC including:

- Organisational Considerations of HC (including team training, agile C2 structures).
- HF guidelines to achieve and maintain HC for all NATO AI applications.
- Systems Engineering methods (including TEV&V for learning systems) to support HC
- Adversary tactics to counter/undermine HC (methods and mitigations).
- MHC for complex socio-technical system of systems (emergent properties, HC propagation).
- Legal, ethical, political, and public perception of HC over AI-based systems.

The proposal for a dedicated workshop in this area was accepted and a workshop (HFM-322 Meaningful Human Control of AI-based Systems: Key Characteristics, Influencing Factors and Design Considerations) will be held in 2020. Although led by HFM, the workshop is intended to be cross-panel and inter-organisational in order to produce comprehensive solutions which draw on the wide range of relevant research being conducted across NATO. Because MHC spans many scientific topic areas across several disciplines but will focus on particularly on:

- Human-autonomy interaction
- Explainable AI
- Teaming research/ distributed intelligence/team structures and roles
- Shared dynamic mental models, maintenance of common ground
- Dynamic task allocation
- Directability and predictability with AI
- Observability/transparency and trust
- Organisational influences and processes
- Joint human-machine decision making/biases
- Joint human-machine learning/training
- Accountability and MHC
- Ethics and morality
- Complexity research

4.0 CONCLUSIONS

- Human control over AI enabled systems is complex and rarely discrete (simply present or absent), rather it is dynamic and multi-dimensional and the level of control required is highly situation and time dependent.
- While MHC has become linked to autonomous weapons it can be applied much more generally to a broader set of military applications where AI based systems are used, such as planning, decision support and ISTAR systems. This is because these applications may be used within the processes that result in the decision to use lethal effects and/or have legal or moral considerations that require human involvement and therefore is essential maintaining human accountability. Human control can also be seen as a critical component of safeguarding the performance and resilience of such systems as the adaptability and flexibility provided by active human involvement can address situations

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It

outside of the capabilities of the technology alone.

- The concept of Appropriate Human Control in AI enabled systems may be more useful to support AI enabled systems design than Meaningful Human Control. But determining the degree of Appropriate Human Control within a system requires consideration of both: Meaningful Human Control (associated with legal, moral and ethical aspects of human control) and Effective Human Control (associated with Human Machine Interaction and Human Machine Team performance and system effectiveness). These two aspects need to be considered in unison as a focus on one at the expense of the other may have negative outcomes.
- Achieving and Assuring human control and accountability in complex Systems of Systems and multi-national organisations will be challenging. Therefore, understanding and managing the risks to achieving appropriate human control can only be achieved through analysis of all aspects of the system design and interactions within it (human-human, human-machine, machine-machine and multiple teams of human machine teams).
- Multiple stakeholder groups and processes will be involved in ensuring that appropriate human control is delivered in fielded AI enabled systems.

5.0 RECOMMENDATIONS FOR NEXT STEPS

The ET recommended that a dedicated workshop should be held to further investigate the most pressing influencing factors regarding MHC. This proposal was accepted and a workshop will be held in 2020.

Given the breadth of AI related research being conducted that may impact on the achievement of MHC it is recommended that an integrating Team should be formed to serve as the centre of a dynamic hub-and-spoke structure between teams so as to enhance cross-team communications, leverage efforts, and maximise overall progress towards integrated MHC solutions across NATO.

6.0 REFERENCES

- [1] TECHNICAL ACTIVITY PROPOSAL HFM-ET-178 Meaningful Human Control Over AI-based Systems (2018).
- [2] Charles E. Billings (1991) Human-Centered Aviation Automation: Principles and Guidelines. NASA Technical Memorandum 110381.
- [3] NATO INDUSTRIAL ADVISORY GROUP NIAG SG.233 FINAL REPORT ON HUMAN MACHINE TEAMING (2019). NIAG-D(2019)0006.
- [4] Andrew P. Williams and Paul D. Scharre (eds) (2015) Autonomous Systems Issues for Defence Policymakers. ISBN 9789284501939.
- [5] Joint Concept Note 1/18: human machine teaming (2018) UK MOD Development, Concepts and Doctrine Centre.
- [6] Endsley, M.R. & Kaber, D.B. (1999). Level of automation effects on performance, situation awareness and workload in a dynamic control task. *Ergonomics*, 42, 462-492.
- [7] Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics – Part A: Systems and Humans*, 30, 286–297.
- [8] Save, L. and Feuerberg, B. (2013) Designing Human-Automation Interaction: a new level of Automation Taxonomy.
- [9] Sheridan, T.B., & Verplank, W. (1978). Human and Computer Control of Undersea Teleoperators. Cambridge, MA: Man-Machine Systems Laboratory, Department of Mechanical Engineering, MIT.
- [10] Ergonomics of human-system interaction — Part 210: Human-centred design for interactive systems (2019) ISO 9241-210:2019.
- [11] iPRAW (2017) “Focus On” Report No. 1 :Focus on Technology and Applications of Autonomous Weapons.
- [12] Santoni de Sio, F., & Van den Hoven, J. (2018). Meaningful human control over autonomous systems: a philosophical account. *Frontiers in Robotics and AI*, 5, 15.

An Exploration of Maintaining Human Control in AI Enabled Systems and the Challenges of Achieving It



Security perspectives on social media exploitation

Nora Bieteniece

NATO Strategic Communications Centre of Excellence, Riga
LATVIA
nora.bieteniece@stratcomcoe.org

Valentina Dragos

Chemin de la Hunière, 91123 Cedex Palaiseau
FRANCE
valentina.dragos@onera.fr

Bruce Forrester

Defence Research and Development Canada, Quebec City
CANADA
bruce.forrester@drdc-rddc.gc.ca

Tomas Krilavičius

Vytauto Didžiojo Universitetas Kaunas, Baltic Institute of Advanced Technology, Vilnius
LITHUANIA
tomas.krilavicius@bpti.eu

Albert Pritzgau

Fraunhofer Institute for Communication, Information Processing and Ergonomics FKIE, Wachtberg
GERMANY
albert.pritzgau@fkie.fraunhofer.de

ABSTRACT

Social media exploration for defence and security applications presents challenges that are beyond the ability of one domain or discipline to address. From online propaganda and disinformation to influence campaigns, social networks and platforms are often the vectors of intentionally distorted and biased narratives. Specific tools, like virtual bots, are used to manipulate public opinion and perceptions by amplifying or repressing political content, disinformation operations, hate and violent speech, and junk or fake news. Effective solutions for cyberspace exploration require synergies from organizational sociology, human computer interaction, communication, information science, and political science to interpret and analyze the evidence. The goal of this paper is to understand how social media streams and mechanisms can be better exploited from defence and security-driven perspectives. The paper discusses main challenges of social media exploration and relevant methods and techniques for use by military and civilian analysts in order to provide situational awareness and intelligence. Practical illustrations and use cases are also presented, explaining the possibilities and potential impact of the exploitation of social media channels for security purposes. The paper concludes by highlighting open questions and sketching future developments in the field.

1.0 INTRODUCTION

The Internet is currently a major source of information used by organizations and citizens to become informed on various subject and situations. The lack of social accountability in many digital platforms yields plenty of incentives for unprecedented forms of misuse. Disinformation, propaganda and fake news are just a few examples of ill-uses lurking in this largely accessible technology.

This paper aims to sample the state of the art in using social media analytics for security applications. First, we provide an overview of main challenges of social streams exploration and discuss methods developed to overcome those challenges. As the Internet is dynamic and heterogeneous, methods are based on hybrid Artificial Intelligence techniques. Secondly, we discuss several use cases of social media exploitation for defense and security applications. The main goal of the paper is to provide technical and practical insights in order to understand current capabilities for social streams exploration for security applications.

2.0 CHALLENGES OF SOCIAL MEDIA EXPLOITATION

Analysis of online streams comes with challenges stemming from both the specific nature of data created on digital platforms and the difficulties of social media exploration, as discussed hereafter.

Specificities of data: data collected from social media are vast, noisy, unstructured, inherently dynamic and heterogeneous in nature. Moreover, they convey reports on real-life facts and events augmented with personal points of view, such as evaluations, attitudes, and emotions. The data is, for the most part, inherently social in nature and reflects the inconsistencies and uniqueness found within humans. Our brains have evolved to be able to de-conflict and disambiguate the natural language that is found within social media posts. However, our ability to program algorithms to do the same is still evolving. Therefore, social data analysis is challenging for traditional data mining approaches that are often too slow and expensive, rely on sample sizes, and come with biases leading to errors.

Limitation of access to data and impact of secondary sources: An important volume of online data is released on DarkWeb [4], a problematic side of Web made of encrypted portions of the Internet that are not indexed by search engines and thus cannot be listed on results pages returned by search engines to user queries. Collection, processing and sharing of such content require specific procedures to be set up or the use of secondary sources. Nevertheless, secondary sources can introduce biases, or even truncate or alter the original set.

Influence of platforms and media-induced bias: Cyberspace is an artificial, manmade environment, with data and interactions framed in a particular manner. For this reason, social media platforms induce bias in how information is viewed by observers. Moreover, those observers can either be a part of the platform when undertaking their analysis or adopting a more direct approach to collect data via technical procedures. Those procedures are built on application programming interfaces (API), subroutines provided by platforms to access their collections of data, which are also stored in proprietary formats. This is a major drawback, as efforts are required to translate data into formats easier to process.

Security and ethical constraints, privacy protection: There are significant security related and ethical constraints to obtain first-hand information about sites, portals or content on social media platforms created by terrorists or extremist groups, from intelligences services, for example. Regarding privacy, data gathering and remote analysis for research purposes requires procedures and techniques that should be employed lawfully, as to make sure the overall process stay within the law.

For example, some countries within NATO do not allow their military intelligence analysis to collect and analyze data on their own citizens. This fact has allowed foreign agents to pose as citizens of these countries by using fake accounts. During the 2016 US presidential election, “Facebook believes 120 fake

Russian-backed pages created 80,000 posts that were received by 29 million Americans directly, but reached a much bigger audience by users sharing, liking and following the posts” [12]. This ploy is akin to an ‘insider threat’ and is very hard to detect.

In addition to those general challenges, social media exploration is also affected by technical bottlenecks, such as: multilingual issues, multimodality content, relevance and coherence of data sets, contextual information, aggregation and correlation of items, etc.. Next section discusses methods and techniques developed to address some of those challenges.

3.0 METHODS AND TOOLS FOR SOCIAL MEDIA EXPLORATION

Effective and efficient intelligence is usually accomplished in a combined human and computer effort. Indeed, the intelligence process heavily depends on combining the human’s flexibility, creativity, and cognitive abilities with the capacity and processing power of today’s computers.

Exploitation of social media generally breaks down into three fundamental objectives: information discovery, situational awareness and predictive analysis. Capabilities in addressing these objectives provide essential estimates of the potential risks faced by communities, economies and the environment. When exploiting social media sources, analysis is commonly limited to two aspects: users as basic units of the network and content as basic elements of communication [19]. These two aspects themselves are already invaluable sources of information. However, social networks additionally offer the context of communication and interaction represented by the network itself, namely, the network topology in the form of entities and relations. In addition to content, a given network structure promotes the derivation of activity and process patterns which can significantly improve situational awareness [20]. As a result, different data representations rendering distinct aspects of content and interactions serve as a means to adapt the focus of the intelligence analysis to specific information requests.

Methods and techniques for social media exploration can be roughly divided into lexicon-based [1] and machine-learning methods [3], and analyse the implicit, explicit and discursive patterns within data sets [2]. More specifically, content analysis and natural language processing (NLP) are two techniques of particular interest for social media exploration.

Content analysis is the quantitative analysis of properties inherent to different forms of communication. It is a scientific method used to systematically and reliably quantify the symbols used in communication. Content analysis has come to be seen as a particularly useful method for analyzing communication techniques associated with computerization and digitization. Indeed, user-generated contents have several interesting properties such as diversity, coverage and popularity that can be used as wisdom of crowds in search processes. Methodical approaches to content analysis, text analysis and text mining examine artefacts of social communication, typically written documents or transcripts, to develop objective inferences about a subject or topic of interest [21], the framing of conversations and many other characteristics of messages. Considering the huge volumes of social media data being created, it becomes clear that automated methods of text analysis make an invaluable contribution.

Promising ways to efficiently extracting intelligence information are NLP methods. These methods usually try to reframe complex language understanding tasks as simpler classification problems. Today’s models produce impressive results on some of the challenging tasks including Question Answering, Sentiment Analysis, and Text Entailment. Since information campaigns and propaganda also use psychological and rhetorical techniques to reach purposes, tools are developed to detect automatically these techniques by reframing them as classification tasks. Of particular interest to intelligence is to understand the collective behavior of certain groups of people, which is often supported by narratives. By automatically categorizing messages into different narratives about events and topics, we aim at raising the analyst’s awareness into the emergent views. The resulting analytical models can be used to reduce huge amounts of data to relevant data streams. The resulting subsets can then be subjected to further investigations such as the detection of propagandistic elements.

From practical perspectives, applications of social media analysis aim at: online hate [5] and violence detection [6], extremist content analysis [7], identification of online communities [10, 13], opinion leaders

Security perspectives on social media exploitation Security Perspectives on Social Media Exploitation

[7, 13] and contagion mechanisms [8], analysis of subjective content [11, 13] and investigation of online behaviour [9].

An approach called authentic chatter [13] was developed to help overcome the problems of large data sets and the need to have a good understanding of both influencers and narratives within topic areas. This Twitter specific method exploits social network analysis research and qualitative analysis. Relevant topic areas are identified by requirements. The prominent influencers (authors) within these areas are determined by using indegree and retweet metrics. This prominent group is then qualitatively analysed and assigned a type (official, news, SPAM, BOT, or Authentic). In this way the posts of between 50 – 100 authors can be analysed in detail in order to provide an effective proxy for the entire topic area. This method significantly reduces the time required for topic awareness while an in-depth understanding of the influencers and narratives is gained by the analyst. As such, appropriate responses and actions can be recommended.

The role of exploratory search for sensemaking in the intelligence process has already been described by Gersh et al. [14]. Exploratory search can be characterized by successive queries interspersed with stages of sensemaking [15]. According to Marchionini, general search activities can be grouped in “lookup”, “learn”, and “investigation” [16]. In particular, Wilson et al. [17] characterized investigation and learning as the two most important goals in exploratory information seeking, both of which require the involvement of the analyst in a range of activities such as comparison, aggregation, and evaluation. To this end, advanced information technologies and tools are developed to improve the efficiency and accuracy of intelligence analyst’s work. Usually, a query is the bridge between the analyst’s conception of the prevailing information need and the information access system [18]. To close simple information gaps, a query specification accomplished using keywords, key phrases, Boolean operators combined with command-based syntax may be sufficient. Complex processes of analytical reasoning, however, call for more elaborated mechanisms to specify queries tuned to a specific research question.

As highlighted by the selection of research efforts above, social media exploration is a challenging task requiring the development and training of domain specific analytical models, the incorporation of diverse knowledge resources- the analyst’s expertise, ontologies, knowledge graphs and more- and the implementation of hybrid methods relying on both data and knowledge driven techniques.

4.0 DEFENSE AND SECURITY USE CASES

This section illustrates several use cases conducted to explore social media for defence and security applications.

4.1 Propaganda detection

Case study: This case study aimed at detecting Russian influence. Two simple filters were developed in order to detect suspected Russian-based tweets. One filter looks at content and the other examines users. The first filter consists of a list of 200 websites taken from www.propornot.com. This site has, and continues to, identify sites that produce or propagate Russian propaganda. The second filter identifies authors who are associated with the Internet Research Agency (IRA) and who have been labelled as Russian Trolls by Twitter. This data set was released by Twitter in 2018.

Application context: These filters were recently applied to a Twitter data collection that had as an aim to detect foreign influence campaigns during the lead up to the 43rd Canadian election in October 2019. Past influence efforts, the 2016 US presidential election, usually involve trying to divide the opinions of the population. The campaign would focus on wedge issues, where there were opposing views. Efforts concentrated on moving one or both sides to a more radical view thus creating a seemingly irreconcilable divide between people.

To do this well-designed BOTs are employed. The BOTs are assigned a certain view – usually one that is in the interest of the foreign influencer. However, sometimes the aim is just to divide the people such that a minority (weaker) government is elected. Once the BOT focus is decided, one technique used is that the

BOTs are employed to support the chosen side using relevant hashtags and linking to local websites that push extreme views for a given topic. This technique makes it seem that there is a lot of support for the views within the extreme website at a local level thus encouraging others to adopt these views.

Methodology: For this case, we first identified potential contentious issues that were likely to be discussed during the election. A data collection, using appropriate hashtags and keywords, was started for each issue several months preceding the election. Care was taken to isolate the issues and ensure that mainly Canadian users were captured. Note that we were not looking at individual Canadians; the focus was on finding foreign BOTs within topic areas. For each issue we applied the two filters to determine if there was links with Russian propaganda or trolls. The first issue we examined was a smaller data set and neither filter produced results. However we were able to identify several BOTs within this topic. Next we examined a much larger topic and the Russian propaganda filter produced results. We examined the users identified by the filter and found a large proportion of BOTs. At this point we were able to use the BOT handles from these two topics as seeds to quickly examine the other wedge issues.

Analysis of results: We uncovered a large BOTNET that had many BOTs involved in several issues and at least two BOTs that posted in all of the wedge issues. Applying the IRA filter did not produce any results. We believe that, since the Russian Troll handles were compromised by their release by Twitter, these handles are no longer in use. A possible replacement for the user-based filter would be to use the BOT handles. Overall, this method proved very effective by producing positive results in less than 3 hours of effort.

4.2 Opinion detection in social streams

Case study: This study aimed at detecting opinions in social media and investigated the use of appraisal categories to explore social data [22]. The approach was grounded on cognitive foundations of the appraisal theory, developed by White and Martin [23]. The appraisal theory introduces three cognitive systems making the distinction between affect, appreciation or judgement, and highlights the mechanisms by which text convey positive and negative attitudes. The approach goes beyond the generally accepted definitions of sentiment and opinions, and focuses on appraisal expressions to describe the way humans express their attitudes, appreciations, and engagement.

Application context: The application context of this work is a cyberspace exploration task designed to support defence and homeland security intelligence practitioners in their efforts to gather valuable data allowing them to understand emergent phenomena, such as online hate proliferation or online propaganda. Linguistic clues of appraisal categories are used as indicators of subjective content to be collected and further analysed in order to understand the way online users express their extreme attitudes, embrace or support extreme ideologies and ideas.

Methodology: Starting with the three systems introduced by the appraisal theory, a semantic resource was build modelling finer categories under each system, and highlighting terms and expressions specific to each category. The methodology relies on a semantic annotation approach augmented with processing methods capable of performing a quantitative analysis of data gleaned on social media. The approach consists of several phases implemented to gather, process and analyse social data.

- Data acquisition: was done by crawling several sets of tweets with specific keywords and additional constraints to select only posts written in English, see tab. 1. This phase was carried out as a straightforward step and relies only on data content of both keywords and posts and the ability of API used to mine the social network.
- Data processing: performs first a cleaning step, by removing URLs, hashtags and any information considered as irrelevant. Processing at paragraph level includes sentences identification according to punctuation marks, tokenizing, part-of-speech (POS) tagging along with identification of words stems.
- Semantic annotation: attaches additional information to various text paragraphs based on their content analysis.
- Data analysis: was carried out based on semantic annotations and estimates for each set of data; the percentage of subjective versus objective tweets; the percentage of tweets having positive

Security perspectives on social media exploitation Security Perspectives on Social Media Exploitation

versus negative orientations; the distribution of tweets according to their low, medium or high strength and also their distributions with respect to *Attitude* and *Engagement* systems, and also with a finer distribution according to their specific concepts.

Table 1: Data collections for opinion analysis

Key word	Remarks	Number of tweets
Western values	Often associated with online hate	7956
White supremacy	Extremist content	7542
Iraq war	Old conflict	7001
Western coalition	Often associated with online hate	6971
Security	Generic term	7509

Results: The analysis of data sets was carried out in order to calculate numerical distribution of subjective and objective tweets, positive and negative orientation, high and low force along with a finer analysis in the light of appraisal categories. Numerical values show a high percentage of subjective tweets for all datasets analysed, regardless of the keyword. Tweets conveying *Attitude* concepts are generally less than 50%. There is a good representation of positive-oriented tweets that account for more than 50% of the collection, while tweets having low or high impact are underrepresented for all data sets. The method also highlights associations of key words and concepts: for example *White supremacy* has strong correlations with *valuation* and *affirm*, and low connections with *impact* and *concede*. At collection level, the analysis shows a rather similar distribution of data sets into categories of appraisal systems, with around 10 classes for *Attitude* and 7 or 8 classes for *Engagement*.

4.3 Dynamics of Lithuanian radical groups in Facebook

Data collection: Data was collected from Facebook using FG graph API [1] (currently not available) on February 12, 2018. Posts from period March 4, 2010 – January 1, 2018 were downloaded. Posts, post creation data and posts ids were collected and stored. The groups were identified using Facebook search engine, mostly based on posts proclaiming nationalism, strong nation, xenophobic ideas as well as hostility to ethnical and sexual minorities, i.e. the following keywords were used *Lithuania, Lithuanians, land sale, European Union, NATO, refugees, refugee crisis, Muslims, Jewish restitution, Jew, Russian, Roma labor, gay pride, gay mountaineering*.

After preliminary analysis, 10 groups were chosen for the in-depth analysis, based on the following criteria: (1) presence of the radical ideology, (2) more than 100 members and (3) the last post was published at least 2 days prior to the analysis (i.e., group is still active).

Two datasets were collected, see tab.2:

- Pro-Russian FB groups: 70 150 posts, 13940 members.
- Other radical right groups: 9578 posts, 6126 members.
- Total: 79728 posts.

Analysis methods: Most frequent words were used as features, and posts timestamps were used to identify change of activity during the all period. Texts were not lemmatized.

Table 2: Data collections for dynamics analysis

	Year							
	2010	2011	2012	2013	2014	2015	2016	2017
Pro-Russian	0	0	0	0	10447	17033	19257	23413
Radical right	58	311	604	1164	2997	791	1914	1739
Total	58	311	604	1164	13444	17824	21171	25152

Analysis of results: Activity (see fig. 1) of two types of groups was quite different:

- Pro-Russian groups were active in FB from 2014, and the peak was 2017.
- Radical right groups were active from 2010, activity growth was very fast, and reached peak in 2014.



Figure 1: Activity of FB groups

Moreover, appearance of different keywords in posts was analysed, e.g. Lithuania was mostly mentioned in 2014 by pro-Russian groups, and 2017-2018 by other groups, but it was quite important all the time, while Russia is mostly discussed by pro-Russian groups (with peak in 2015) and only scarcely by other groups, mostly in relation to sanctions and Ukrainian crisis.

Basically, the results show that FB groups' activity is strongly influenced by different events, involving Lithuania and Russia, and most of radical groups are moving to Facebook from forums and other media channels. Monitoring of such groups could be a good indicator for different groups' activities, dynamics of such groups, and related risks.

4.3 Target identification

Target identification, i.e., identification of domains or information actors in the online environment, is an important step in social media exploration for defense and security. It narrows in on outlets or online actors that exhibit behaviour of interest or belong to a real-life group of interest. It also reduces the amount of information to be processed when exploring social media.

Open source intelligence (OSINT)¹ methods and tools are often used for target identification. This includes manual monitoring and processing of social media streams, user data and metadata or the use of automated OSINT tools² for monitoring and processing such information. Both approaches can be generally divided into text-based methods/tools, geographic information and remote-sensing

¹The process of visual and textual monitoring of and deriving information from publicly available data in the intelligence context is a common.

² Maltego, TinEye, IntelTechniques.

Security perspectives on social media exploitation Security Perspectives on Social Media Exploitation

methods/tools, network science methods/tools and visual forensics³ none of which are discussed in this paragraph.

Application context and methods: Method frequently used in conjunction with OSINT tools and techniques is crowdsourcing, the practice of obtaining information or input by enlisting the services of a large number of people via the Internet. Much of Bellingcats'⁴ work is an example of crowdsourced target identification. In 2014, Bellingcats' crowdsourced analysts used open-source tools to discover which Russian unit shot down the MH17 flight in Ukraine⁵. For this analysis Bellingcat volunteers used images and social media posts taken and posted by the local public prior and/or after the accident. They also used satellite imagery to verify the locations identified from social media posts and images, and searched and found their suspect actors online. Their analysis was sufficient to be used as evidence by Dutch Safety Board⁶.

A less commonly known target audience identification method in conventional defense and security is honey-potting. In cybersecurity honeypot is a closely monitored machine serving several purposes: it can distract adversaries from more valuable machines on a network, provide early warning about new attack and exploitation trends, or allow in-depth examination of adversaries during and after exploitation of a honeypot. In the application of social media exploration for defense and security, honey-potting is a space designed to "lure" possible suspects or targets. In NATO StratCom COE recent study⁷, Facebook page run by the red-team was launched and advertised to possible military exercise participants via Facebook Ads. After gathering a significant following, the red-team created a closed group for exercise participants and advertised it via the said page. This group served as a "honey-pot" for identifying exercise participants; moreover, it provided a space to interact with the participants without being challenged by the general public. This method of target identification is more "invasive" in comparison with OSINT techniques that merely gather, correlate, verify and create new evidence from existing data points such as social media posts, images etc. However, it serves potentially as a source for new information that suspects or targets otherwise would not have disclosed on their social media feed.

The variety of use cases discussed above clearly illustrates the potential of social media to facilitate the creating and sharing of information useful for defence and security applications. However, those use cases also emphasize the crucial importance of social, linguistic and cultural indicators, the difficulty of developing efficient approaches for social media exploitation and the exploratory nature of the task.

5.0 CONCLUSION AND PERSPECTIVES

This paper provides an overview of challenges and methods for social stream analysis. This an important and timely endeavour because social media does not only play a significant role in shaping people's attitudes and beliefs, but it also has far-reaching consequences for societies in general, such as increasing tendency to violating social norms and disrupting democratic processes. The paper analyses several approaches used for social media analysis. The first was based on filtering for content or users that are known to be used by Russia. The second and third case studies use linguistic cues that are based on sociology principles and aim at understanding human behaviour online and the propagation of stories of interest to various target audiences. The last use case discusses the identification of targets on social networks and the analysis of interactions between agents in the real and virtual environments for intelligence tasks.

As a dynamic and heterogeneous environment, social media exploration deserves further study and leads to interesting new perspectives when combined with relevant research in related areas such as information

³<https://edam.org.tr/wp-content/uploads/2018/07/Digital-Open-Source-Intelligence-Bosch-2nd-Report.pdf>

⁴<https://www.bellingcat.com/>

⁵Ibid.

⁶<https://www.bellingcat.com/news/uk-and-europe/2015/10/15/how-the-dutch-safety-board-proved-russia-faked-mh17-evidence/>

⁷<https://www.stratcomcoe.org/current-digital-arena-and-its-risks-serving-military-personnel>

retrieval, event detection or trust analysis. Certainly the major improvements in hardware power and categorization algorithms have led to significant advancements in the research. However, as researchers and analysts become better at detection of influence and deception, those wishing to influence and deceive are continuously creating new techniques making it a ‘cat and mouse’ game. Challenges still exist. Natural language processing, sentiment analysis and content analysis continue to be problematic when applied to the messy nature of social media data. Further, the always increasing and never ending stream of data taxes our analytical resources. We must continue our efforts to increase the signal to noise ratio allowing analysts to concentrate on pertinent signals. Finally, our analytical tools need to be precise thus permitting for reliable, valid, and accurate assessments that can translate into actionable intelligence for decision makers.

6.0 REFERENCES

- [1] M. Taboada, J., Brooke, M., Tofiloski, K., Voll, and M., Stede. Lexicon-based methods for sentiment analysis. *Computational linguistics*, 37(2), 267-307., 2011.
- [2] F., Villarroel Ordenes, S., Ludwig, K, De Ruyter, D., Grewal, & M., Wetzels, (2017). Unveiling what is written in the stars: Analyzing explicit, implicit, and discourse patterns of sentiment in social media. *Journal of Consumer Research*, 43(6), 875-894.
- [3] E., Cambria, C., Havasi, and A. Hussain, SenticNet 2: A Semantic and Affective Resource for Opinion Mining and Sentiment Analysis. In *FLAIRS*, 2012.
- [4] Weimann, G. (2016). Going dark: Terrorism on the dark web. *Studies in Conflict & Terrorism*, 39(3), 195-206.
- [5] Isbister, T., Sahlgren, M., Kaati, L., Obaidi, M., & Akrami, N. (2018). Monitoring Targeted Hate in Online Environments. *arXiv preprint arXiv:1803.04757*.
- [6] Alvari, H., Sarkar, S., & Shakarian, P. (2019). Detection of Violent Extremists in Social Media. *arXiv preprint arXiv:1902.01577*.
- [7] Wei, Y., & Singh, L. (2017, May). Using network flows to identify users sharing extremist content on social media. In *Pacific-Asia Conference on Knowledge Discovery and Data Mining* (pp. 330-342). Springer, Cham
- [8] Ferrara, E. (2017). Contagion dynamics of extremist propaganda in social networks. *Information Sciences*, 418, 1-12
- [9] Lara-Cabrera, R., Gonzalez-Pardo, A., Barhamgi, M., & Camacho, D. (2017, August). Extracting radicalisation behavioural patterns from social network data. In *Database and Expert Systems Applications (DEXA), 2017 28th International Workshop on* (pp. 6-10). IEEE.
- [10] Rossetti, G., Pappalardo, L., Pedreschi, D., & Giannotti, F. (2017). Tiles: an online algorithm for community discovery in dynamic social networks. *Machine Learning*, 106(8), 1213-1241.
- [11] Ruhrberg, S .D.,Kirstein, G., Habermann, T., Nikolic, J. and Stock, W.G.(2018) #ISIS—A Comparative Analysis of Country-Specific Sentiment on Twitter. *Open Journal of Social Sciences*, 6, 142-158.
- [12] Solon, O., & Siddiqui, S. (2017). Russia-backed Facebook posts 'reached 126m Americans' during US election. *The Guardian*. 31 October, 2017. Accessed 9 September 2019.
- [13] Forrester, B., (In Press) Authentic Chatter, *Computational and Mathematical Organization Theory Journal*.
- [14] Gersh, J., Lewis, B., Montemayor, J., Piatko, C., & Turner, R. (2006). Supporting insight-based information exploration in intelligence analysis. *Communications of the ACM*, 49(4), 63-68.
- [15] P. Pirolli and S. Card, “The sensemaking process and leverage points for analyst technology as identified through cognitive task analysis,” *Proc. Int. Conf. Intell. Anal.*, vol. 2005, pp. 2–4, 2005.
- [16] G. Marchionini, “Exploratory search: from finding to understanding,” *Commun. ACM*, 2006.
- [17] L. Wilson, “From Keyword Search to Exploration: Designing Future Search Interfaces for the Web,” *Found. Trends® Web Sci.*, vol. 2, no. 1, pp. 1–97, Jan. 2010.
- [18] M. Hearst, “Search user interfaces,” 2009.

Security perspectives on social media exploitation Security Perspectives on Social Media Exploitation

- [19] H. Kwak, C. Lee, H. Park, and S. Moon, "What is Twitter, a social network or a news media?," in *Proceedings of the 19th international conference on World wide web - WWW '10*, 2010, p. 591.
- [20] D. Helbing *et al.*, "How to Save Human Lives with Complexity Science," p. 67, Feb. 2014.
- [21] K. H. Krippendorff, *Content Analysis: An Introduction to Its Methodology*, 3rd ed. 2012.
- [22] Rasa Kasperienė, Tomas Krilavičius. Application of Social Network & Content Analysis Methods for Assessing the Dynamics of Facebook Groups. The 24th International Conference on Information Technology (IVUS 2019). Kaunas, Lithuania, Apr 25, 2019.
- [23] Graph API in Facebook for Developers. <https://developers.facebook.com/docs/graph-api>, Accessed 19 Apr 2018

Propaganda Filters: Tracking Malign Foreign Interventions on Social Media

Bruce Forrester

Defence Research and Development Canada, Quebec
City Canada bruce.forrester@drdc-rddc.gc.ca

Akiva Bacovcin

Nexalogy, Montréal, Canada
abacovcin@nexalogy.com

Zachary Devereaux

Nexalogy, Montréal, Canada
zdevereaux@nexalogy.com

Stefany Bedoya

Nexalogy, Montréal, Canada
sbedoya@nexalogy.com

ABSTRACT

Democracy requires that the people have access to factual information and that they be able to participate in open and honest discussion on issues of importance. Unfortunately, there are forces in this world that would benefit from weak democracies. Many such forces are currently very active on social media spreading propaganda, disinformation, and trying to increase divisions within democratic populations. Identifying foreign-backed users and BOTs that are being exploited to interfere in democratic processes is an important first step to understanding effects. This research aimed to describe a methodology that can be used to identify suspicious users from which indicators of foreign narratives can be inferred. To help describe the network effects of BOTS, the BEND (build, engage, neutralize, distract) framework is used. To make the research both timely and real, data collection and analysis was conducted on divisive issues found within Canada. We identified evidence of BOTs, with attribution to Russia, meddling in Twitter discussions on these wedge issues. This meddling was mapped onto the BEND framework. In particular, this meddling seemed to be focused in Boosting and Enhancing extreme viewpoints.

1.0 INTRODUCTION

Evidence of foreign interference in numerous democratic processes, perpetrated using social media, is no longer surprising [1-4]. In fact, Bradshaw and Howard [5] have found that in at least 70 countries there is evidence of social media manipulation to shape public attitudes. NATO is also well aware of the 2013 “Gerasimov Doctrine”, Russian Chief of the General Staff that promotes indirect and asymmetric methods [6] that uses the information environment to attack human cognition centering on distraction and manipulation [7]. As an example of manipulation, according to a NATO study [1], Russia had four strategic (read political) goals for their propaganda effort against the Ukraine:

1. To promote Russia as a crucial player in the polycentric world in the process of international peace;
2. To claim Russia’s superiority over the US;
3. To prevent Ukraine’s transformation into being part of the external border of NATO and the European union; and,
4. To soften and, in the nearest future, achieve the lifting of the sanctions regime against Russia.

Within Canada, the focus is more about distracting the population. Upsetting well-established democracies by increasing the divisions between citizens’ with opposing views is an effective method; while the people of that country are busy “fighting” each other, Russia is able to move with greater freedom with less scrutiny. This research was conducted with regard to the several of the objectives of the IST-177 research task group. In particular:

1. Exploration of attack vectors and effects of social media influence attacks;
2. Fake News and Propaganda Recognition;
3. Understand tools for injection of military-relevant information into social media streams;
4. Exploration of relevant technologies and tools; and,
5. Analytics for segmenting target populations for maximum influence opportunity.

The goal of this research was to try to detect Russian influence within divisive Canadian issues. Additionally, we hoped to decrease analyst time to insight for these issues by using AI to quickly treat large amounts of data e.g. big data and automation.

In order to find evidence of Russian interference on social media, analysts’ must understand the likely target areas within the information environment. They must collect, sort, and analyse hundreds of thousands of social media posts. A daunting challenge within the world of data analysis, for situational awareness, public affairs, and numerous other purposes, is the amount of time it takes to analyse enormous groups of social media posts. Indeed, in many cases, there is not enough manpower available to make manual inspection possible in a useful timeframe. This research aimed to help address this challenge by using filters and AI based treatment of big data via neural networks which have been developed to significantly decrease the time to insight with large data sets of social media.

The novel methodology derived here is based on knowledge of Russian (Maskirovka) techniques. Their information warfare tactics focus on promoting division in foreign populations by finding wedge issues to push people to the extremities. In studying the behaviour of Russian BOTs on social media, McKelvey and Dubois found “amplifiers, dampeners, transparency, and servant BOTs” [8, 9].

In order to better understand the motivation and tactics of the identified BOTs, the BEND framework [10] was used. This framework is based on research on disinformation and Russian information manoeuvres and describes strategies that can be used by actors in the information environment. Information activities are characterised along two axis: community manipulation versus content manipulation and positive manipulation versus negative manipulation. Each letter of BEND corresponds to one quadrant produced by

the combination of the two axis Figure 1. For example, the positive community manoeuvre ‘Build’ helps to create a group of users that supports a brand, topic, or theme. A negative content manoeuvre such as ‘Distort’ manipulates content through omitting facts, disinformation, or lies, in order to alter an audience’s perception. For a full discussion of this framework and description of the 16 manoeuvres refer to [10].

The BEND Framework

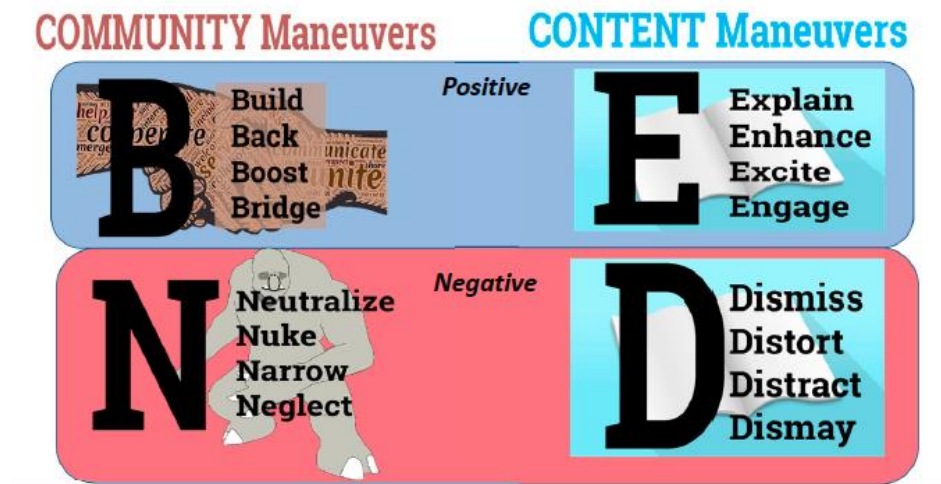


Figure 1. The BEND framework

2.0 METHODOLOGY

Keeping in mind the research objectives identified in the introduction, a repeatable and simple method was devised that allowed for iteration over the data and spiral analysis of interim results between steps. We were interested in determining if there was evidence of Russian interference in Canadian based Twitter discussions that dealt with divisive issues. We assumed that Russian information warfare tactics would be centred on the further trying to divide (encouraging more extreme views and opinions) the Canadian population by using community and content manoeuvres as identified by the BEND model. As such, the first step was to determine topics that were potentially divisive. These issues were chosen based on their importance to many Canadians and because there exists strong opposing opinions within the issues, often based on ones’ personal beliefs.

A recursive and iterative, discovery method was used for this research, which was applied to several different, but related, datasets. The overall method used for this research was as follows:

1. Identify possible wedge issues for foreign interference and related search terms;
2. Perform initial discovery & filtering;
3. Apply predetermined filters and Authentic Chatter map to detect suspicious users;
4. Analyse initial results;
5. Identify and classify users;
6. Update filters to include suspicious users (likely to be BOTs);
7. Apply updated filters to detect associated users and related narratives;
8. Identify network effects according to BEND; and,
9. Apply updated filters to examine inter-issue narratives.

Starting in June 2019, we identified a number of potential wedge issues based on current media political discussion and our own knowledge of the Canadian political scene. Those issues were: immigration, ethical issues, bill21, Canada-China relations, Canada-USA relations, climate change, pharmacare, indigenous issues, and pipelines. A control group using an overarching hashtag was also collected to permit discovery of unidentified issues.

Initial data collect based on the Twitter API (about 1500 tweets) was performed and the results were examined. Boolean search queries were created to capture the relevant terms, checking the initial version for accuracy, and then performing necessary edits until the data capture was accurate and comprehensive. Top Hashtags and Top Words, for example, contributed to search strategy refinement and data veracity was achieved by iterating through this process. Contradicting hashtags, on either side of an issue, was evidence of a wedge issue.

Data collection occurred from 11 June to 10 Aug 2019. For this short paper, the research investigate only three topics: ethical issues, pipelines, and climate change. It is always easier to capture a wider swath around a topic and then filter and refine the results. Stop words included “Trump” and geolocations derived from Twitter based on latitude and longitude or on proprietary location derivatives (likely biographical, discussion topics) were used to increase the probability that tweets were coming from Canada. The goal was to create Canadian specific data collections without, as much as possible, noise from the USA.

The rest of this section will be a walkthrough of the methodology. The method was very iterative, often examining the summary ‘reports’ provided by the NexaIntelligence software. The software used allowed for quick examination of the overall content through these numerous reports that were automatically produced. Ranks and clustering of the data facilitated the understanding of overall themes and groupings. Several reports (automatically produced by the NexaIntelligence software) are presented in the results section.

Starting with a chosen data set (ethical issues which was just under 30k tweets) tweets were examined to see if they contained issues that would be a target of Russian interference, such as anti-western rhetoric. Next, two filters were applied to the dataset. The first filter consists of users that have shared Russian propaganda content appearing on a list of 200 websites taken from www.propornot.com. This site has, and continues to, identify sites that produce or propagate Russian propaganda. The second filter identifies authors who are associated with the Internet Research Agency (IRA) and who have been labelled as Russian Trolls by Twitter. This data set was released by Twitter in 2018. If the filters produce results, the tops users were scrutinised (see next paragraph) and the tweets’ contents were further examined for extreme narratives. What content might support the Russian agenda? To whom would a group of BOTs, promoting such a view, be beneficial?

The next step took a closer look at the users identified by the filters. Users were examined to determine the nature of their account (i.e. is it a BOT? Where are they from?). A prototype analysis tool was applied Figure 2, based on Authentic Chatter [11], that classifies users as either BOT or Official (an official NATO, Government or organizational account) using machine-learning. As well, the users’ Twitter sites were qualitatively examined to confirm the automated findings. The following steps were used to verify that a suspected BOT, Figure 3, was actually a BOT after detection by the Authentic Chatter method:

1. Check if user exists;
2. Look at user photo;
3. Look at user name for oddities;
4. Look at number of tweets compared to date of creation;
5. Look at following vs. followers (are they similar?);
6. Check for posting behaviour e.g. all retweets?
7. Look at likes;
8. Use common sense regarding naming e.g. ‘nationalist’;

9. Use common sense regarding the bio; and,
10. Reverse image lookup on suspicious profile pictures and do a Google search.

Once suspicious users were classified as BOT, a new set of filters were developed to include these BOTS. As additional BOTS, or users that interacted with the BOTS, were found in the various datasets, these too were added to the filter. The new filters were then run using previously examined datasets. Again content derived from this new filtering was examined and compared to the previous results.

The above steps were designed to filter the dataset down to the suspected Russian content. At this point the content was again examined as were the top hashtags and top words. The user interactions were examined to look for network effects; was there evidence of amplification or any of the BEND manoeuvres?

The dataset was then filtered to include only the users which were highly suspected of being Russian controlled or users that were interacting with the BOTS. The content was examined for narrative imbalances. Finally, the three data sets were cross-examined for inter-narrative similarities and for BOTS acting across wedge issues.

3.0 RESULTS

3.1 Ethical issues Dataset

The period of analysis was from the 11th of June to the 10th of August. In total, there were 7,174 results from 3,694 publishers (users). The initial attempt to identify Russian propaganda in this dataset used pre-set filters (blacklists) to look for examples of tweets produced by known sources of Russian propaganda (again, derived from the Twitter IRA dataset and based on a previous blacklist constructed from users that frequently link to Russian propaganda sites identified by PropOrNot). In this dataset, there were no examples of data from or mentioning users from either of these blacklists.

Since the pre-existing blacklists failed to identify and users, a prototype tool developed for identifying different categories of users on Twitter was employed: the Authentic Chatter chart, Figure 2. This tool relies on machine learning algorithms trained on examples of official accounts (media personalities, government agencies, corporate accounts) and BOT accounts (as identified by Botometer, BotSentinel and botcheck.me). It gives a probability of any account being an official account, a BOT, both or neither. Attention was focused on accounts with more than a 50% chance of being a BOT.

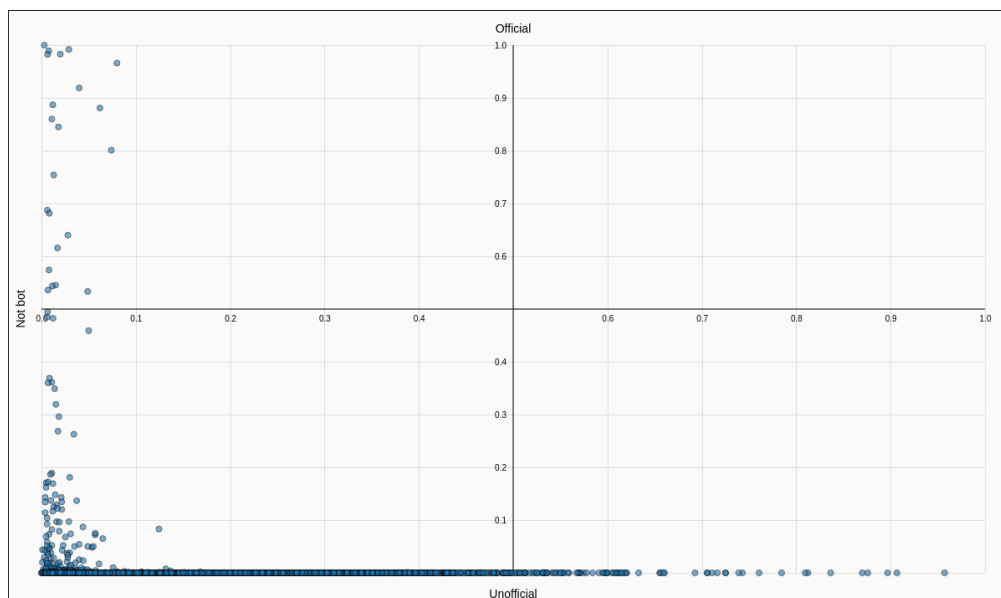


Figure 2. Automatic detection of Official or BOT users. The users in the bottom left quadrant are likely to be ‘authentic users’ not controlled by others not part of an official organisation.

The result of this filtering gave 43 actors that have at least 50% probability of being a BOT or Troll, producing 316 posts out of 23,819 overall collected results. Hashtags from these accounts are strongly negative towards a left of centre viewpoint. Some of the most active accounts, including @Working_Cdn2019 have (as of September 2019) been taken down although the Nexalogy software still preserves the data that was collected from Twitter before an account suspension. Another suspected BOT was LisaMar91564392, Figure 3. This account was examined for BOT-like qualities.



Figure 3 The Twitter page for user LisaMar91564392

Looking at the content of the suspected BOTs, it was evident that they focused on a right leaning stance within ethical issues. The BOT’s narrative focus can be seen by (a) the list of hashtags used, and (b) by manually examining the 75 tweets from the set (as this was such a small number of tweets, manual examination was feasible). While a set of automated accounts that were pushing a pro-Russian narrative was

identified, the next step was to attempt to determine if there is a foreign connection to these BOTs.

3.2 Pipelines Dataset

The pipeline dataset contained 80,903 posts and 57,620 publishers from June 11th to August 10th. To look for evidence of Russian interference, the propaganda filters were again applied. While the IRA list returned no results, the PropOrNot list found 83 posts by 73 users, confirming that at least some of the users were associated with Russian propaganda sites in this Canadian pipeline discussion. Investigating the top hashtags, Figure 4, and manually inspecting a sample of tweets indicated that these users were pushing a largely leftward (politically) leaning set of themes e.g. destroying organic farms or pro-Iranian and anti-Saudi messages. There were several discussions of Cold War themes that seemed more suspicious given the original topic of discussion.

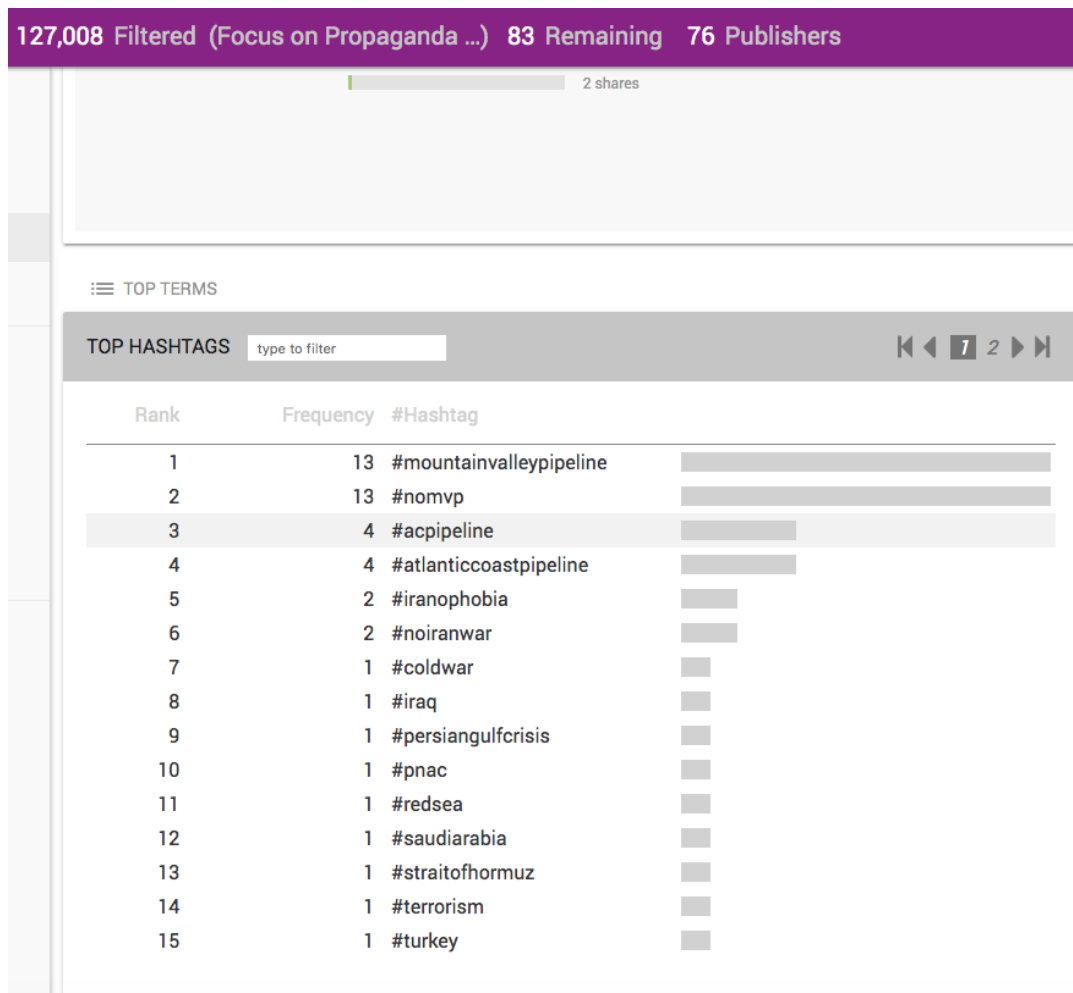


Figure 4. The top hashtags associated with the 73 users identified using the Russian propaganda content filter.

In order to examine if the BOTs from the ethical issues dataset also occurred in the Pipeline dataset, a search for tweets that mentioned or were produced by the same set of BOTs from ethical issues was conducted. Taking the top 15 most prolific BOTs from ethical issues, a filter was made and applied to the pipeline data for these users. The result found 22 posts and 13 users of which 5 users were the same. All of the 8 (13-5) new users were examined. An interesting tweet was found, Figure 5. It contained a long list of '@users'

(with sort messages and link at the end - this link leads to a version of God bless the USA by Lee Greenwood) that looked suspiciously BOT-like as well. Upon verification by searching for each users' Twitter page, it was found that many of these suspicious users have been deleted by Twitter.

```
. description: "@BrentDgls @Danbalkwill1 @BobRae14 @georgewcowles
@MonicaFibonacci @Canadian_Chris_ @lassiter1550 @OldManDuke @muchmore2cents
@pacopilbakalao @phil_rack @leighgt @indivcan @backmarker3 @JohnAll63815927
@grossw04 @MarkMe60 @GirouxRheal @trains72 @torontoensis @Blueyes9445
@vesnalaaurie8 @HaveWeAllGoneM1 @NewImproved9 @jay_slatter @daveb2561
@AsuAdanac @defiantcanuck @Friends0Science @Canadian_logic_ @IsExtortion
@RRidley11 @LDawg05 @kstokesvies @SusanIverach @lambert_pp @JIsaacsonV2
@laShawner951 @nemo_gratis @FayMary3 @jacksurfs204 @UpKeeks @molly6342
@Mim19561 @ThedeplorableM @DawsonMdhoust @JonDavi87753992 Sounds just like
the Pipeline expansion ??? https://youtu.be/Q65KZIqay4E"
```

Figure 5. A suspicious tweet that has a large sender list and a short message with a link.

Having found this interesting connection and its corresponding tweet, the original 15 users from the ethical issues set and the @user mentions from the suspicious BOTs were combined to create a new Ru specific blacklist filter. In order to see if there was a connection between this blacklist and Russian propaganda, the intersection between the Ru blacklist and the PropOrNot blacklist within the pipeline data was examined. The results, showed 50 users making 75 tweets for the time period. This showed overlap in terms of BOT and suspicious activity between the Canadian pipeline and the ethical issues discussions.

3.2.1 BOT Narratives in Pipeline Dataset

To evaluate the BOT narratives, the pipeline dataset was filtered for users or mentions of users present on one of the two blacklists (the new Ru blacklist and the PropOrNot blacklist). There was a high-level discussion of why climate change, as provoked by the Anthropocene, is an incorrect hypothesis. People were actively engage with the @friendsoscience, which was being promoted by the BOT network. This promoting consisted of replying to those whom reply to the account, sharing anti-renewable energy memes, and anti-left messaging. Using the BEND framework, it was found that:

1. DISTORT was present in response to the 12 year window for the point of no return on climate change;
2. AMPLIFY of Albertan pundits such as Friends of Science Barry Cooper was a pattern of behaviour by pro-Russian accounts, perhaps in order to foment Alberta separatism or undermine Canadian federalism? Previous research also showed that Ezra Levant was promoted by Russia-friendly accounts. There is also anti-protester and pro-pipeline narrative shaping. For instance, @LisaMar91564392 was present in this cross-section as an anti-left, but pro-pipeline oil & gas, and this BOT was present in the ethical issues topic issue. BOTs also pushed a Libertarian viewpoint.
3. AMPLIFYING climate change skeptics and far-right politics including anti-left sentiment in Canada is the main theme in this cross-section, in particular identified BOTs are AMPLIFYING @justinshafer1 whose profile links CPC, NRA and US Republicans.

3.3.2 Cross-Section Inter-Narrative Analysis

A filter that combined the BOTs from ethical issues and the Pipeline datasets was now applied to the previously examined ethical issues data. This was testing the attribution of Russian activity derived from the method to this point. There were 13 users with 59 tweets. These users were part of BOT nets discovered in both the ethical issues and Pipeline discussions. Looking at the top hashtags and words, Figure 6, the BOTs' expressed extremist views towards the left. The BOTS that were associated with Russian propaganda in this data sample were anti-left.

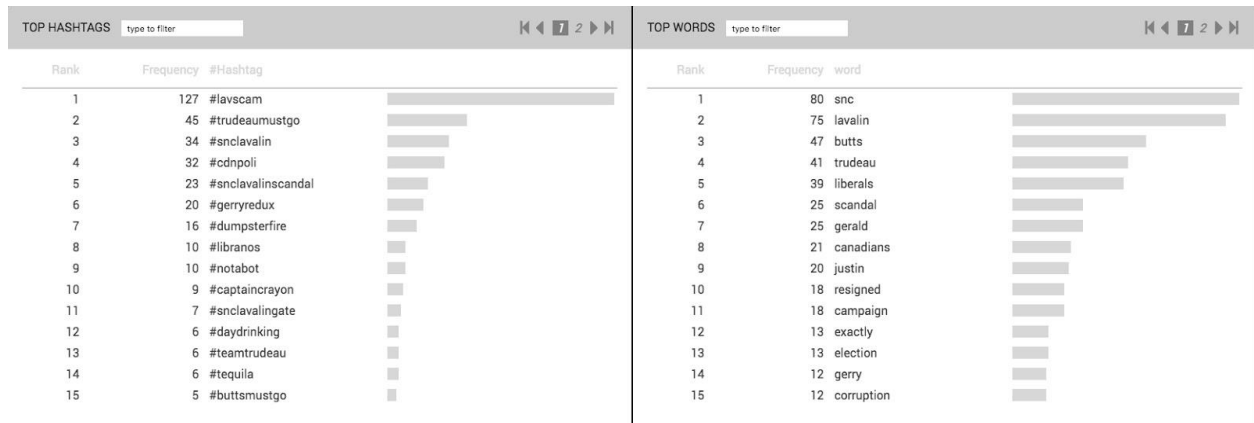


Figure 6. Top Hashtags and Words associated with the combined ethical issues and Pipeline BOT filters

3.3 Climate Change Dataset

In order to clarify the understanding from pipelines, the broader wedge issue of Climate Change, containing 500,499 tweets, was used for validation. The original BOT list from the ethical issues dataset was applied to Climate Change resulting in 9 publishers making 16 Tweets. This revealed some new BOTS but also some ‘useful idiots’. Useful idiots are defined as ordinary users who naively and unknowingly follow and spread the messages promoted by BOTs. Next the propaganda content filter revealed 445 tweets from 400 publishers. This filter was combined with the filter from ethical issues and Pipelines, which revealed 227 publishers making 491 Tweets. An actor interaction map, Figure 7, was produced. The actor interaction map shows the social network of conversations and their directionality on Twitter, with Influencers in green that are referenced more than they mention others, and loudspeakers in blue who reference others more often than they are referenced themselves. It shows that there were two tiers of targets in Green, e.g. larger volume influencers and smaller volume influencers within the core malign network. The targeting, it appears, was done by two BOT networks, visible in the upper and lower quadrants of the map. There were BOTs that made tweets mentioning many people and then the targets’ replies to those tweets, and those replies often included the original @ list. This technique effectively amplified the propagation of the message, as the initial list was made by a BOT but the replies were made by real human users. This is a known technique used to game the Twitter algorithm. This is an ENGAGE strategy according to BEND. There was also BOOST in terms of amplifying the voice of climate change skeptics. This is another way to use Nexalogy to discover narratives, for example the top two most mentioned accounts, mentioned by BOTS are climate change sceptics, but ones that have ‘nuanced’ positions against climate change as opposed to simple refusal. What this shows is the DISTORT and DISMISS techniques also being applied.

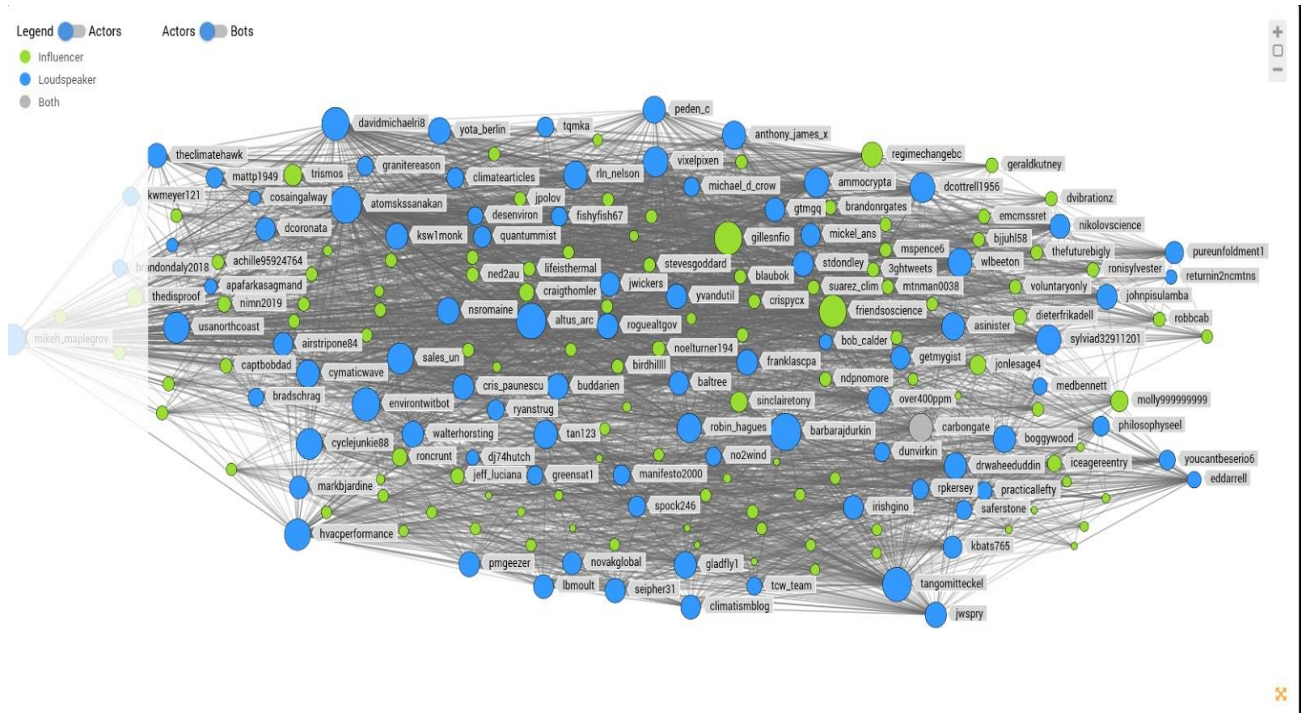


Figure 7. The actor interaction map showing the connections between BOTs identified in the ethical issues and Pipeline datasets applied to the Climate Change dataset.

4.0 DISCUSSION

A set of suspicious users was found using two Russian propaganda filters. A new AI BOT detection technique identified additional suspicious users, some that were identified as overlapping. The narratives pushed by this BOT network emphasised an anti-western, far-right approach. In particular, it demonstrated an anti-western perspective in the ethical issues and in the pipeline topics. In the climate change discussions they the BOTs supported pipelines and opposed the attribution of climate change to human activity. This resembles previous research on the action of Russian agents in the 2016 United States presidential election [12].

As mentioned in the introduction, the BEND information manoeuvres framework, Figure 1, allows analysts to map information activities to tactical categories. In our investigation, we observed the following types of BEND tactics being implemented.

In the Content axis of BEND, ENHANCE was being used in relation to anti-tarsands (the oil producing region of Canada) and anti-western content, and the narratives being promoted involve negative emotions - a DISMAY tactic. In the Community axis of BEND, BOOST and BUILD were observed across wedge issues using the BOT network. The BACK manoeuvre takes place in @mentions of the party leaders. BRIDGE occurred between wedge topic areas. To confirm BEND findings (especially with regards to the Community aspects) a SNA tool like ORA-pro could be used to increase our confidence levels - for example in order to see propagation within the network and compare BOT scoring across tools.

The results from the propaganda filters were used to attribute Russian influence in the results. However, there is a potential of bias stemming from the use of geolocation in the original queries. The benefit of geolocating is less overwhelming data, but the con is that it can bias the dataset towards Twitter’s definition of a location source. Twitter queries involving geolocation take advantage of a combination of Machine

Learning attribution of the author of a tweet to the selected location, and actual geocoding of tweet. Since the searches looked specifically for data from inside Canada, tweets originating from Venezuela or Russia would have been eliminated. Hence, leading to a bias within the sample against foreign influence not pretending to originate within Canada. To verify the degree of bias, the ethical issues dataset was also captured globally as of Sept. 18th and, as it is a very Canadian specific issue, will be checked for potential users from outside of Canada. This post-analysis tweaking reinforces a general methodological point that the queries for any dataset should be re-examined every few months and modified if necessary (to check if bias has been detected or new indicative terms have entered the discussion).

This method can only produce an inferred attribution due to the use of the Russian propaganda filter. This attribution may not be valid. It is very hard to trace exactly who controls a BOT. There were 5 BOTs that overlapped between Pipelines and ethical issues. These two issues are quite different, and one might not expect the users to be similar between topics. This adds to the indication of interference by a third party. It makes sense that Russian BOTs would promote anti-pipeline (#tarsands is the only hashtag in the overlap sub-set) and also promote anti-western discussions, in order to promote Russian geopolitical interests. This makes sense from a Russian perspective because less oil produced by Canada, and being against users seen as pro-pipeline, means less Canadian oil would make it to market.

Many of the BOT accounts that were found in the datasets has since been deleted by Twitter. So, the fact that accounts like @Working_Cdn2019 have been disabled by Twitter shows that there is uptake on Twitter's behalf and actions undertaken in relation to foreign interference within Canadian discussions (or perhaps these same BOTs are being used in other topic areas). Because tweets persist on the NexaIntelligence software, this research was able to detect these now deleted accounts, whereas the removed accounts would not be obvious to everyday users.

5.0 CONCLUSION

We identified evidence of BOTs, with attribution to Russia, meddling in Twitter discussions on Canadian wedge issues. This meddling was mapped onto the BEND framework. In particular, this meddling seemed to be focused in BOOSTING and ENHANCING extreme viewpoints on wedge issues. Focusing on the BOT net identified during this analysis, the propaganda promoted a populist/far-right perspective on establishment politics and pipeline/climate change development (human caused climate change denial).

This research was primarily interested in determining the effectiveness of the method and tools for identifying possible foreign influence. Follow on research will look at suspicious users and narratives and how these change over time. While the methodology will be further streamlined, the current NexaIntelligence tool can already be used in identifying foreign information warfare activity on Twitter. In particular, using a combination of blacklist filters, AI driven tools and analytic methodology, an analyst can quickly detect BOT activity, determine probable attribution, identify key messages and narratives, and determine the tactics using the BEND model. Without these tools, this type of activity would take significantly longer to detect and understand.

One of the current challenges to this methodology is correctly identifying the relevant narratives and then determining the prevalence and development of those narratives across the Twitter community and through time. To enhance narrative analysis, a tool is in development that will detect 'Tweet Caboodles' (individual tweets that are tightly connected in meaning). These caboodles can be combined to identify higher-level issues and topics. Caboodles, issues, and topics will be integrated to help identify how different types of narratives evolve and are propagated by the social network.

6.0 REFERENCES

- [1] NATO, *The dynamics of Russia's information activities against Ukraine during the Syria campaign*. . 2016: NATO Strategic Communications COE, Riga.
- [2] Berzins, J., *Russia's new generation warfare in Ukraine: Implications for Latvian defense policy*, C.f.S.a.S. Research, Editor. 2014, National Defence Academy of Latvia. p. 15.
- [3] Paul, C. and M. Matthews, *The Russian "Firehose of Falsehood" Propaganda Model*. 2016, Rand Corporation.
- [4] Shane, S., *The Fake Americans Russia Created to Influence the Election*, in *The New York Times*. 2017: New York.
- [5] Bradshaw, S. and P.N. Howard, *The Global Disinformation Disorder: 2019 Global Inventory of Organised Social Media Manipulation. Working Paper 2019.2*, in *Project on Computational Propaganda*. 2019: Oxford, UK.
- [6] Bartles, C., *Getting Gerasimov Right*. Military review, 2016. 96: p. 30-38.
- [7] Allen, T.S. and A.J. Moore, *Victory without Casualties: Russia's Information Operations*. Parameters, 2018. 48(1).
- [8] McKelvey, F. and E. Dubois, *Computational Propaganda in Canada: The Use of Political Bots*, in *Computational Propaganda Research Project*. 2017, University of Oxford: Oxford, UK.
- [9] Gorwa, R. and D. Guilbeault, *Unpacking the Social Media Bot: A Typology to Guide Research and Policy*. Policy & Internet, 2018.
- [10] Beskow, D. and K.M. Carley, *Social Cyber Security: An Emerging National Security Requirement*. American Military Review, 2019.
- [11] Forrester, B., *Authentic Chatter*. Computational and Mathematical Organization Theory Journal, In Press.
- [12] Ruck, D.J., et al., *Internet Research Agency Twitter activity predicted 2016 U.S. election polls*. First Monday, 2019. 24, Number 7 - 1 July 2019.

Towards Big Data in the Tactical Domain

Frank T. Johnsen

Norwegian Defence Research Establishment (FFI)

P.O. Box 25, 2027 Kjeller
NORWAY

frank-trethan.johnsen@ffi.no

ABSTRACT

With the Internet of Things comes new possibilities for harvesting information, but the potential vast amount of data sources makes this a prime candidate for big data analysis. Further, with such information needing to be connected to the tactical domain, perhaps even originating in the battlefield itself, there is a need to either transmit data for processing elsewhere, or to apply smart algorithms and approaches that can manage and analyze the data where it originates. In tactical networks, the limitation on available communications makes this a challenge. Typically, tactical networks are referred to as disadvantaged grids or DIL (short for Disconnected, Intermittent and Limited) environments. These communications challenges mean that stock big data approaches likely cannot be applied as-is, but need to be tailored specifically for the tactical domain, or the algorithms perhaps be employed there as other places, but will require supporting services that are tailored for DIL environments. Examples here include tactical edge cloud computing and delay and disruption tolerant data dissemination services. Preferably, one should promote interoperability by leveraging open standards and APIs where applicable.

1.0 INTRODUCTION

For military forces operating at the tactical level, shared situational awareness is a requirement for efficient decision making. Simply put, *situational awareness is knowing what is going on around you* [1].

With emerging and evolving technology, new and faster ways for information sharing are made possible. Having reliable and up-to-date information about the operating environment, including the position and status of friendly forces and other assets available when and where needed, is essential to situational awareness. This type of information can be gathered in a number of different ways, and one civilian trend which is gaining traction also for military use is the deployment of cheap sensor systems as a means to augment the information already available to military decision makers today. This civilian trend is known as the Internet of Things (IoT), can be defined as follows [2]:

“IoT describes the revolution already under way that is seeing a growing number of internet enabled devices that can network and communicate with each other and with other web-enabled gadgets. IoT refers to a state where Things (e.g., objects, environments, vehicles and clothing) will have more and more information associated with them and may have the ability to sense, communicate, network and produce new information, becoming an integral part of the Internet. A widespread Internet of Things has the potential to transform how we live in our cities, how we move, how we develop sustainably, how we age, and more.”

The reason why IoT has become commonplace during the course of the last five years is that three important enablers have come into place:

- Availability of cheap sensors

- Cloud computing
- Powerful smartphones

Cheap sensors allow for a low-threshold approach to rapid prototyping. These days, sensors are easily procured from eBay, deal extreme, and other online vendors. A large number of sensors can potentially generate huge amounts of data. Gartner [3] defines *big data* as *high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing that enable enhanced insight, decision making, and process automation*. Such data cannot be handled by traditional means like a single database or server, so cloud computing serves as a back-end for IoT systems and can handle challenges posed by big data. Finally, smart devices like phones and tablets provide a common control panel for consumers in the IoT context.

The contribution of this paper to the IST-178 inter-panel / inter-group workshop on big data challenges, situation awareness and decision support is a summary of relevant, related IST activities to promote inter-group discussions and potential collaboration.

2.0 THE INTERNET OF THINGS

Figure 1-1 shows how different components come together, forming an IoT ecosystem. Here, we see the “Things”, which are connected to some form of local network. Devices can for example be deployed in homes (so-called smart homes where the sensors are typically privately owned) or on/in buildings and other infrastructure (so-called smart cities where the sensors are typically owned by the government). Typically, IoT systems are connected to the Internet via their local networks, where they can report their data to (and be controlled and updated from) services deployed in the Cloud. Smartphones are often used as end-devices for visualizing data from (and controlling) the sensors, as IoT vendors typically release apps that can be used to connect to the Cloud services. These services may offer a standardized application programming interface (API) or it can be proprietary.

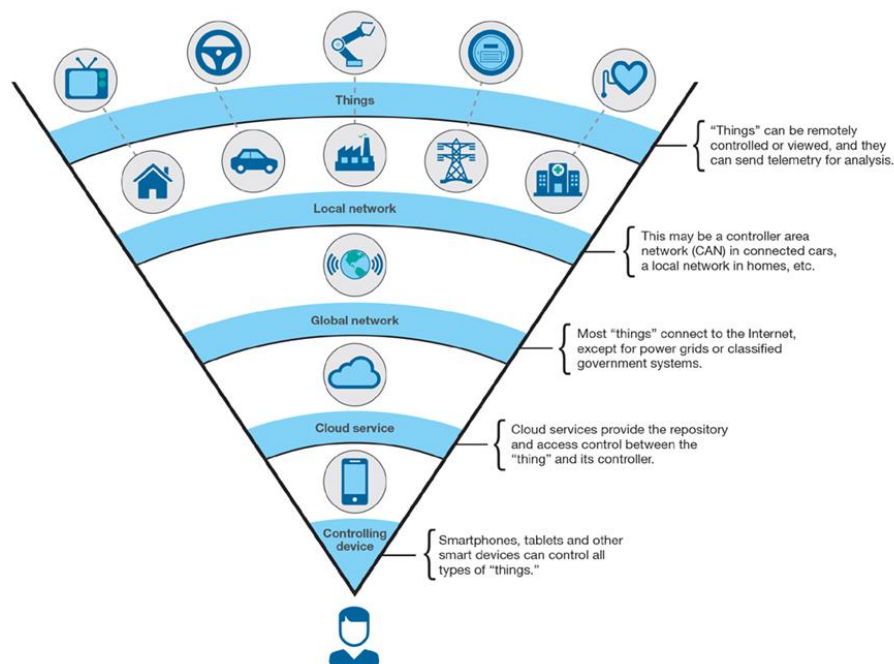


Figure 1-1: Human view of IoT (from [4]).

IoT can potentially be used for a diversity of applications. There have been many business ideas in the healthcare sector, logistics and other areas that give rise to a number of applications fueling the current IoT trend. IoT as a concept is definitely relevant in a defense context. An example of this is the work described in [5], which deals with the use of sensor networks and lightweight processing platforms that require low power. IoT includes several disciplines, as one needs networking, embedded hardware, software architectures, sensors, information management, data analysis and visualization to fully leverage the concept. A key component within IoT is the use of distributed online devices that communicate using Internet protocols. A “thing” in IoT may be any device that is able to communicate, gather data or offer some kind of control. With this wide interpretation of “things”, IoT may include, but is not limited to: Vehicles, appliances, medical equipment, power grids, transport infrastructure and production equipment. Military organizations can exploit IoT deployed in battlefields and operational theaters to improve situational awareness, mission performance and achieve information superiority [6]. Within NATO, the Research Task Group (RTG) IST-147 “Military Application of Internet of Things” has investigated how to a coalition force can use IoT to augment situational awareness in military operations in smart cities [7].

3.0 CHALLENGES AND OPPORTUNITIES

There are many challenges posed when considering employing IoT for military purposes. However, one can envision two major focus areas that need to be considered and their challenges addressed before IoT can start being using also in the tactical domain. These areas are:

1. Interoperability
2. Data dissemination

Let us explore each of these in turn.

3.1 Interoperability

Today there is an emphasis on using commercial off-the shelf (COTS) products where possible, because it is considered a cost-effective way of acquiring a capability. This idea is well rooted in NATO, and has been considered foundational for an effective Network Enabled Capability (NEC) as identified in the NATO NEC Feasibility Study [8]. This study also pointed out that the principles of a service-oriented architecture must be taken into account when building distributed systems. These observations can be continued within the IoT venture, as there will also be a need to build large, efficient and interoperable systems. Following the conclusion of IST-147, the recently started IST-176 group aims to continue where IST-147 left off, and further investigate scalability and interoperability issues of combining civilian and military sensors with Command and Control (C2) systems. The IST-176 group, aptly titled “Federated Interoperability of Military C2 and IoT Systems”, is currently pursuing two main lines of work: IoT security and architecture. Security is all too often neglected when IoT devices are designed and manufactured [9]. When seeking to leverage available sensors, this may include a mix of military sensors that members of the coalition control and own, government owned sensors in the area of operations (e.g., installed in a smart city), or even privately owned sensors that may be in the area. Given this mixture of potential data sources, trust in the data becomes an issue. Hence, it is important to have some approach to identifying the data source and be able to take this into account when assessing the value of the information. As for the interoperability, there are today many different competing specifications and standards, as well as a large number of proprietary approaches to IoT, so that identifying suitable approaches from a technical architecture and interoperability viewpoint will be a major part of the group’s undertaking [10].

3.2 Data dissemination

Civilian IoT approaches are centered around the Internet as the backhaul network (see Figure 1-1) to reach the cloud where big data can be stored and processed. For tactical military use, with sensors deployed in the field, it would be necessary to convey data over tactical data links. As opposed to the Internet, with high reliability and high throughput, tactical links operate in a contested environment leading to disconnections, intermittent connectivity and severe limitations on throughput (aka DIL environments). Services and data exchange in tactical networks have been investigated in the IST-090 “SOA challenges for real-time and disadvantaged grids”, IST-118 “SOA recommendations for disadvantaged grids in the tactical domain”, and IST-150 “NATO Core Services profiling for hybrid tactical networks” series of RTGs, which have shown that approaches that work for the Internet very often need to be adapted or optimized to become usable in the tactical domain [11]. Also, IST-161 “Efficient Group and Information Centric Communications in Mobile Military Heterogeneous Networks” has investigated several popular dissemination mechanisms, and have shown that these work poorly (or not at all) in tactical networks [12]. A prominent example in this respect is *Kafka*, which is used *for real-time streams of data, to collect big data, or to do real time analysis or both* [13]. *Kafka* was among the frameworks IST-161 tested and found unsuitable in tactical networks. This indicates that there is a need to further investigate approaches to, and possibly develop new mechanisms and approaches for, big data in tactical networks. For example, IST-150 has performed several experiments featuring the MQTT protocol in tactical networks, which shows promise as a pure dissemination mechanism since it is more lightweight than competing standards [14]. Conversely, IST-161 work shows that some proprietary protocols are even more efficient, which makes sense since these solutions were tailored for DIL environments like tactical networks [12].

Dissemination mechanisms aside, one approach to optimizing information processing could be to attempt to process as much as possible near where the data originates, and so only allow certain identified “events” to propagate across the tactical network. The analogy here would be the civilian approach of so-called *fog computing* where IoT data can be processed near where it originated and before the processed information is sent to the cloud [15]. One can anticipate that such an approach would be beneficial in tactical networks, since the inherent volume of big data will likely be too much for the network to handle. As for cloud computing, one cannot expect to have (or even desire to have) access to civilian cloud providers over the Internet from the tactical battlefield. Hence, cloud computing resources must be provided near the tactical edge or from devices (for example hosted as tactical cloudlets on vehicles) in the battlefield itself. The RTG IST-168 “Adaptive information processing and distribution to support Command and Control” is currently investigating cloud computing approaches for military tactical use. The group pursues two main research tracks: 1) Interoperability where clouds under different ownership domains, typically in a coalition, need to be able to access services and data deployed across different clouds. 2) The suitability of standardized solutions and approaches for cloud computing to providing tactical cloud computing resources. The group is currently limited to pursuing Kubernetes as the cloud computing technology for these purposes [16]. Kubernetes (K8s) is an open-source system for automating deployment, scaling, and management of applications, where emphasis is on ease of management and service discovery [17]. If IST-168 is successful in these endeavors, then the results would be a valuable contribution towards interoperable, tactical coalition clouds, which again can be seen as another step towards leveraging big data in the tactical domain.

4.0 SUMMARY

Big data can provide new insights, enhance situational awareness and be a valuable aid in decision support. Combining civilian Internet of Things (IoT) and military sensor assets makes a lot of information available, which needs big data approaches to handle. At the same time, increased information loads and processing needs tax the available resources at the tactical level. This paper has introduced and raised such concerns regarding big data in tactical networks as food for thought in the IST-178 inter-panel / inter-group workshop on big data challenges.

REFERENCES

- [1] M. Endsley, “Theoretical underpinnings of situation awareness: A critical review.” Situation awareness analysis and measurement, Jan. 2000, pp. 332.
- [2] IoT Special Interest Group, “Technology Strategy Board”, 2013.
- [3] Gartner. “IT Glossary – Big Data”. <https://www.gartner.com/it-glossary/big-data> (accessed 23 September 2019)
- [4] X-Force Research and Development, “IBM X-Force Threat Intelligence Quarterly 4Q 2014,” Doc # WGL03062USEN, Publish Date: Nov 2014. <http://www.ibm.com/security/xforce/downloads.html>
- [5] Wind River Systems. “The Internet Of Things For Defense.” White Paper, 2015.
- [6] N. Suri et al. “Analyzing the Applicability of Internet of Things to the Battlefield Environment.” IEEE ICMCIS 2016, Brussels, Belgium, May 2016.
- [7] F.T. Johnsen et al. “Application of IoT in Military Operations in a Smart City.” IEEE ICMCIS 2018, Warsaw, Poland, 22nd – 23rd May 2018.
- [8] P. Bartolomasi et al. “NATO network enabled capability feasibility study.” V. 2.0, October 2005.
- [9] R. Sfar et al. “A Roadmap for Security Challenges in Internet of Things.” Digital Communications and Networks. 4. 10.1016/j.dcan.2017.04.003. 2017
- [10] F.T. Johnsen et al. “Using Open Standards for Utilizing IoT Sensors in a Smart City Scenario.” International Command and Control Research and Technology Symposium (ICCRTS) 2018
- [11] P.P. Meiler et al. “Improving Integration between Tactical and HQ Levels by making SOA applicable on the Battlefield”, ICCRTS 2017, Los Angeles, CA, USA
- [12] N. Suri et al. “Experimental Evaluation of Group Communications Protocols for Data Dissemination at the Tactical Edge.” IEEE ICMCIS 2019.
- [13] J.P. Azar. “What Is Kafka? Everything You Need to Know.” <https://dzone.com/articles/what-is-kafka>, published 9 August 2017
- [14] M. Manso et al. “Mobile Tactical Force Situational Awareness: Evaluation of Message Broker Middleware for Information Exchange.” ICCRTS 2018
- [15] IEEE Standards Association. “IEEE 1934-2018 - IEEE Standard for Adoption of OpenFog Reference Architecture for Fog Computing.” <https://standards.ieee.org/standard/1934-2018.html>, Published 2 August 2018
- [16] H. Bastiaansen et al. “Adaptive Information Processing and Distribution to Support Command and Control in Situations of Disadvantaged Battlefield Connectivity.” IEEE ICMCIS 2019.
- [17] B. Burns et al. “Borg, Omega, and Kubernetes.”, Queue volume 14 number 1, January-February 2016, pages 10:70-10:93



Assessment of IoT Data Ingest Reliability for Urban Environments

James R. Michaelis

U.S. Army Research Laboratory
2800 Powder Mill Road, Adelphi, Maryland 20783
UNITED STATES

james.r.michaelis2.civ@mail.mil

Jade Freeman

U.S. Army Research Laboratory
2800 Powder Mill Road, Adelphi, Maryland 20783
UNITED STATES

jade.l.freeman2.civ@mail.mil

Adrienne Raglin

U.S. Army Research Laboratory
2800 Powder Mill Road, Adelphi, Maryland 20783
UNITED STATES

adrienne.raglin2.civ@mail.mil

ABSTRACT

Globally, Internet of Things (IoT) technology has seen significant growth in adoption and deployment, with total data generated by IoT devices forecast to exceed 850 Zettabytes by 2021. Civilian IoT infrastructures consist of a collection of constituent sensing, networking, and computational components, and can be viewed as data processing pipelines aimed at providing services for civilian benefit (e.g., environmental monitoring). Data ingest from this civilian IoT space has received interest from militaries worldwide, particularly to support establishment of situational awareness and management of urban operations. However, a number of challenges to military data ingest are known to manifest at different sections of corresponding IoT pipelines.

To support exploitation and utilization of IoT data for tactical purposes and integration with C5ISR (Command, Control, Computers, Communications, Cyber, Intelligence, Surveillance and Reconnaissance) systems, there has been a growing interest in assessment of IoT and supporting network protocols along dimensions of scalability, cost, range, energy efficiency, and ability to deploy rapidly, particularly in smart cities with massive IoT networks. This submission presents ongoing research into a collection of challenges to data ingest from civilian IoT infrastructures, placing focus on two current efforts: (1) Assessment of IoT communication protocol reliability in urban environments, (2) Methods to support assessment and prioritization of data obtained from civilian IoT infrastructures, according to both intrinsic quality assessment and value to mission needs.

1.0 INTRODUCTION

The Internet of Things (IoT) represents an emerging technological paradigm, consisting of numerous constituent networking, sensing, actuation, and computing systems. Globally, IoT technology has seen significant growth in adoption and deployment, with total data generated by IoT devices forecast to exceed 850 Zettabytes by 2021 [1]. A key driver for IoT growth lies in the emergence of smart city ecosystems

[2], aimed at providing Information and Communications Technology (ICT)-enabled services for the benefit of citizens and civilian organizations. Such services may range in scope (e.g., environmental monitoring [3], traffic management [4]), and can be coordinated by government organizations at the city and national levels. Likewise, citizen-led efforts have led to the emergence of grassroots IoT efforts, as exemplified by open IoT infrastructures such as The Things Network [5]. Management of these forms of ICT infrastructure within smart city environments represents a key sub-domain within existing big data research [6].

Continued advances in IoT technology have prompted new investigation into its usage for military operations under the emerging Internet of Battlefield Things (IoBT) paradigm [7]. Research in IoBT has sought to assess viability of Commercial-off-the-Shelf (COTS) IoT technology to augment and complement existing military sensing assets [8], provide decision support and mechanisms to establish situational awareness and understanding, as well as support next-generation artificial intelligence and machine learning systems.

Despite the potential of civilian IoT infrastructure as a data source for supporting military operations, a number of challenges impacting requisite data ingest are known to exist which include [7, 9]:

- Dynamic, potentially degraded networking conditions.
- Constraints on the ability of networks to deliver and process information for military consumers.
- Cognitive constraints on military personnel in handling large-scale data collections.

To support exploitation and utilization of IoT data for tactical purposes and integration with C5ISR (Command, Control, Computers, Communications, Cyber, Intelligence, Surveillance and Reconnaissance) systems, there has been a growing interest in assessment of IoT and supporting network protocols along dimensions of scalability, cost, range, energy efficiency, and ability to deploy rapidly, particularly in smart cities with massive IoT infrastructures. This submission presents ongoing research into a collection of challenges to data ingest from civilian IoT infrastructures, placing emphasis on two current efforts: (1) Assessment of IoT communication protocol reliability in urban environments, focusing on recent test and evaluation of the LoRaWAN (Long Range Wide Area Network) protocol [10, 11]; (2) Methods to support assessment and prioritization of data obtained from civilian IoT infrastructures, according to both intrinsic quality assessment and value to mission needs.

2.0 RESEARCH ISSUES FOR IOT DATA INGEST

IoT infrastructures are commonly defined to encompass collections of devices (e.g., sensing and actuation assets), designed to communicate with one or more centralized servers via requisite networking components (e.g., device gateways) [12]. Once aggregated, data from IoT infrastructures can in-turn be processed and made accessible to consuming parties via a collection of means (e.g., software APIs, web interfaces) [12]. In this form, civilian IoT infrastructures can be viewed as data processing pipelines aimed at providing services for civilian benefit (e.g., environmental and traffic monitoring [3, 4]).

Towards facilitating exploitation and utilization of civilian IoT data for supporting mission needs, a collection of research efforts are presently assessing IoT and supporting network protocols, particularly in smart cities with massive IoT networks. This section covers two case study research efforts: (1) Assessment of IoT communication protocol reliability in urban environments, (2) Methods to support assessment and prioritization of data obtained from civilian IoT infrastructures, according to both intrinsic quality assessment and value to specific mission needs. Following review of these efforts, a brief overview of alternate research topics of relevance to IoT data ingest will be provided.

2.1 Focus Area 1: Communication Protocol Coverage

Commercial IoT systems rely upon a collection of communication protocols to facilitate data ingest, which vary based on factors including: geographic coverage, supported data payload sizes, and power consumption [13]. Within smart city environments, the potential for obstruction of transmissions from IoT devices and receiving gateways can emerge as density of urban infrastructure increases [14, 15], potentially impacting collection of data from devices corresponding to particular areas of interest. Towards mitigating these challenges, research investigating IoT communication protocol reliability becomes of interest.

Case Study - Coverage for the LoRaWAN Protocol

Among existing commercial IoT protocols, LoRaWAN has gained significant adoption in civilian IoT infrastructures through combined support for low power, long range transmissions within IoT infrastructures [10]. Prior research has established LoRaWAN's support for IoT device transmissions at distances beyond 10 Km in ideal conditions. However, a key knowledge gap for LoRaWAN has been limited empirical knowledge on its coverage in the presence of dense, potentially obstructing urban infrastructure [11, 14].

Towards addressing this knowledge gap, recent research by scientists at the U.S. Army Research Laboratory involved a comprehensive test of LoRaWAN coverage in the city of Montreal through use of vehicle-mounted IoT transmitters [11]. A series of vehicle routes were driven, covering multiple sections of Montreal featuring varying geographic terrain and urban infrastructure. To support the LoRaWAN testing conducted, an ARL-developed IoT architecture was utilized based on a collection of COTS hardware and software. The supporting IoT architecture was originally developed as a means to support integration of COTS assets into broader C2 systems, capable of supporting expanded Situational Awareness and decision support capabilities in Tactical Operations Centers.

Figure 1 provides a diagram of LoRaWAN coverage established during the driving tests. For the LoRaWAN coverage testing, maximum transmission distance could reliably be established 5 Km from the receiver across Montreal's central business district. This maximum transmission distance was achieved for three separate LoRaWAN data transmission rates (each capable of transmitting different sized IoT device messages) on the North American ISM (Industrial, Scientific, and Medical) 915 MHz band. These findings were seen to help reinforce the reliability of LoRaWAN in the presence of dense urban infrastructure [11].

Summary of Identified Research Issues

Following from experimentation conducted in Montreal, a series of follow-on research issues were identified [11]:

- **Coverage Gap Analysis**, to assesses conditions where LoRaWAN coverage may be obstructed in urban environments (e.g., in the presence of skyscrapers or dense infrastructure).
- **Coverage per Configuration**, to assess impact of varying LoRaWAN transmission configurations (e.g., varying data rates) on coverage in urban environments.
- **Coverage Comparison Across Protocols**, involving comparison of LoRaWAN coverage in urban environments with alternate communication protocols, including Narrowband (NB)-IoT and SigFox.

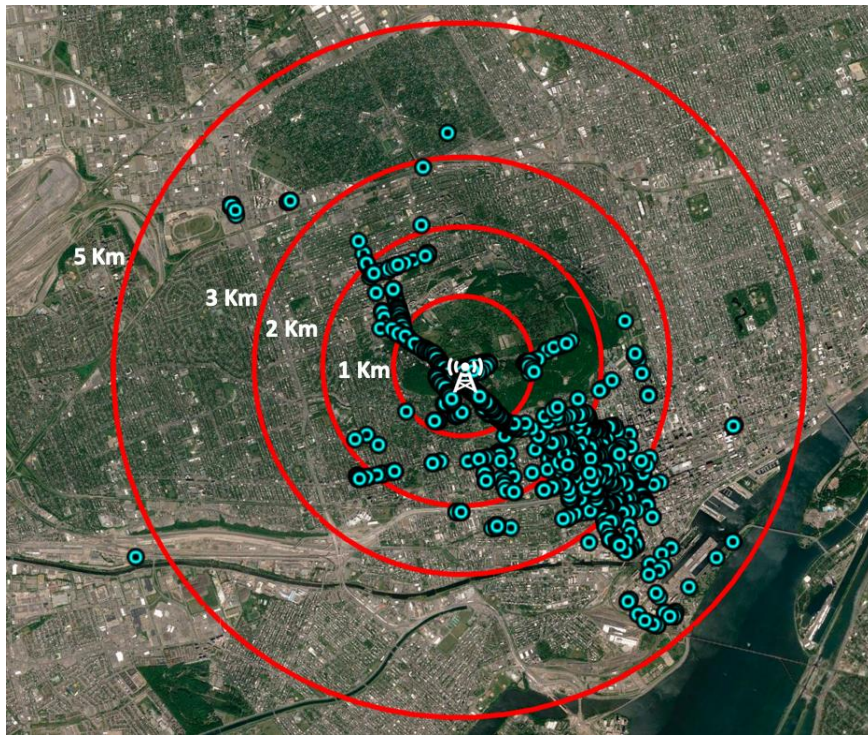


Figure 1: Visual plot of LoRaWAN testing results in Montreal, depicting locations of messages received by a centralized gateway. The circles (in red) denote the following distances from a centralized LoRa gateway: 1 Km, 2 Km, 3 Km, and 5 Km. Dots indicate locations where LoRaWAN messages were received during vehicle tests.

2.2 Focus Area 2: Value Assessment and Prioritization of IoT Data

To effectively leverage continued growth in civilian IoT infrastructures and accompanying data generation, military networking and information management systems must address a collection of known challenges, including [7, 9]:

- Dynamic, potentially degraded networking conditions.
- Constraints on the ability of networks to deliver and process information for military consumers.
- Cognitive constraints on military personnel in handling large-scale data collections.

Under operating conditions featuring limited network bandwidth for data transmission, as well as limitations on the ability of personnel to review corresponding information, it is not desirable to misuse these resources on low-value or under-processed data. Thus, methods to properly assess the value of particular units of information, prior to their transmission over networks and review by personnel, becomes an important element in facilitating data ingest by C5ISR systems to support mission needs. Towards supporting information delivery in tactical networks (e.g., [9]), methods for Value of Information assessment can make distinctions on what should and should not be delivered to decision makers and also place emphasis on delivering high-valued information with greater speed and precision. Such approaches aim to prioritize and filter units of data from operational environments (e.g., from IoT devices and services) accounting for both intrinsic quality assessment and Value of Information to mission needs.

Defining Value of Information for Information Objects

The concept of Value of Information can differ depending on one's goal. Typically, one wants to know the Value of Information when considering obtaining that data with *finite* resources. Here, such resources

could be time, budget, or other types of constraints and limitations. Even though the metrics for the value depend on the evaluator's goal, the common thread for the Value of Information can be supported by analytic methodology for qualifying or quantifying the potential benefit of the information in the face of uncertainty.

The progress of Value of Information study in various research domains, such as information theory and economics, have been made in providing its own problem-specific approach. In tactical networking research (e.g., [9, 16]), Value of Information has been defined using both intrinsic vs. extrinsic attributes of Information Objects – defined here as units of data derivable from IoT infrastructures. Intrinsic attributes can be viewed as measuring the inherent quality of an Information Object. For instance, an Information Object corresponding to audio data could have intrinsic quality attributes of bit rate and sample rate. Here, intrinsic attributes can help establish the intrinsic quality for a particular Information Object. By contrast, extrinsic attributes apply toward measuring the utility of an Information Object to meet a specific consumer's needs. Within the context of tactical operations, examples of extrinsic attributes could include geographic relevance (*is this information from a mission-relevant location?*), temporal relevance (*will I need this information soon for my mission tasks?*), and source reliability (*did the information come from a sufficiently trustworthy source for mission needs?*). Additionally, extrinsic attributes could measure presence of relevant information (*does this image contain mission-relevant features?*). Prior efforts in tactical networking research has viewed Value of Information as inherently building upon intrinsic quality assessment [9], while emphasizing the inherent difference between these assessment classes (i.e., an Information Object with high image quality may not have mission-relevant information, thereby having low inherent value).

Within tactical networking systems (e.g., [9, 16]), quantitative Value of Information assessment has previously been applied to prioritize Information Object delivery to Soldiers, through weighted averages of evaluation metrics each corresponding to particular Information Object attributes. An example of a weighted metric average for Value of Information assessment [17] takes the following form:

$$\text{VoI} = (\text{GR} * w_{\text{GR}}) + (\text{TR} * w_{\text{TR}}) + (\text{E} * w_{\text{E}}) + (\text{I} * w_{\text{I}}) + (\text{IC} * w_{\text{IC}}) + (\text{SR} * w_{\text{SR}})$$

For each evaluation metric, a quantitative value is calculated along with a corresponding weighting of importance. In turn, the metrics listed in this equation can be defined as follows:

- **GR (Geographic Relevance):** Estimated based on where particular data for an Information Object was obtained, relative to a consumer's mission location(s). For example, the distance between where an image was taken and a location of mission relevance.
- **TR (Temporal Relevance):** Estimated based on when an Information Object will be needed by a consumer for mission tasks.
- **E (Expiration):** Estimates when the content of an Information Object will become too stale for mission needs.
- **I (Importance):** A value provided by a Subject Matter Expert (SME) or automated process, denoting an individual Information Object's importance specific to particular consumers and mission tasks.
- **IC (Information Content):** An assessment of the intrinsic significance of an Information Object's content for particular mission needs, as defined by an SME or automated process.
- **SR (Source Reliability):** An assessment of the reliability / trustworthiness of an Information Object's source or provider, as defined by an SME or automated process.

Another approach in current research in Value of Information quantification is contextual adaptive learning [18]. This approach is based on the historical utility of an Information Object, building on information derived from the context of a consuming user. Historical and contextual factors are

determined dynamically via learning algorithms, enabling the utility of particular Information Objects to be predicted.

In summary, Value of Information assessment is seen as having a particularly rich set of research challenges, which include development of models for both Soldier context (e.g., concerning environmental/physiological factors) as well as mission state. In both the policy and learning based methodologies described above, the following factors are seen as critical elements in designing Value of Information assessment models:

- **Factor Identification**, identification of appropriate metrics to apply to Value of Information assessment calculations.
- **Factor Assessment**, identification of methods to support calculation of Value of Information metric values and accompanying weightings of importance.
- **Definition of Supporting Models**, aimed at representing mission state, including mission tasks and operational conditions. Recent efforts tied to semantic models of mission planning and execution (e.g., [19, 20]) are of particular relevance.

2.3 Additional Relevant Research

Towards supporting IoT data ingest from smart city ecosystems, a number of additional research areas are presently under investigation. These include efforts to account for heterogeneity in IoT service APIs [16], techniques to account for IoT asset security and trustworthiness [21], as well as methods to support assessment and presentation for Uncertainty of Information derived from IoT sources. Uncertainty of Information values can be generated at various steps along IoT data pipelines, but consideration is needed on how that uncertainty is both quantified and communicated to military personnel based on the underlying properties of the data [22].

4.0 CONCLUSION

The Internet of Things (IoT) represents an emerging technological paradigm, consisting of numerous constituent networking, sensing, actuation, and computing systems. Continued advances in IoT technology have prompted new investigation into its usage for military operations under the emerging Internet of Battlefield Things (IoBT) paradigm. Research in IoBT has sought to assess the viability of Commercial-off-the-Shelf (COTS) IoT technology to provide decision support and mechanisms to establish situational awareness and understanding, as well as support next-generation artificial intelligence and machine learning systems. This submission has aimed to cover a collection of relevant research directions aimed at addressing challenges in data ingest from civilian IoT infrastructures. Addressing these challenges, in turn, is seen as key to facilitating civilian IoT spaces as a source of big data in supporting military operations.

REFERENCES

- [1] Cisco global cloud index: Forecast and methodology, 2016-2021. (2018). Cisco. Online: <https://www.cisco.com/c/en/us/solutions/collateral/service-provider/global-cloud-index-gci/white-paper-c11-738085.html> [Last Accessed: 26 September 2019]
- [2] A. Zanella, N. Bui, A. Castellani, L. Vangelista, and M. Zorzi, 2014. "Internet of things for smart cities," *IEEE Internet of Things Journal*, 1(1), pp.22-32.
- [3] B. Ahlgren, M. Hidell, and E. Ngai, "Internet of things for smart cities: Interoperability and open data," *IEEE Internet Computing*, 20(6), pp. 52-56, 2016.
- [4] P. Ta-Shma, A. Akbar, G. Gerson-Golan, G. Hadash, F. Carrez, and K. Moessner. "An ingestion and analytics architecture for iot applied to smart city use cases," *IEEE Internet of Things Journal*. 2017 Jun 30; 5(2):765-74.
- [5] The Things Network. (2019). Online: <https://www.thethingsnetwork.org/> [Last Accessed: 26 September 2019]
- [6] R. Kitchin, "The real-time city? Big data and smart urbanism," *GeoJournal*, 79(1), pp.1-14, 2014.
- [7] A. Kott, A. Swami, and B.J. West, "The internet of battle things," *Computer*, vol. 49, no. 12 (2016): 70-75.
- [8] T. Abdelzaher, N. Ayanian, T. Basar, S. Diggavi, J. Diesner, D. Ganesan et al. "Will distributed computing revolutionize peace? The emergence of Battlefield IoT," In 2018 IEEE 38th International Conference on Distributed Computing Systems (ICDCS), pp. 1129-1138. IEEE, 2018
- [9] N. Suri, G. Benincasa, R. Lenzi, M. Tortonesi, C. Stefanelli, and L. Sadler, "Exploring value-of-information-based approaches to support effective communications in tactical networks," *IEEE Communications Magazine*, 53(10), pp.39-45, 2015.
- [10] LoRaWAN Specification (2019). Online: <https://loro-alliance.org/about-lorawan> [Last Accessed: 26 September 2019]
- [11] J. Michaelis, A. Morelli, A. Raglin, D. James, and N. Suri, "Leveraging LoRaWAN to support IoBT in urban environments," In *2019 IEEE 5th World Forum on Internet of Things (WF-IoT)* (pp. 207-212). IEEE, 2019.
- [12] A. Al-Fuqaha, M. Guizani, M. Mohammadi, M. Aledhari, and M. Ayyash, "Internet of things: A survey on enabling technologies, protocols, and applications," *IEEE communications surveys & tutorials*, 17(4), pp.2347-2376, 2015.
- [13] R.S. Sinha, Y. Wei, and S.H. Hwang, "A survey on LPWA technology: LoRa and NB-IoT," *ICT Express*, 3(1), pp.14-21, 2017.
- [14] P.J. Radcliffe, K.G. Chavez, P. Beckett, J. Spangaro and C. Jakob, "Usability of LoRaWAN technology in a central business district," in *2017 IEEE 85th Vehicular Technology Conference (VTC Spring)* (pp. 1-5). IEEE, 2017.
- [15] F. Adelantado, X. Vilajosana, P. Tuset-Peiro, B. Martinez, J. Melia-Segui, and T. Watteyne, "Understanding the limits of LoRaWAN," *IEEE Communications magazine*, 55(9), pp.34-40, 2017.

- [16] F.T. Johnsen, Z. Zieliński, K. Wrona, N. Suri, C. Fuchs, M. Pradhan, J. Furtak et al. “Application of IoT in military operations in a smart city,” In 2018 International Conference on Military Communications and Information Systems (ICMCIS), pp. 1-8. IEEE, 2018.
- [17] L. Sadler, J. Michaelis, S. Metu, R. Winkler, N. Suri, A. Raj, and M. Tortonesi, “A distributed value of information (VoI)-based approach for mission-adaptive context-aware information management and presentation,” Technical Report No. ARL-TR-7674. US Army Research Laboratory Adelphi United States, 2016.
- [18] Y. Gai, B. Krishnamachari and R. Jain, “Learning multiuser channel allocations in cognitive radio networks: a combinatorial multi-armed bandit formulation,” in 2010 IEEE Symposium on New Frontiers in Dynamic Spectrum (DySPAN), pp. 1-9, 2019.
- [19] J.R. Michaelis, M. Tortonesi, M. Baker, and N. Suri, “Applying semantics-aware services for military IoT infrastructures,” in Proceedings of the 2016 International C2 Research and Technology Symposium (ICCRTS 2016), London, UK, 2016.
- [20] P.H. Deitz, J.R. Michaelis, B.E. Bray, and M.A. Kolodny, “The missions & means framework ontology: matching military assets to mission objectives,” in Proceedings of the 2016 International C2 Research and Technology Symposium (ICCRTS 2016), London, UK, 2016.
- [21] I. Agadakos, G.F. Ciocarlie, B. Copos, J. George, N. Leslie, and J. Michaelis, “Security for resilient IoBT systems: emerging research directions,” Proceedings of Workshop for Internet of Things in Adversarial Environments (co-located with 2019 IEEE INFOCOM Conference), Paris, France, 2019.
- [22] A. Raglin, “Presentation of information uncertainty from IoBT for military decision making,” In International Conference on Human-Computer Interaction (pp. 39-47). Springer, 2019.

Storytelling Exploratory Visual Analytics for Counter-Improvised Explosive Device Incidents

Valérie Lavigne

Defence R&D Canada
2459, route de la Bravoure,
Québec (QC)
CANADA

valerie.lavigne@drdc-rddc.gc.ca

Shivas Jayaram

BrainCradle
322 Birchwood Dr
Hartsville, SC 29550
UNITED STATES

shivasj@gmail.com

Marius Panga

Microsoft
Microsoft Campus
Thames Valley Park, Reading
UNITED KINGDOM

marius.c.panga@gmail.com

ABSTRACT

In order to better understand the dynamics of Counter Improvised Explosive Devices (C-IED), we developed an exploratory visual analytics tool coupled with storytelling techniques. Analyzing datasets about incidents can help derive activity level assessment, perform trend analysis and gain better situation awareness.

Exploratory visual analytics allows a user to explore a dataset and discover interesting patterns. Visual analytics tools can offer complex interactions, but they may be intimidating to a new user. Storytelling techniques has the potential to help new users get started with an interactive visualization application.

The use of our exploratory visual analytic tool allowed the discovery of interesting patterns in the C-IED incidents dataset produced by the NATO Counter-IED Center of Excellence (NATO C-IED COE), and we conveyed those insights to users by applying storytelling techniques as part of the interactive visualization tool. After the users have explored the provided insights, they are invited to continue exploration on their own, allowing them to make more discoveries.

This paper explains the exploratory visual analytics capabilities implemented into the tool, as well as how storytelling techniques were applied to convey insights to a user and subtly teach them how to use the tool. Then, we provide an overview of the insights discovered with the tool about C-IED events.

INTRODUCTION

We developed an exploratory visual analytics application that was coupled with storytelling techniques.

The collection and analysis of datasets about incidents can help analysts derive activity level assessment, perform trend analysis and gain greater insight about the problem they are monitoring. In order to better understand the dynamics of Counter Improvised Explosive Devices (C-IED), we decided to apply an exploratory visual analytics approach with storytelling elements.

Exploratory visual analytics is meant to allow a user to explore a dataset and discover interesting patterns and insights. Some visual analytics tools afford complex interactions that can be intimidating to a new user. Storytelling techniques can help a new user get started with a new interactive visualization application in order to understand what a dataset contains and how to leverage the exploratory visual analytics' tool capabilities to perform new analyses on the dataset.

We found that the use of an exploratory visual analytic tool allowed the discovery of interesting patterns in the Ukraine C-IED incidents dataset produced by the NATO Counter-IED Center of Excellence (NATO C-IED COE), and we conveyed those insights to users by applying storytelling techniques as part of the interactive visualization tool. Once, the users have started exploring the provided insights, they are invited to continue their exploration allowing them to make additional discover from the data.

After presenting the related work, this paper explains the exploratory visual analytics capabilities implemented into the tool, as well as how storytelling techniques were applied to convey insights to a user and subtly teach them how to use the tool. Then, we provide an introduction to the C-IED events analysis and the insights discovered with the tool are presented for this application domain.

DATASETS DESCRIPTIONS

NATO Ukraine IED Incidents Data

This is a NATO Unclassified IED events spreadsheet from the NATO C-IED COE. It contains 665 events, with 15 in 2001-2013, about 230 in 2014 and the rest in 2015. It contains the following column headers: Date, Type, KIA (Killed in Action), WIA (Wounded in Action), City, Region, Country, Details, and Remarks. There are missing values in the dataset.

Ukraine Census Data

We also included various statistics about Ukraine to see if we can find patterns between the IED events and these statistics. The data was cleansed and transformed in order to have use the same common attributes as our primary data set (the NATO IED dataset). Transformations applied in order to facilitate the data merge with the other datasets: removed irrelevant data from the data set, converted all fields to the correct type, and translated the region names to the ones used by the primary data set.

EXPLORATORY VISUAL ANALYTICS STORYTELLING TOOL

Design Goal and Approach

This project's goal was to allow the exploration of over 600 IED incidents in Ukraine mostly over the past 2 years for the purpose of highlighting and better understanding the temporal, geographical and political patterns in that data. We also considered additional census data and election statistics to uncover potential regional patterns, as well as publicly available data about the existing conflict in Eastern Ukraine. Our intent was to employ interactive visualization to generate better insights about the Ukraine IED situation. We documented how we gathered and transformed the data, our interactive visualization design process and inspirations, and user feedback evaluations in our process book [1].

In order to drill down into different aspects of the dataset, we decided to split it in an overview introducing the purpose of the tool (why) and leading to three specialized views aiming at answering the where, what and how questions about IED incidents.

We repeated the timeline banner and the storytelling boxes at the top of each view, and reused the same casualties icons across the visualizations as a way to keep consistency and help user orientation in the tool. Each

perspective offers multiple coordinated views that can be filtered temporally with using the interactive timeline.

Dataset Overview

The dataset overview approach aligns with the Visual Information-Seeking Mantra of “Overview first, zoom and filter, then details-on-demand” [2]. It introduces the purpose of the visual analytics tool and provides links to the other perspectives (Figure 1).

Geospatial View

The geospatial view shown in Figure 2 provides insights into the geo-temporal aspects of IED incidents. The user can drill down into each region to see how IED incidents happen over time down to city level. The map display shows of individual incidents are spread over each area and the incidents dots can be coloured according to their type or effect on casualties. We also included additional statistical data about Ukraine that allows the regions to be coloured by population density or Russian/Ukrainian ethnicity in addition to the incidents frequency data.

Incidents Type View

In Figure 3, the incidents types and properties views provide a Sankey diagram that focuses on the incidents type and casualties. Brushing over incidents types with the mouse displays examples of incident of that type to the right of the diagram.

Text Analysis View

The text analysis in Figure 4 view exploits Natural Language Processing (NLP) to analyze the unstructured text descriptions of IED incidents. In the graph, events that share similar words are linked together and the most prominent words in each group are displayed.

Storytelling Exploratory Visual Analytics for Counter-IED Incidents

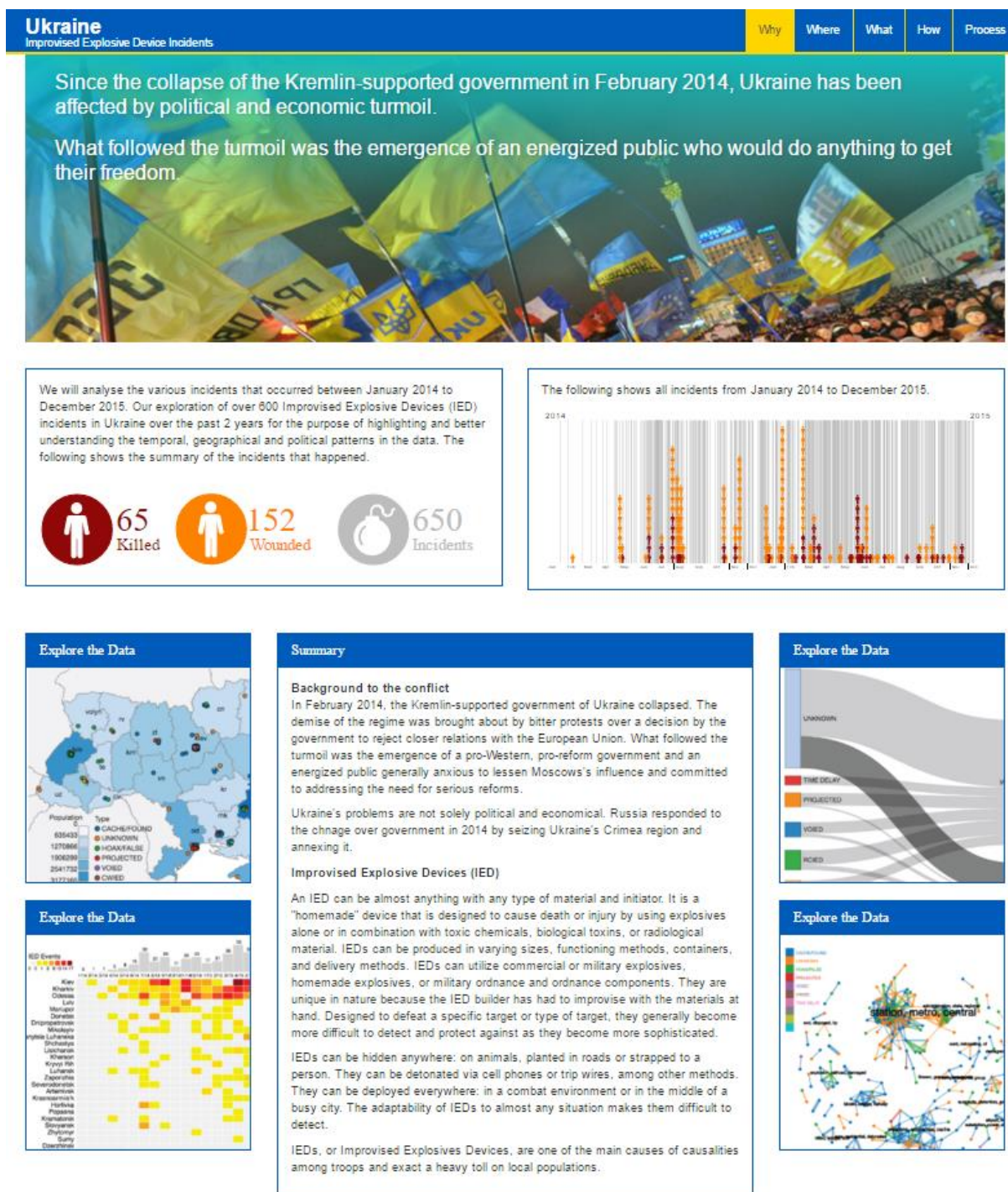


Figure 1: C-IED events analysis tool introduction view.

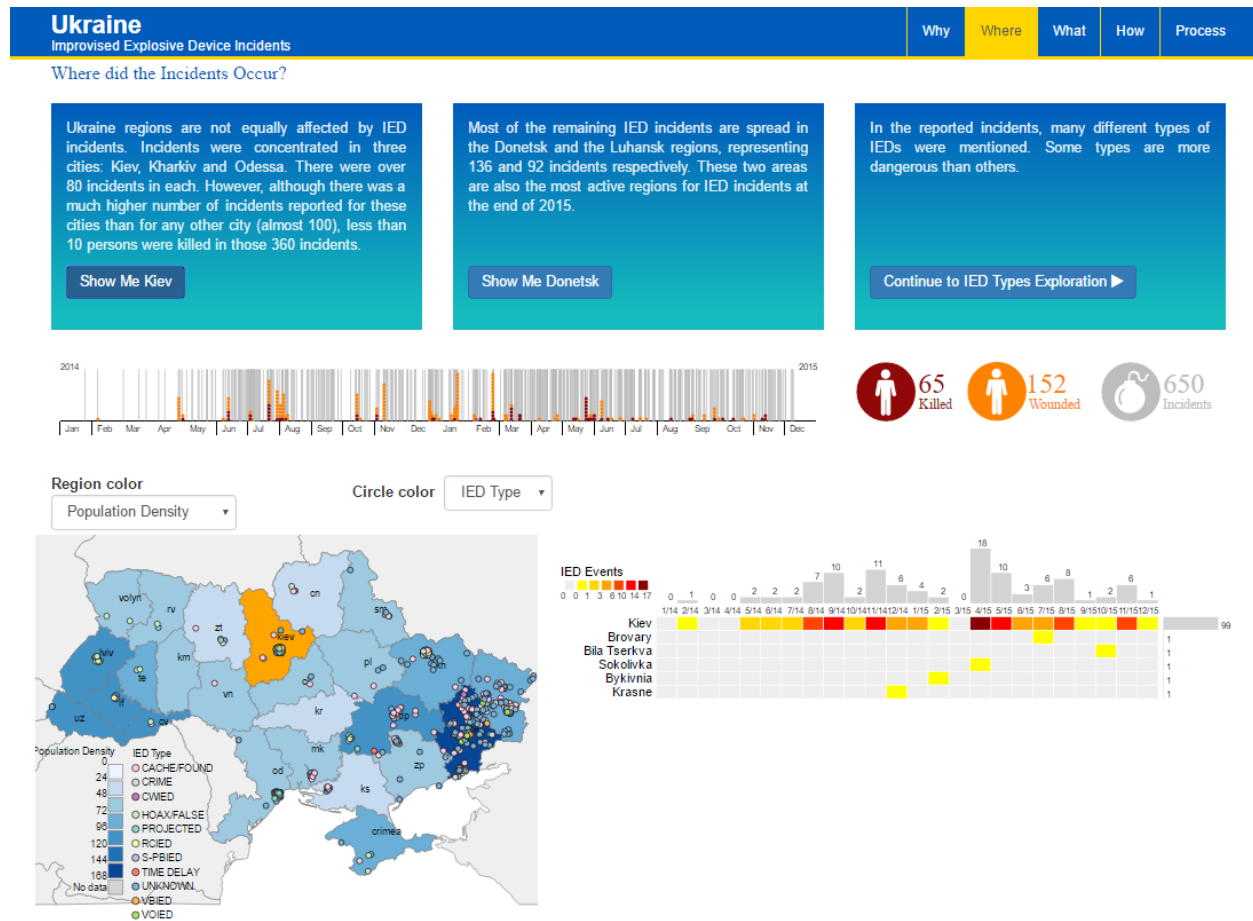


Figure 2: C-IED events analysis tool geospatial view.

Storytelling Exploratory Visual Analytics for Counter-IED Incidents

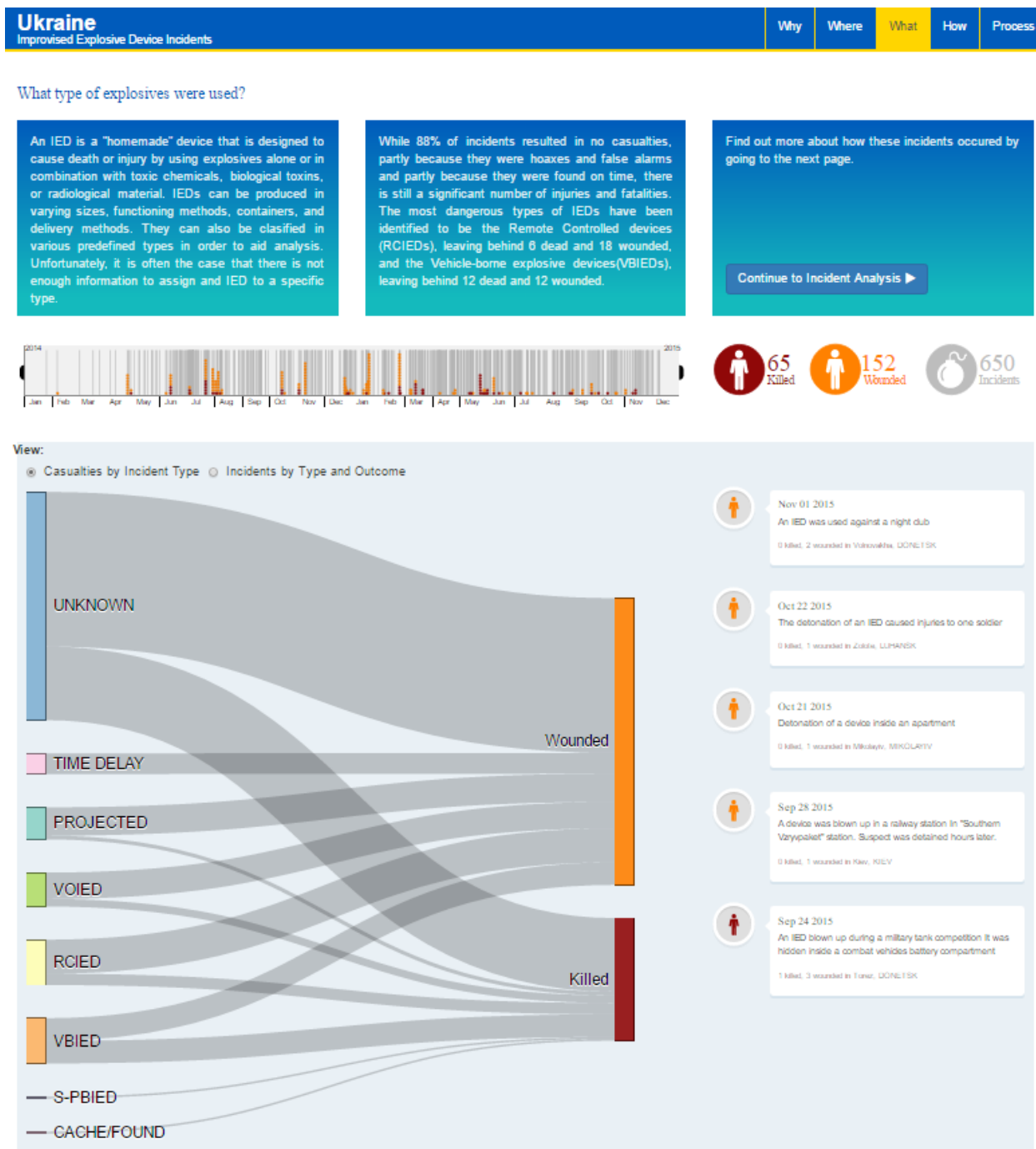


Figure 3: C-IED events analysis tool incident view.

Ukraine Improvised Explosive Device Incidents

Why Where What How Process

How these Incidents Occurred? - Interpreting the Free Text Reports

In this exploration we want to analyze the free text summary written for each incident. For the process of finding report similarity, we first used the TF-IDF and Cosine Similarity algorithms to find similarity between all the incident reports. Then for our analysis we consider two reports to be similar only if their threshold is 50% or higher. To gather important words across report clusters, we first find the three most important words in each of the reports. This was found using TF-IDF (Term Frequency, Inverse Document Frequency) scores. Then for each cluster of nodes we gather all the important words from each node and take the top three scored word. The font size of the words is determined by the frequency of occurrence in the cluster.

By clustering the incident reports in this method we can see that most similar incidents that were reported in metro stations, administration buildings were hoax or false reports which led to no casualties.

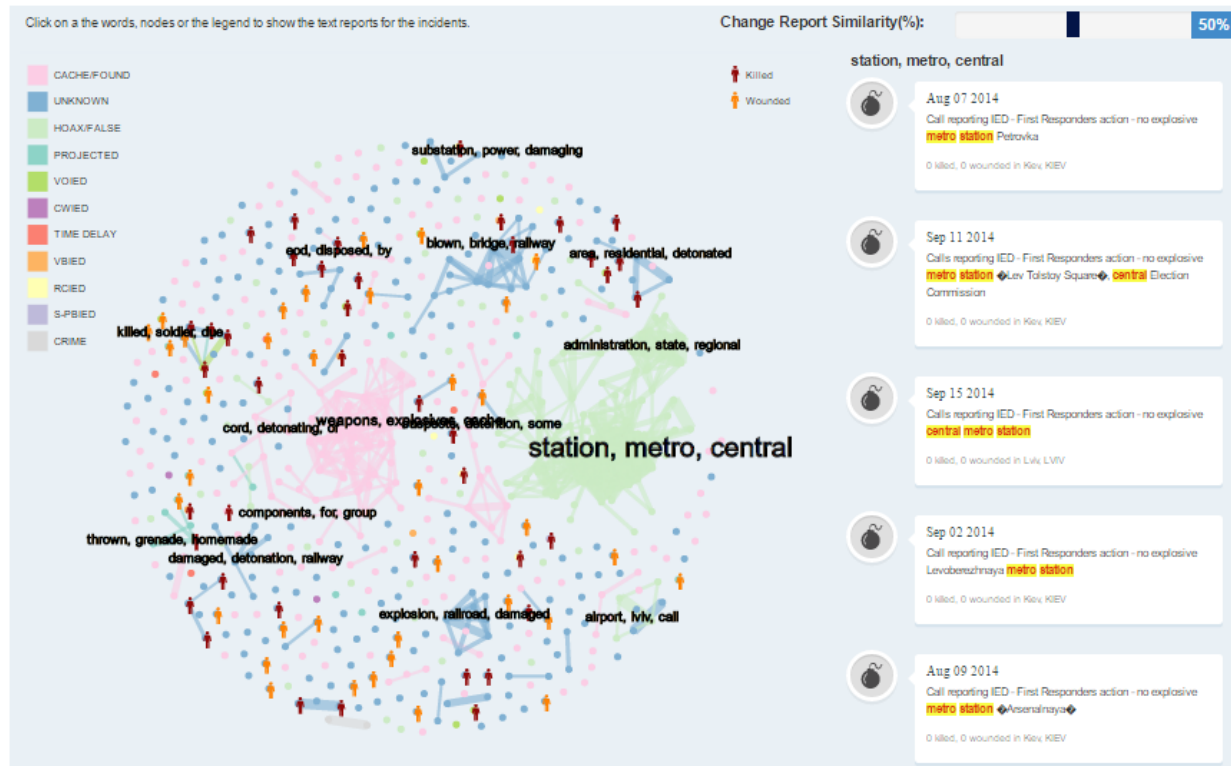
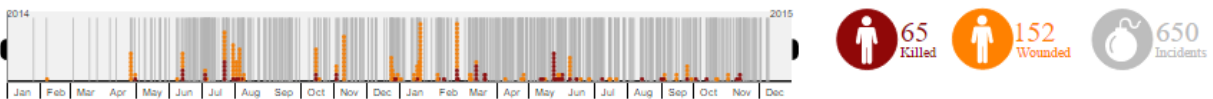


Figure 4: C-IED events analysis tool text analysis view.

STORYTELLING TECHNIQUES

Segel and Heer [3] performed a design space analysis of narrative visualizations. According to this classification, our C-IED events analysis tool uses the following three techniques, to balance between an author-driven and a reader-driven perspective on the dataset:

- Drill-Down Story,
- Martini Glass Structure, and
- Interactive Slideshow.

Drill-Down Story

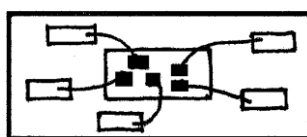


Figure 5: Drill-Down Story [3]

“The Drill-Down Story visualization structure presents a general theme and then allows the user to choose among particular instances of that theme to reveal additional details and backstories.” [3]

In our tool, this is what happens in the overview page (Figure 1), which offers a description of the C-IED context and the incident dataset surrounded by small frames that link to more specialized views to explore this dataset.

Martini Glass Structure

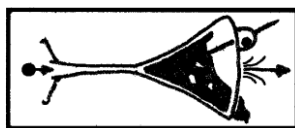


Figure 6: Martini Glass Structure [3]

“The Martini Glass visualization structure begins with an author-driven approach, initially using questions, observations, or written articles to introduce the visualization.” [3]

The Martini Glass Structure is very efficient for showing the user how to use the different features of the interactive visualization tool. The top part of Figures 2 contains boxes that offer a text providing insights linked to this view of the data and when the user clicks the corresponding “Show Me”, the tool automatically renders the appropriate values and settings to show the insight described with the data.

Interactive Slideshow

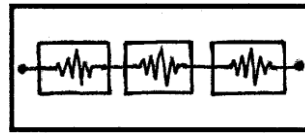


Figure 7: Interactive Slideshow [3]

“The Interactive Slideshow structure follows a typical slideshow format, but incorporates interaction mid-narrative within the confines of each slide. This structure allows the user to further explore particular points of the presentation before moving ahead to the next stage of the story.” [3]

This pattern appears at the top part of Figures 2, 3, and 4, where the blue boxes discuss particular aspects of the data that stand out from the current view, before offering to continue to the next view. This technique also facilitates navigation between the different views of the tool, along with the top menu.

INSIGHTS ABOUT C-IED INCIDENTS

The dataset about the Ukraine IED situation was produced by the NATO C-IED COE. The classic way to convey the information from the C-IED dataset employed by the CIED COE is through static visualizations and tables that can be incorporated into MS PowerPoint slides, as can be seen from Figures 8 and 9. Those graphics do not allow an in depth analysis of the dataset and it can be argued that it is not their purpose. Still, it is interesting to explore what are the additional insights that can be extracted from a storytelling visual analytics tool.

Figure 8 makes it obvious which regions have the highest number of IED incidents and provides the total number of incidents for the country. We can see that Kiev(Kyiv) region has a lot of incidents. This is also confirmed in the geospatial view from Figure 10. However, if we choose to color the regions according to the lethality of IED incidents, we quickly realize that incidents that happened in Luhansk and Donetsk were more dangerous than the incidents in Kiev and Odesa. The incidents dots also show how the incidents are spread in each region, showing that incidents in Kiev and Odesa are grouped mostly in the cities of the same name, whereas incidents in Luhansk and Donetsk are spread throughout the area.

In Figure 9, we can see over time how many incidents happened, and how many persons were either wounded or killed. A pie chart of incident types and a table of events by target type are provided. The interactive timeline banner at the top of each view (shown in Figure 12) also provides the information about incidents over time, but using a visual representation. The Sankey diagram shows the relative prevalence of each type of incidents, but adds the information about which ones caused higher casualties. And if we turn the focus around on casualties, we see in Figure 14 the effects of each type of IED. Note that Figure 9 shows data from 2013, which was not available to us, while Figure 13 and 14 are based on 2014/2015 data. Using the interactive timeline filter, we can select a sliding window of a few months and discover that hoaxes incidents increase over time. Also, the relative portion of incidents that killed people is increasingly of unknown type.

Finally, the details of IED incidents are not exploited at all in Figure 8 and 9. Natural language processing allowed us to explore the incidents text descriptions to find out that most hoaxes and false IEDs involve the

words “station, metro, central”.

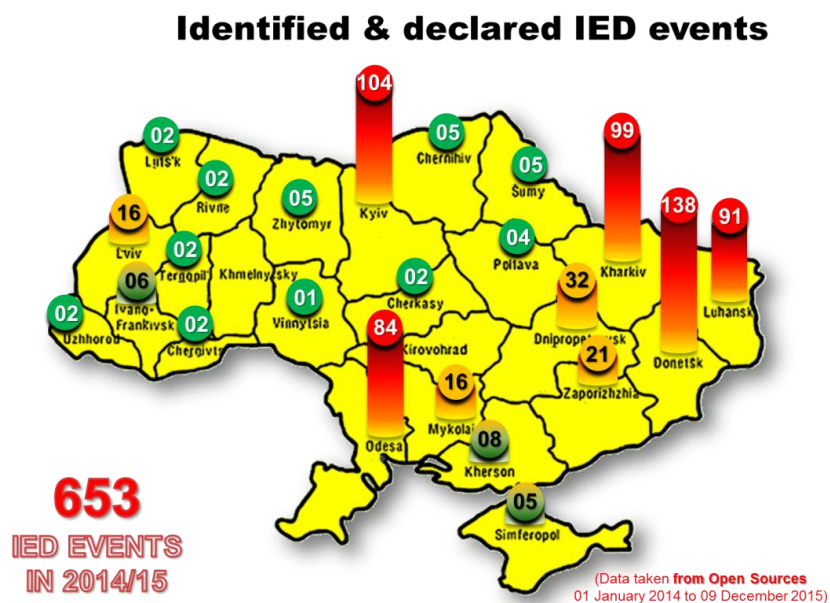


Figure 8: Geospatial summary slide featuring a map of Ukraine IED incidents in 2014- 2015.

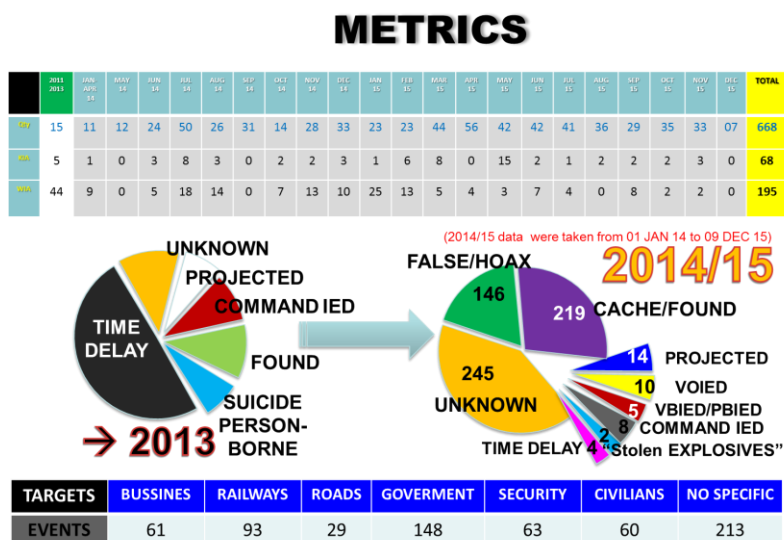


Figure 9: Categorical and temporal summary slides featuring statistical data about Ukraine IED incidents in 2014-2015.

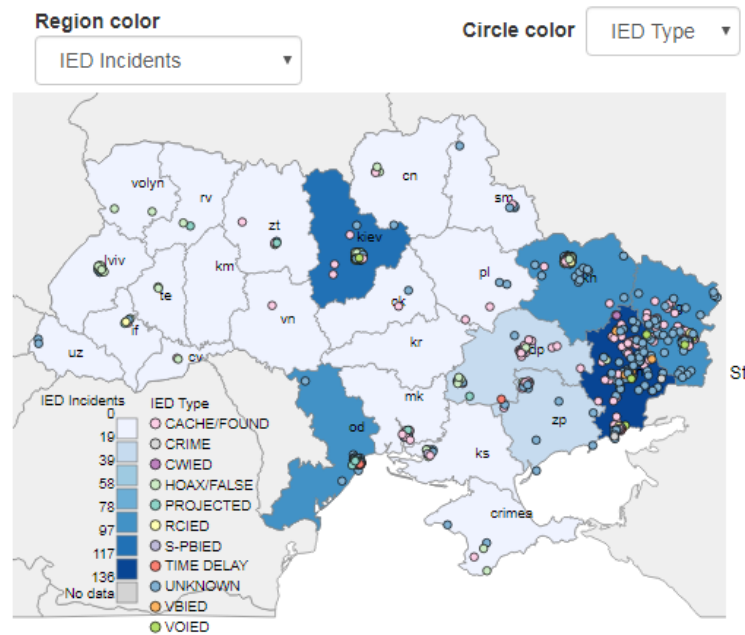


Figure 10: Regions colored according to level of IED incidents.

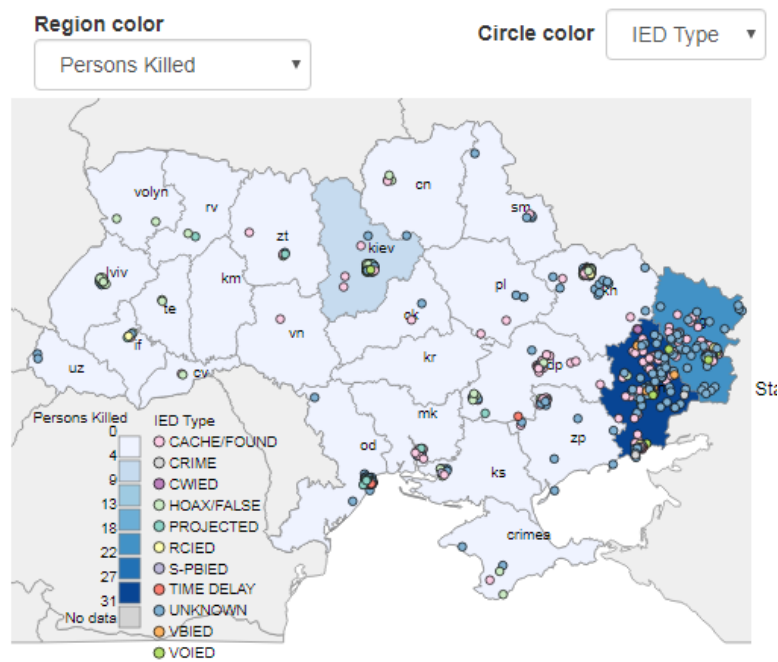


Figure 11: Regions colored according to lethality of IED incidents.

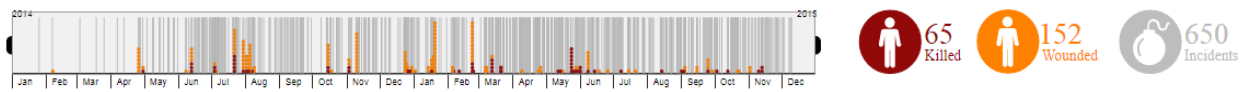


Figure 12: Interactive timeline banner and filter at the top of each view.

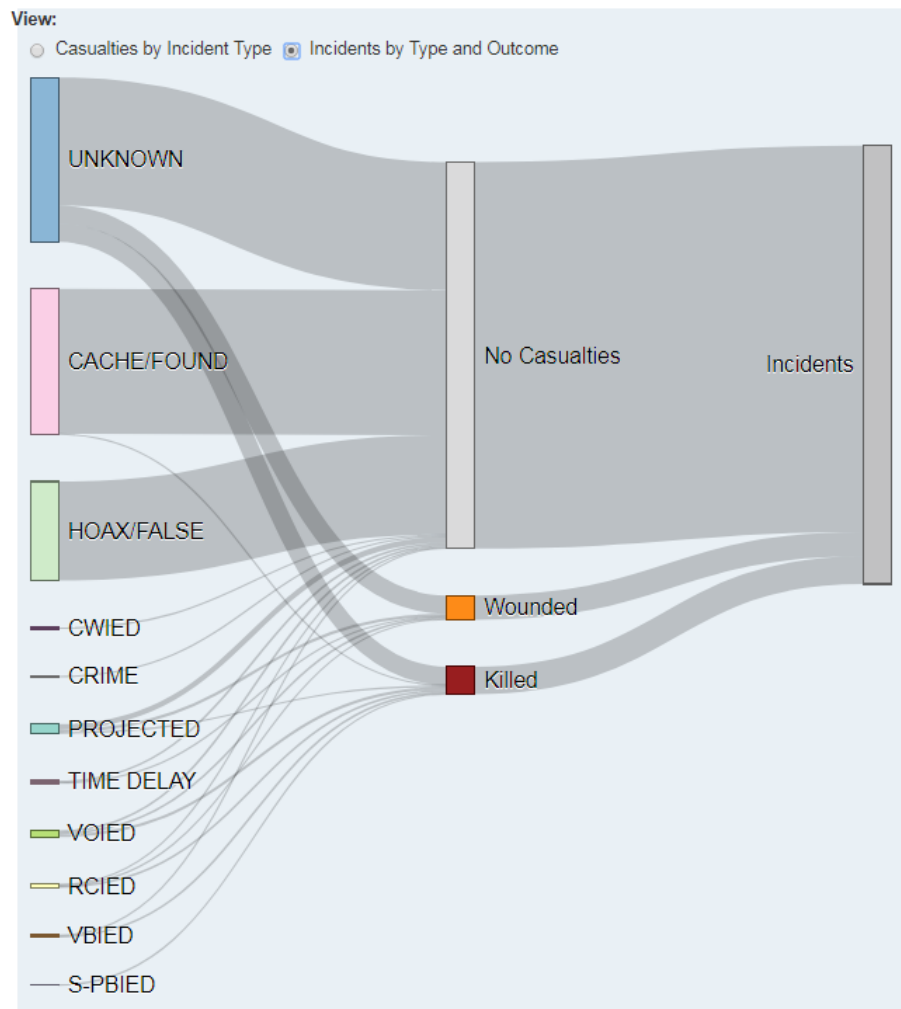


Figure 13: Sankey diagram showing incidents by type and outcome.

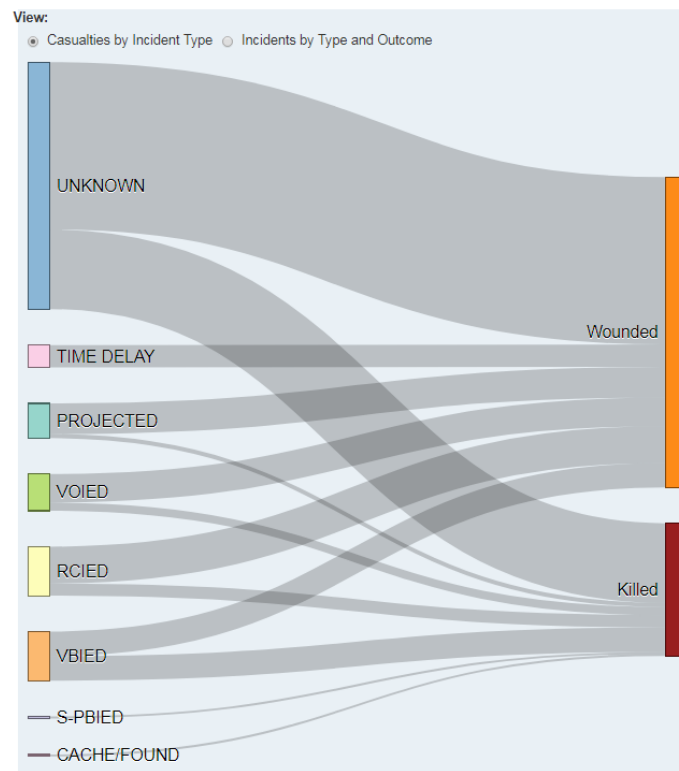


Figure 14: Sankey diagram showing casualties by incident type.

CONCLUSION

In this paper, we described the design approach and storytelling elements behind the exploratory visual analytics tool we developed for analyzing a dataset about IED incidents in Ukraine. We demonstrated that a visual analytics approach can highlight insights that cannot be discovered from summary tables and the basic visualizations provided in presentation slides. The tool can be web based and accessible at (<http://nato-project.github.io/v2/index.html>). The code for the tool is also available freely.

REFERENCES

- [1] Lavigne, V., Panga, M. and Shivas, J. (2016). Process Book (CS171 Week 13). Project Title: Ukraine Improvised Explosive Devices. [http://nato-project.github.io/v2/Process%20 Book.pdf](http://nato-project.github.io/v2/Process%20Book.pdf). Accessed on September 9, 2019.
- [2] Shneiderman, B. (1996). The eyes have it: a task by data type taxonomy for information visualizations. Proceedings 1996 IEEE Symposium on Visual Languages, 336-343.
- [3] Segel, E., & Heer, J. (2010). Narrative Visualization: Telling Stories with Data. IEEE Transactions on Visualization and Computer Graphics, 16, 1139-1148.



A Preliminary Experimentation for Meteorological Data Visual Analytics

Pisano Francesca Maria, Cembalo Assuntina, Romano Gianpaolo

Italian Aerospace Research Centre,
via Maiorise, 81043 Capua (CE)
ITALY

f.pisano@cira.it, a.cembalo@cira.it, g.romano@cira.it

ABSTRACT

Everyday sensors and detectors register huge quantities of geo-referenced meteorological data up to 40-45 GB a day, available in different formats (.csv, .nc, .grb, .xlsx, .tiff, .shp, etc.), at different frequencies; that data is to be appropriately analyzed to extract information as support in critical domains such as aviation. Meteorologists have expressed analysis requirements referring to visualization and graphical representations of data as univariate or multivariate spatial distributions of physical observations and simulated variables. Meteorologists can potentially gain advantage in adopting the innovative approach of Visual Analytics based on interactive and visual data representations to identify spatio-temporal trends and correlations. In the present work, a preliminary Visual Analytics solution, combining Tableau and in-memory database Exasol, is designed and implemented to offer worksheets e dashboards, whose visualization requires few seconds versus hours and days needed to generate traditional meteorological maps and graphs. Satisfactory experimental results are shown for COSMO (Consortium for Small-scale Modeling) data and ECMWF (European Center Medium Weather Forecast) SYNOP (Surface synoptic observations) data for a total volume of almost 1 GB and more than 21 millions of records.

1.0 INTRODUCTION

Weather information is nowadays extremely relevant for safety in a number of domains, because it can effectively help to avoid incidents and mitigate consequences related to extreme weather events. Both infrastructures and transportation systems can benefit from weather awareness in order to manage adverse weather conditions. The European Union Aviation Safety Agency has a Weather Information to Pilots project team [1] that “has reviewed the existing means to mitigate against the effects of the weather, including the use of on-board weather radar, other on-board sensors, the information provided to pilots pre-flight, and the information that is available in-flight”. In the U.S.A., the Aviation Weather Center delivers consistent, timely and accurate weather information for the world airspace system [2].

Meteorology is a complex scientific area, making use of both observation and prediction data, covering large geographical areas, with updating period from 10 seconds to 1 day. The visual approach is usually adopted by meteorologists, who are used to work with graphs and maps to investigate and communicate complex phenomena in a synthetic and integrated manner [3], nevertheless they experience the limit of depicting static information.

Numerical Weather Prediction systems can take advantage from the remarkable progressive growth registered in recent years by the hardware computational power [4], that allows the storage of huge amount of data in various domains. Unfortunately, existing analytical systems can meet difficulties in managing such increasing quantity of data, so innovative solutions become necessary to adequately exploit all the available information coming from different sources, in different formats, at different frequencies: the so called *Big Data* [5]. Visual Analytics (VA) can effectively helps as it is the *Science of analytical reasoning facilitated by interactive visual interfaces* [6]. VA aims at stimulating the human cognitive abilities, in identifying

patterns and detecting relationships in visualizations, in combination with the data management and analysis methods and technologies. Meteorology can just gain great benefits from the Big Data Visual Analytics [7], providing the breakthrough technical solutions to be applied in order to extract relevant information for experts and end-users from new datasets.

In the present work, an experimental approach to manage huge quantity of meteorological data is proposed using unconventional storage solution like a NewSQL database [8] and a Visual Analytics tool in order to satisfy meteorologist requirements with respect to the interactivity and the time performance of analytics.

2.0 RATIONALE

Traditional Information Visualization objectives [9] are about the effective representation of analytical activity results or the confirmation of domain expert hypothesis. Towards the end of 90s, a group of visionary researchers at the Pacific Northwest National Laboratory [10] detected the gap between the existing Information Visualization capabilities and the emergent demanding analytics requirements. After 09/11, growing national security and safety needs gave birth to projects dedicated to the elaboration of innovative visual approach to analytics worldwide ([11], [12], [13], [14]). Visual Analytics pushes forward the application of visualization techniques towards the creation of analytical models supporting information exploration and extraction by means of graphs and maps. Visual Analytics is a multidisciplinary research area ([15], [16], [17]) integrating Data Management and Analysis, Statistics, Spatio-temporal Data Processing, Visualization and Human-machine Interaction, which aims at passing over the simple representation of available data, facilitating meaningful interactions so that the end-user experiences innovative information understanding tools. In this sense, the right choice of the graphs and maps and of the interactive solutions has to be strictly related to user needs. Visual Analytics is able to effectively manage heterogeneous data formats by means of different representations; spatial and time data are especially suitable to visual depiction. VA mainly pursues the tasks of Exploration and Analysis: the former includes data selection, pattern and trend identification, outlier detection; the latter faces issues as visual support to analytical algorithms and models evaluation and the effective results representation. Therefore, VA can expand users' explorative analytical experience and possible creative processes. A particularly relevant aspect of Visual Analytics is about the communication of the obtained analytical results; methods and tools supporting the data storytelling [18] are necessary to improve domain expert abilities to disseminate the outcomes of their work.

Visual Analytics is applicable to issues featured by huge volume of information, high variety, heterogeneous sources and large number of parameters; therefore it seems to be desirable for domains as Climatology and Meteorology, which are featured by high quantity of data, daily collected, to be compared over long time periods. In recent years, some applications of Visual Analytics to the meteorological domain have been elaborated to support situation and weather awareness. The visualization solution named "VIDa" (Visual Interactive Dashboard) is proposed in [19]; multiple views of time sequences, spatial filters and maps support end-users seeking spatio-temporal patterns for short-term forecasts. The MEVA (Multifaceted Environmental Data Visualization Application) [20] offers interactive visual interfaces for meteorological observation and simulated data to users without advanced skills in 3-D visualization systems. The Ship and Weather Information Monitoring decision support system [21] aims at specifically applying the visual approach to weather awareness for naval transportation system. Naval fleet routes are elaborated taking into account timely meteorological conditions; in order to do that, data from ECMWF and updated route information are integrated and alternative routes can be depicted by means of interactive visual interfaces.

3.0 DESCRIPTION

In the present work, an experimental approach to manage huge quantity of meteorological data is proposed using unconventional storage solutions like in-memory analytic database [8] and a Visual Analytics tool in

order to satisfy meteorologist requirements with respect to the interactivity and the time performance of analytics. A brief introduction to specialist requirements is provided together with the description of the data sources selected for the experimentation. Then the solution architectural design and the adequate tools selection on the basis of expert needs are discussed.

3.1 Domain requirements

In collaboration with the Italian Aerospace Research Centre Meteorology Laboratory, expert requirements have been gathered; they are firstly about the exploration of the spatial distribution of one or more physical and numerical variables. Typically, meteorologists work with graphs and maps to depict available information in a synthetic and integrated approach; the power of visual representation is limited by the static nature of images. Indeed, specialists recursively elaborate graphs to gain exhaustive insight into data. Experts ask for dynamically and interactively selecting the geographical region and the variables values of interest, and they request to download the obtained graphs and maps as the most widespread image formats. Moreover meteorologists need plots representing observations and simulated data as scatterplots and temporal series to seek for correlations and trends; often mean values and other statistical indices of collected registrations are to be shown. Specialists also require faster analytical tools enabling the elaboration of graphs and maps dynamically, following the expert own needs during the user exploration and analysis activities. All the operations such as processing, analysis and visualization are strongly time-consuming when recurring to traditional analytics methods and technologies.

3.2 Meteorological data

The activity of requirements analysis is joined to the analysis of data sources, which have been selected by meteorologists as a representative specimen of relevant information to process. Different kinds of data have been provided, including observations and numerical datasets of different variables.

A brief description follows:

- **SYNOPSIS data:** observed data stored in the MARS (Meteorological Archival and Retrieval System) [22] repository managed by the ECMWF (European Center Medium Weather Forecast) [23]. Registrations are collected hourly and include one alphanumeric code to represent date, geographical area, significant phenomena occurrence, etc. and 29 fields (station coordinates, temperatures, wind speed, pressures, etc.). Records are gathered by 877 monitoring stations, positioned in Europe and Northern Africa. The dataset includes 4 months of registration in 2016, for a total amount of more than 2 millions of records and 250 MB volume.
- **COSMO data:** simulated data from the Consortium for Small-scale Modeling model [24]. Data represents temperature and rainfall registered by 51520 monitoring stations in Central Europe during the year 2016. It amounts to about 19 millions of records and 700 MB.

3.3 Analysis and design

Architectural solution has been derived from the domain expert needs and data sources analysis. Such an integrated design approach drives to a solution able to take into account expert and end-users requirements combined with the effective available information. In the following, the technical issues to be faced and the proposed solution in terms of architectural components and related tools are described.

3.3.1 Architectural design

Analytics requirements from meteorologists are focused on demanding data visualization and representation needs with respect to spatial and statistical distributions of one or more variables in order to seek temporal trends and possible correlations. As mentioned before, the possibility of storing huge quantity of data, both

for physical observations and numerical forecasts, covering long time periods, represents an outstanding opportunity to elaborate more accurate forecasting models for meteorologists. The complexity of analytics activities is often faced using graphs and maps to catch synthetic and integrated view of data, eventually for subsequent more detailed analysis.

In this context, Visual Analytics offers unheard-of interactive visualization solutions, enabling a rich and customizable experience of data exploration, that can effectively empower expert activities. Indeed, *Visual exploration* can support the information interpretation process in a user-friendly manner, using geographical maps, which easily represent multivariate problems in two and three-dimensions. Points stand for single observations corresponding to positions in a specified reference system; rectangles represent observation groupings whose extension is related to the number of included points; lines link observations and polygons can aggregate rectangles. All those geometric elements can be combined to create graphs, that can in turn form analytical interactive dashboards. The selection of specific elements and graphs has to be driven by the expert needs. Geographical maps represent positions of interest, related to monitoring stations or sensor coordinates, and use shapes, dimensions and colors to depict one or more variables values. In addition, operators as sum, average, maximum, minimum can be displayed and more textual information is provided passing the cursor over the map. Filters, highlights and dynamic animations facilitate the data exploration experience. Map portion selection is performable to focus on a reduced geographical region of interest, synchronizing statistical indices and graphs.

The temporal trend of a variable can be represented by means of linear graphs that can be aggregating with respect to the time dimension (year, month, day, hour, etc.) updating the aggregation operators (sum, average, etc.). More variables are added using side-by-side graphs, and/or using colors and shapes. Time is depicted using discrete or continuous approach. Box-plots to compare variable distributions for different time spans and scatterplots compare two variables for correlation identification; they can be enriched with information using colors and shapes again. The user takes advantage of the graph synchronization, so that selections are applied to all available visual representations on the screen.

On the other hand, the time performance challenging requirement fulfillment can be sought applying innovative solutions for the management of huge volume of data. Actually, COSMO simulated data amounts to about 700MB, not adequately managed by Visual Analytics solutions. It results in exceeding processing time that dramatically affects the user interactivity. In order to overcome such a drawback, *in-memory database* technology [8] is investigated to achieve faster data management, exploiting the computer central memory. In-memory databases belong to NewSQL database category [25], which assure as high performances as the NoSQL databases, keeping the well-known relational approach and SQL query language. Such databases effectively support the Volume, Velocity, and real-time Big Data analytical needs [26].

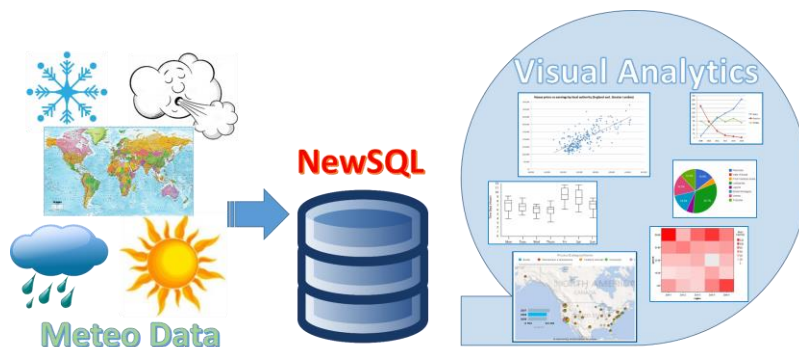


Figure 3-1: Architectural solution

One last issue is worthy to be faced: the solution release. As a matter of fact, it can be helpful for the end-users to have easily shareable analytical results at disposal, but this could be still an issue when high volumes of data are to be shared. A shareable solution including analytical results and data of interest should be investigated.

3.3.2 Tools

For the experimentation of the above described architectural solution (Figure 3-1), a technological benchmarking activity has been carried out. It has resulted into the selection of the free in-memory database Exasol [27] and the commercial Visual Analytics tool Tableau [28], which has native connection to Exasol database [29].

Exasol is a high-performance in-memory database, able to manage data much faster than traditional relational databases making use of mass memory. Exasol requires the installation of a Java Virtual Machine: Oracle Virtual Box has been exploited. Exasol can be used both from command line or by means of the SQL editor console *ExaPlus*.

Tableau Desktop is one of the most widespread Visual Analytics solutions worldwide. It is ranked as the second “Analytics and Business Intelligence Platform” for completeness of vision and ability to execute in the corresponding Gartner Magic Quadrant [30]. Tableau offers many graphs to be combined in effective worksheets including filters, highlights, legends and parameters and advanced features to create dashboards from them. Tableau lets save Visual Analytics applications as collection of worksheets and dashboards, optionally joining the associated data for release to users who do not have access to source data. Moreover, Tableau capabilities include *Self-service Data Preparation*, so user can create derived fields, dimensions and measures by themselves.

In the end, *Tableau Reader* [31] is a free software tool allowing the access and utilization of the visual analytical applications produced using Tableau Desktop also to users who do not own Tableau license.

4.0 RESULTS

COSMO data has been cleaned eliminating special characters and about 19 million of rainfall records have been stored in the Exasol database, using the ExaPlus editor, taking less than 40 seconds. The Tableau connection to Exasol allows to accelerate data visualization from minutes to few seconds. The most relevant advantage is about the interactive visual approach to data exploration and analysis; end-users can navigate worksheets and dashboards investigating temporal trends and comparing variables over geographical maps, dynamically selecting statistical indices and filtering values. In this way, experts can exploit the interpretation power of the visual approach and save hours of waiting for traditional graphs and maps elaboration and finally speed up the analysis work. Both for the COSMO and SYNOP data, interactive geographical maps have been created to represent the spatial distribution of meteorological variables statistical indices, aggregated on the selected time interval, using shapes like circles and squares to depict the value: bigger is the circle, higher is the index value. Filters are applicable to select a specific week or month of observation.

In Figure 4-1, the worksheet related to the animation of the COSMO monthly-based average rainfalls is represented. The legend on the upper right corner lets user select a specific month or to play the visual animation of the monthly average rainfalls sequence. The color legend indicates the depicted average values. The map appears as a solid shape because of the dense grid of monitoring stations; zooming in the background image, single points corresponding to monitoring station become visible.

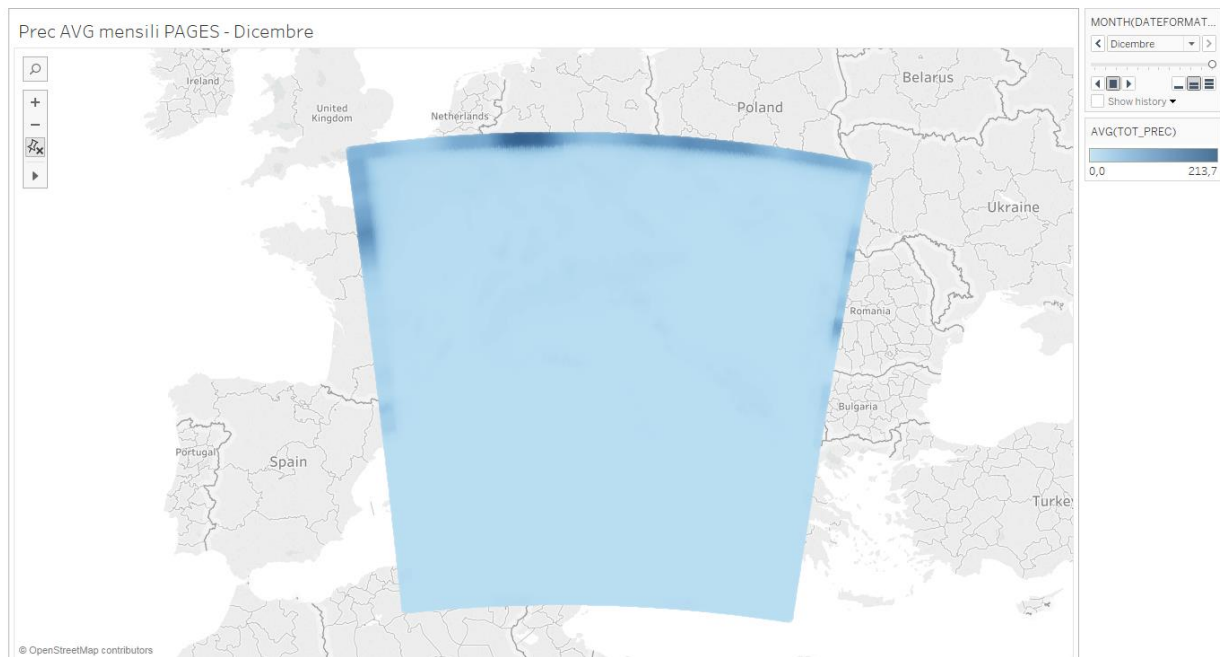


Figure 4-1: Animated sequence of simulated rainfall average value spatial distribution

In Figure 4-2 a SYNOP variable spatial distribution is shown; the circle area is proportional to the minimum value in a specific week/month of observation that user can select using the filter on the right.

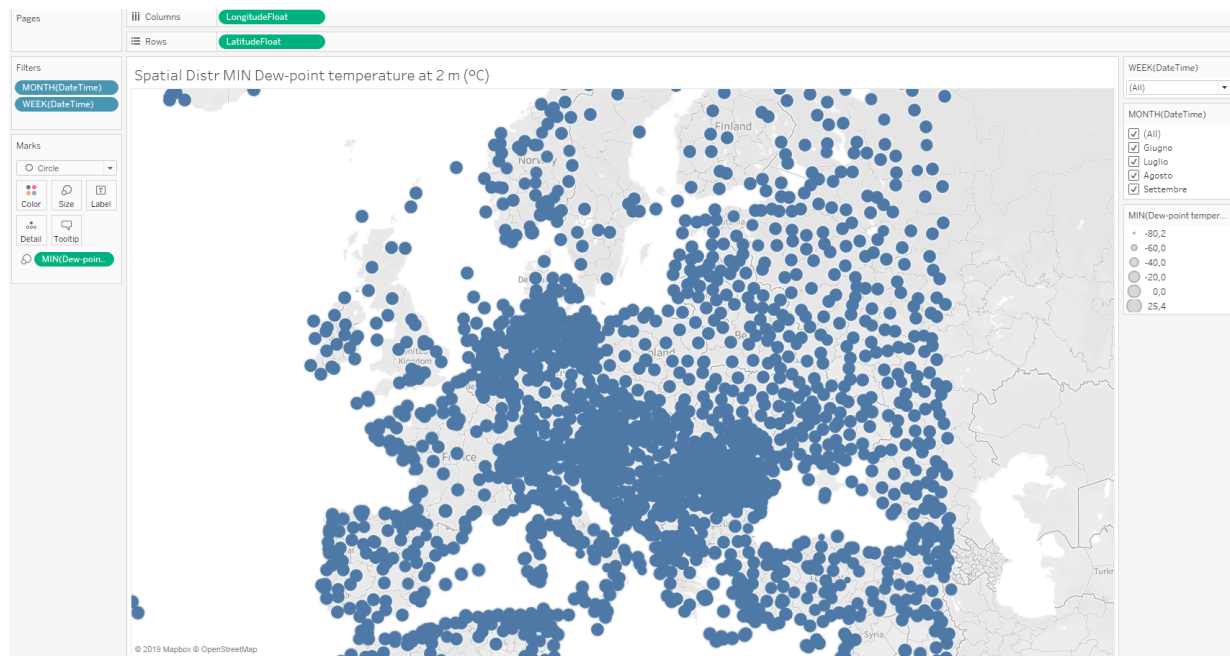


Figure 4-2: One variable spatial distribution with time filter

On the map in Figure 4-3, a second physical variable has been added to let user catch information about two observed dimensions at one glance. The first variable values are represented by the circle areas, while the second variable values are depicted by colors as indicated by the palette color legend. Interactivity is

empowered by filters over the latitude and longitude values, which allow the user to select geographical area of interest, and by animated sequence of monthly average value spatial distributions of the two variables.

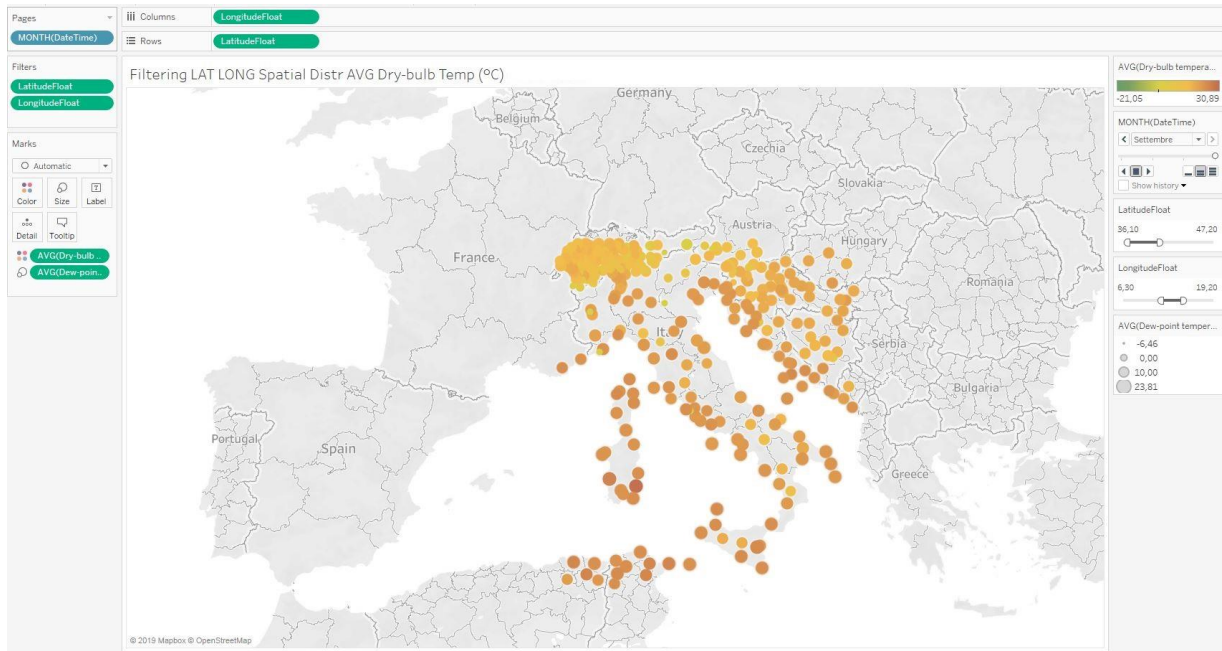


Figure 4-3: Animated sequence of two variable spatial distributions with spatial and time filters

Analytical dashboard can be built by suitably combining worksheets. Starting from a map (Figure 4-4), user can select a region of interest using the cursor and such a selection works as a filter to a second dashboard representing analytical graphs (Figure 4-5). In this way, user is supported exploring an increasing level of detail, because a visual selection is much simpler and more intuitive than numeric filters application.

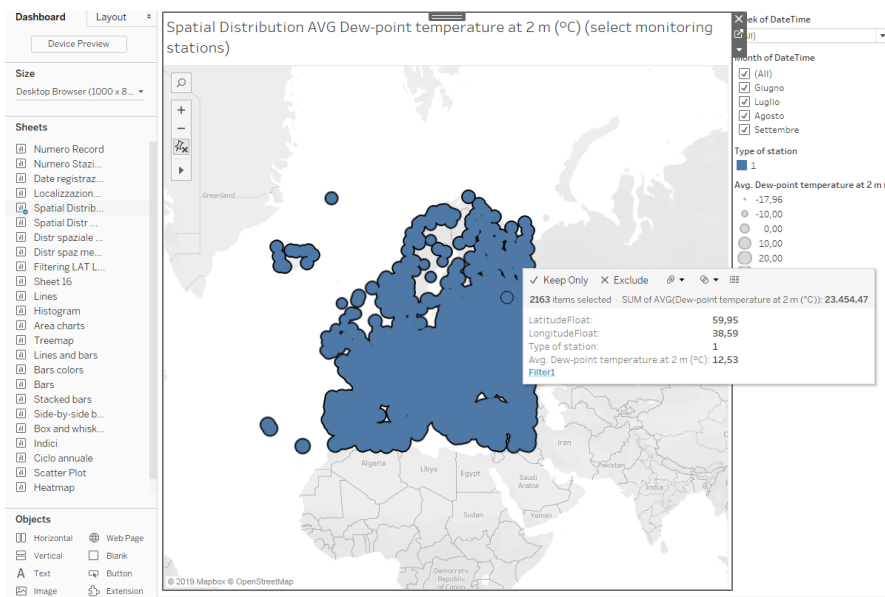


Figure 4-4: - Spatial filtering



Figure 4-5 – Filtered linear and box-plot graphs

Besides some easy worksheets are provided to support the initial exploration of data; they let user navigate the registration time spans and the spatial distribution of the monitoring stations.

Record date			
Year of Dat...	Quarter of ...	Month of D...	Day of Date...
2016	T2	Giugno	1
			2
			3
			4
			5
			6
			7
			8
			9
			10
			11
			12
			13
			14
			15
			16
			17
			18
			19
			20
			21
			22
			23
			24
			25
			26
			27
			28
			29
			30
		T3	Luglio
			1
			2
			3
			4
			5
			6

Figure 4-6: Records time span

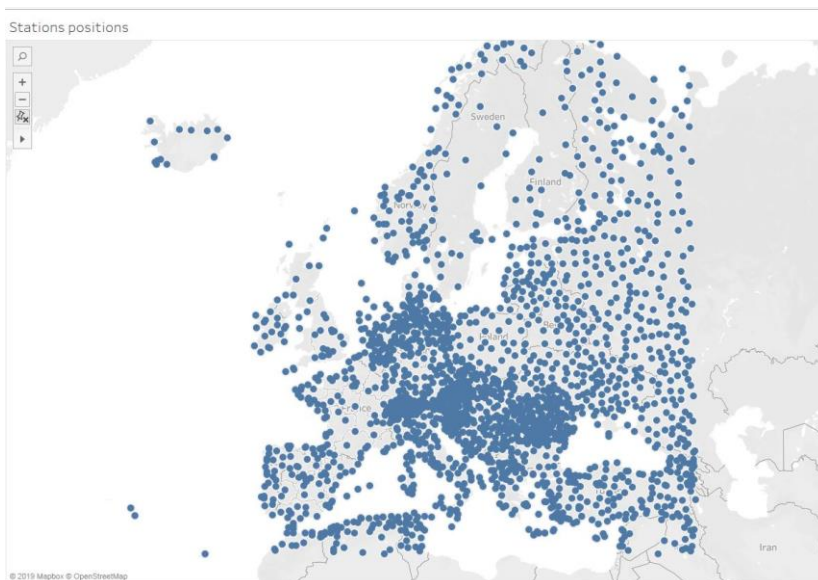


Figure 4-7: Spatial distribution of monitoring stations

5.0 CONCLUSIONS

In the present work, elements for a Visual Analytics solution design have been described aiming at assuring ease of use for the end-user so that the information understanding and perception process from huge amount of heterogeneous data can be significantly empowered by interactive analytical interfaces. Meteorologists requirements about interactivity and time performance have been met using graphs and maps and in turn combining them into worksheets and dashboards providing effective representations of available information and facilities for eventually subsequent more detailed analysis. Using technologies based on the data storage in the central memory instead of traditional mass memory, faster elaboration and visualization of high amount of data are achieved. Hopefully, future work will be focused on the integration of advanced analytical models in the Visual Analytics application, moving forward towards the empowerment of analytics.

REFERENCES

- [1] <https://www.easa.europa.eu/easa-and-you/air-operations/weather-information-pilots#group-easa-related-content>
- [2] <https://www.aviationweather.gov/>
- [3] Manzi P., Zollo A., Mercogliano P., Galdi C., “Aviation weather awareness: Development of algorithms for the detection of weather hazard through the use of EUMETSAT products”, IEEE Metrology for Aerospace (MetroAeroSpace) Conference, 4-5 June 2015, Benevento, IT (2005)
- [4] Mylne K., Big Data challenges in NWP and its Application, EMS Annual Meeting Abstracts Vol. 12, EMS2015-172 (2015)
- [5] Big Data - A New World of Opportunities, NESSI (The Networked European Software and Services Initiative) White Paper, (2012)
- [6] Thomas, J., Cook, K., *Illuminating the Path: Research and Development Agenda for Visual Analytics*, IEEE-Press (2005)
- [7] Gaztelumendi S et al., “Big data in the context of Euskalmet activities”, EMS Annual Meeting Abstracts Vol. 12, EMS2015-245, (2015)
- [8] Stimmel C.L., *Big Data Analytics Strategies for the Smart Grid*, CRC Press, (2016)
- [9] Spence R., *Information Visualization: an introduction*, 3rd Ed., Springer, (2014)
- [10] https://www.pnnl.gov/computing/visual_analytics/index.stm
- [11] <http://www.vismaster.eu/>
- [12] Keim D., Kohlhammer J., Ellis G. and Mansmann F. (Editors), *Mastering the Information Age Solving Problems with Visual Analytics*, VisMaster - European Coordination Action Project, Eurographics Association (2010)
- [13] Dill J., Earnshaw R., Kasik D., Vince J., Wong P.C., *Expanding the frontiers of Visual Analytics and Visualization*, Springer, (2012)

- [14] <http://www.va-sa.net>
- [15] Wong P. C., Thomas J., “Visual Analytics, Computer Graphics and Applications”, IEEE (Volume:24 , Issue: 5), (2004)
- [16] Keim D., Mansmann F., Schneidewind, Thomas J., Ziegler H., “Visual Analytics: Scope and Challenges”, Lecture Notes in Computer Science, No.4404, pp. 76-90 (2008)
- [17] Keim D., Mansmann F., Stoffel A., Ziegler H., “Visual Analytics”, Encyclopedia of Database Systems, Springer, (2009)
- [18] Segel E., Heer J., “Narrative Visualization: Telling Stories with Data”, IEEE Transactions on Visualization and Computer Graphics, Vol. 16 , Issue: 6 (2010)
- [19] Diehl A., Pelorosso L., Delrieux C., Saulo C., Ruizk J., Gröller L., Bruckner S., “Visual Analysis of Spatio-Temporal Data: Applications in Weather Forecasting”, In Eurographics Conference on Visualization (EuroVis) – Carr H. , Ma K.L., Santucci G. editors, Volume 34, Number 3, (2015)
- [20] Helbig C., Bilke L., Bauer H.S., Böttinger M., Kolditz O., “MEVA - An Interactive Visualization Application for Validation of Multifaceted Meteorological Data with Multiple 3D Devices”, <http://dx.doi.org/10.1371/journal.pone.0123811>, (2016)
- [21] Lundblad P.; Eurenus O.; Heldring T., “Interactive Visualization of Weather and Ship Data”, 13th International Conference Information Visualisation, 379 – 386, (2009)
- [22] <https://www.ecmwf.int/en/forecasts/datasets/archive-datasets>
- [23] <https://www.ecmwf.int/>
- [24] <http://www.cosmo-model.org/>
- [25] Moniruzzaman A. B. M., Hossain S. A., “NoSQL Database: New Era of Databases for Big data Analytics – Classification, Characteristics and Comparison”, International Journal of Database Theory and Application, Vol.6, No.4, (2013)
- [26] Dillon S., Brodeur J., “Supporting the Smart Grid 2.0 using In-Memory Database System”, (2015) <https://www.slideshare.net/StephenDillon/supporting-the-smart-grid-with-imdbs-48353189>
- [27] <https://www.exasol.com/en/>
- [28] <https://www.tableau.com/it-it>
- [29] https://help.tableau.com/current/pro/desktop/en-us/examples_exasolution.htm
- [30] <https://www.gartner.com/en/documents/3900992/magic-quadrant-for-analytics-and-business-intelligence-p>
- [31] <https://www.tableau.com/it-it/products/reader>

Microbial Genomic Data Analysis for Infectious Diseases

Dr. Margaret Varga, Dr. Charlene Rodrigues, Dr. Keith Jolley, Dr. Holly Bratcher, Dr. Jenny MacLennan, Dr. Odile Harrison, Dr. Fran Colles, Dr. Alison Cody, Dr. James Bray and Prof. Martin Maiden

Department of Zoology, University of Oxford, South Parks Road, Oxford, OX1 3SY,
UNITED KINGDOM

margaret.varga@zoo.ox.ac.uk

ABSTRACT

Infectious diseases are caused by pathogenic micro-organisms which can be bacteria, viruses, parasites or fungi. The diseases can be spread through many different routes, either directly or indirectly. Military personnel are at high risk of contracting infections, in particular vector-borne and zoonotic infections, during overseas deployments, where they may be exposed to endemic or emerging infections to which they do not have immunity. Additionally, overcrowded settings with poor sanitation are high risks for disease.

Genomics is having a transformational impact on medicine. It is enabling advances in accurate diagnosis of infectious disease, development of effective and targeted treatment strategies and opportunities to assess pathogenicity. Further, it supports the detection, surveillance of infectious diseases, the development and assessment of vaccines, as well as the assessment and prediction of anti-microbial resistance. These capabilities are all key military needs to protect personnel in this inter-connected world.

The advances in sequencing technologies have resulted in an explosion of genomic data. However, making sense of genomic data requires advances in computational analysis technologies together with cross-disciplinary scientific approaches, skill sets and people. There are extensive reference databases of genomic data. One such open access database is PubMLST.org: it contains well curated genomes for more than 100 microbial species and genera integrated with provenance and phenotype information. All levels of sequence data, from single gene sequences up to and including complete, finished genomes can be accessed on this platform. This data is, however, both large and complex and intractable to analyse and understand using traditional analysis tools.

This paper will discuss the challenges of analysing such genomic data for bacterial infections and consider the application of bioinformatics tools and techniques to analyse and communicate microbial genomic data in healthcare.

1.0 INTRODUCTION

There are three different categories of military casualties, namely, battle casualties, non-battle injuries and those suffering from non-battle related diseases or infections. Infectious diseases are caused by pathogenic micro-organisms which can be bacteria, viruses, parasites or fungi, with varying global epidemiology. Diseases can be spread through many different routes, including direct person-to-person spread, or indirectly through the air, food, water, and vectors such as arthropods or animals. Military personnel are at high risk of contracting infections, in particular vector-borne and zoonotic infections, especially during overseas deployments, as they travel to areas with different exposures to the home region against which they had developed immunity. Additionally, living in barracks or settings where sanitation may be poor and living condition crowded can lead to other public health risks [12].

The health status of military personnel has direct bearing on the success of a military campaign's outcome. The 1918 influenza pandemic (January 1918–December 1920), termed the Spanish flu, occurred during World War I and represented the most severe pandemic in recent history. The war and the pandemic intertwined. The infection was caused by an H1N1 virus with genes of avian origin which spread worldwide during 1918-1919. It was first identified in US military personnel in the spring of 1918. The compact living conditions, poor hygiene, malnutrition and the global troop movement contributed to the spread of the infection. The lack of vaccine and treatments resulted in an estimated 500 million people, or one-third of the world's population, becoming infected with this virus. The number of deaths was estimated to be at least 50 million worldwide, mostly young previously healthy adults, greater than the almost 20 million death toll of World War I [5, 20 and 32]. The US Army medical department concluded that secondary bacterial pneumonia, a common secondary infection complicating influenza, were the cause of nearly all the deaths rather than the primary infection [5]. Indeed, in World War I more deaths in the US armed forces (army and navy) engaged in the battlefield were due to pneumonia associated with the Spanish flu than from direct combat injuries [5]. The epidemic resulted in the diversion of personnel, resources, human attention as well as energy from the military campaign. Even during peace time, infectious diseases may spread easily in military bases and training camps resulting in epidemics with material effect on the military power of the units or the country in question. Vaccination and other preventative treatments play an important role in protecting the armed forces [3, 4, 5, 13, 18, 31 and 41].

2.0 MICROBIAL GENOMICS DATA

The sequencing of disease-causing microorganisms and study of genetic code is having a transformational impact on medicine, enabling advances in accurate diagnosis of infectious diseases, development of targeted treatment strategies and opportunities to assess pathogenicity of disease-causing organisms. Further, it supports the detection and surveillance of infectious diseases, vaccine development and post-implementation assessment and prediction of anti-microbial resistance. As the military, civilians and animals move with ease around the world, genomic technology can help in the areas of disease prevention, surveillance and management. These capabilities are all key military needs so as to protect personnel in this inter-connected world.

Multilocus Sequence Typing (MLST) was the most widely used portable, reproducible sequence-based approach for bacterial typing in the 21st century, it is used in identifying relationship among bacteria and cataloguing these using stable nomenclatures [25]. Seven MLST loci are indexed in most of the MLST schemes where an arbitrary and unique 'allele' number is assigned to each unique sequence for each locus. The designations for each of the loci are incorporated into a sequence type (ST), e.g. *Campylobacter jejuni* ST-21 or *Neisseria meningitidis* ST-11. Each ST summarizes thousands of base pairs of information; for some species, many hundreds of alleles at each locus and thousands of STs have been identified. Further STs can be grouped by similarity at more than four of the seven loci, into clonal complexes.

The Bacterial Isolate Genome Sequence Database (BIGSdb) platform is designed to store and analyse sequence data for bacterial isolates in which any number of sequences can be linked to isolate records [19]. It extends the principle of MLST to genomic data, where large numbers of loci can be defined, with alleles assigned by reference to sequence definition databases. Loci can also be grouped into schemes, most commonly of biological relevance, so that types can be defined by combinations of alleles (allelic profiles), a concept similar to MLST, Figure 1. This provides an effective mean to conduct gene-by-gene typing [26].

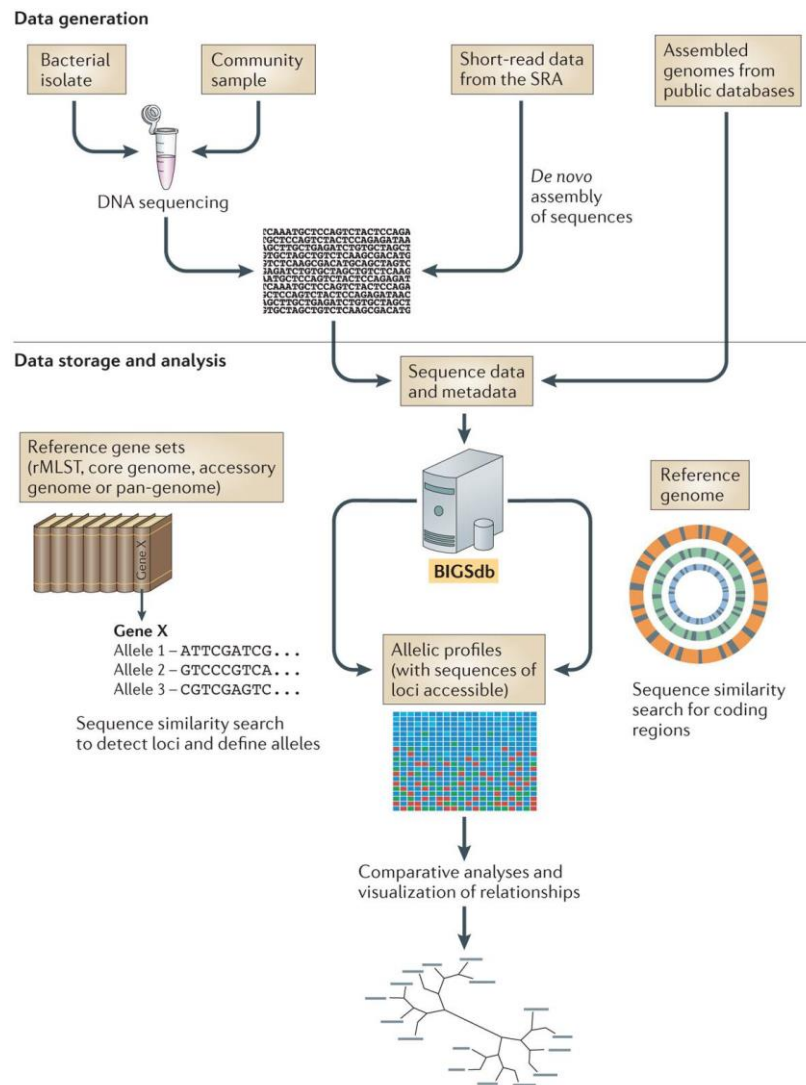


Figure 1: Illustrative representation of the gene-by-gene approach to whole genome sequence analysis using the Bacterial Isolate Genome Sequence Database (BIGSdb) platform. Reproduced with permission (Maiden et al. 2013) [26].

There are extensive reference databases of microbial genomic data. One such open access database is PubMLST.org which hosts databases using the Bacterial Isolate Genome Sequence Database platform (BIGSdb) [19 and 26]. Users can download genomic and associated metadata from those databases via PubMLST.org (or indeed from other public databases) and can analyse the data using the open source BIGSdb platform. PubMLST contains well curated genomes for more than 100 microbial species and genera integrated with provenance and phenotype information. It includes all levels of sequence data, from single gene sequences up to and including complete, reference genomes, see Figure 2. The allelic variants in the MLST data support sequence-based analysis. At the time of writing, PubMLST.org held more than 750,000 isolates from different countries and of different species.

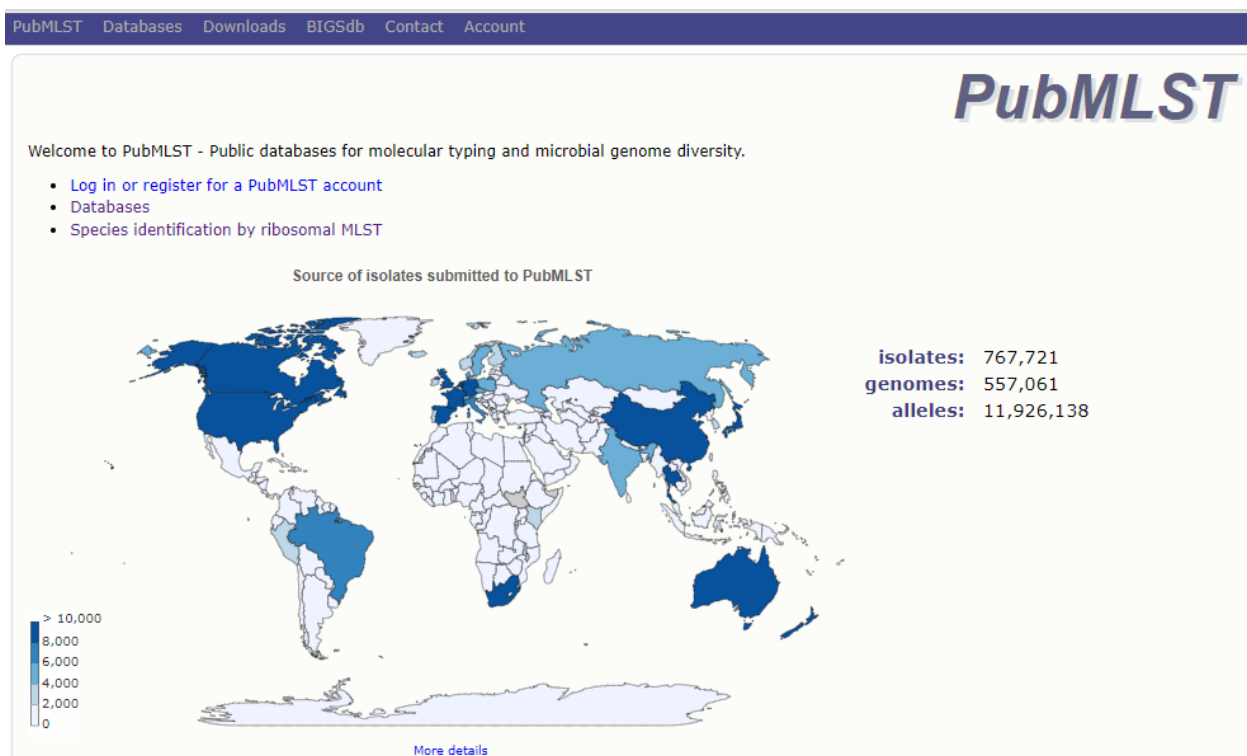


Figure 2: PubMLST (<https://pubmlst.org/>) The colour coded geospatial map shows from where the isolates were submitted; the darker the blue the more isolates submitted.

This data is both large and complex and is intractable to analyse and understand using traditional analysis tools.

3.0 MENINGOCOCCAL DISEASE AND EPIDEMIOLOGY

Significant progress has been made in developing effective vaccines against meningococcal disease and treatment in terms of antibiotic therapy and wider healthcare, but invasive disease remains a life-threatening disease. Meningococcal disease is considered to be a medical emergency; it requires immediate action for accurate diagnosis, rapid stabilisation and appropriate antimicrobial treatment [17 and 35]. A patient can present with non-specific symptoms and can progress to death in just a few hours. Permanent disabilities such as brain damage, hearing loss, learning disabilities and loss of limbs can result from meningococcal disease in up to one third of survivors [47]. The infection is spread from person-to-person; human carriers rarely become sick although they can be infectious. The epidemiology of meningococcal disease is highly variable, for reasons that are incompletely understood. It is a major cause of meningitis and sepsis in industrialised settings but the meningitis belt in Africa continues to have the highest burden of disease [40].

Military personnel and recruits are high risk groups for meningococcal disease. It has been reported that the incidence is between four and ten times higher than the in general population [30]. There have been many cases of meningococcal disease outbreaks in the armed forces across different countries, including those of the United Kingdom [29] and the USA [28 and 39]. The living conditions, in particular overcrowding on military bases and training camps, were implicated as far back as 1917 [10].

Neisseria meningitis has thirteen significant serogroups, among them, six serogroups, namely A, B, C, W, Y and X are responsible for the majority of cases of disease [14 and 15]. The distribution of each serogroup varies geographically and also changes over time [35 and 40]. Vaccines have been developed targeted at

specific serogroup(s). In view of the increasing use of these vaccines worldwide, high-throughput genomic methods are required to understand variations in capsular type and meningococcal protein diversity [36].

Vaccines have been developed to control the diseases, with development initiated at the Walter Reed Army Institute in the 1960's to protect military personnel [2 and 11]. The US military was involved in the development of the first licenced polysaccharide vaccine for serogroup A in 1970 followed by a combined serogroup A/C vaccine in 1978. Polysaccharide vaccines were later developed against serogroups W and Y, with X in development, and all have prevented countless cases of meningococcal disease globally. Meningococcal serogroup B vaccine was harder to develop as its polysaccharides are similar to the human neuronal proteins [9]. However, through the application of reverse vaccinology, a vaccine was designed utilising other expressed antigens predicted from the genome sequence of a meningococcal B strain (MenB) [8]. In the UK different types of vaccines are offered to different age groups to prevent meningitis and septicaemia; the 6-in-1, MenB, pneumococcal, Hib/Men C vaccines for babies, and meningococcal ACWY conjugate vaccine offered to teenagers and first-time university students. The UK armed forces personnel are offered meningococcal conjugate ACWY vaccine [46]. Vaccines can also be deployed in an outbreak setting, due to the high mortality and morbidity rates, and so the provision of accurate information is crucial in understanding, analysing and communicating an outbreak 'situation' to allow the infection prevention and control measures to be implemented in a timely manner [23].

4.0 VISUAL ANALYTICS AND VISUALIZATION

It is often said that a picture is worth a thousand words; in the case of genomic data a picture is worth millions of isolates. Humans have long used visual aids to help them define, understand, analyse and navigate their way through their problems, so as to understand the situation and thus enable informed decision-making. Visual analytics is the science of analytics reasoning supported by interactive visualizations [42]. The analytics approaches include statistical analysis, knowledge discovery, data management and knowledge representation [21]. Visual analytics and visualization are widely used to enable the exploration and analysis of the genomic data to gain insight of the situation. Visualization transforms inherently non-visual data into a form that supports efficient exploration, analysis, discovery, understanding and communication of large volumes of complex data. Visualization can be a static or interactive presentation of data that reinforces human cognition. Interactive visualization enables dynamic exploration and analysis of huge amounts of data that support users not only to detect expected patterns, trends, or correlations but also to discover unexpected associations more efficiently than using tradition approaches. When given a dataset, there are many different ways to represent the data and encode the various variables. However, certain displays are more effective than others for a given task, user, data and / or user, due to the different manner in which they exploit the underlying human perception and cognition [1, 21, 24, 37, 38, 4, 44 and 45]. In short, disease situational awareness encompasses the human perception and cognition skills as well as the processing of data by the machine. In such a complex and dynamic environment, acute situational awareness is likely to greatly enhance the rate and the quality of the decision-making process and infection management [6 and 7].

A wide range of visualization approaches have been developed and applied to support visual analytics of genomic data [33, 34 and 48]. Commonly, the large amount of data is aggregated to provide an overview of the situation and more detailed information can be 'drilled down' into for further analysis.

Figure 3 shows an exploration of *Neisseria meningitidis* for the period 2011-2017 in the United Kingdom using an integrated dashboard. In the integrated dashboard, each display represents a different aspect of the data: clicking on any element in any of the display will result in filtering of all the related elements in the other parts of the dashboard display. This provides an effective means of exploring different aspects of the data and their inter-relations and impacts. The top left hand area shows geographic data and the top right shows the temporal trend of the different serogroups and one non-group (NG), namely A, B, C, E,

W, W/ Y, X, Y and Z, these are colored coded so as to facilitate identification of the serogroup. The lower left stacked bar chart shows the temporal patterns of the clonal complexes. The lower right heatmap shows the nested relationship between clonal complex and serogroup. It can be seen that there is a noticeable decrease in serogroup B and a noticeable increase in serogroup W.

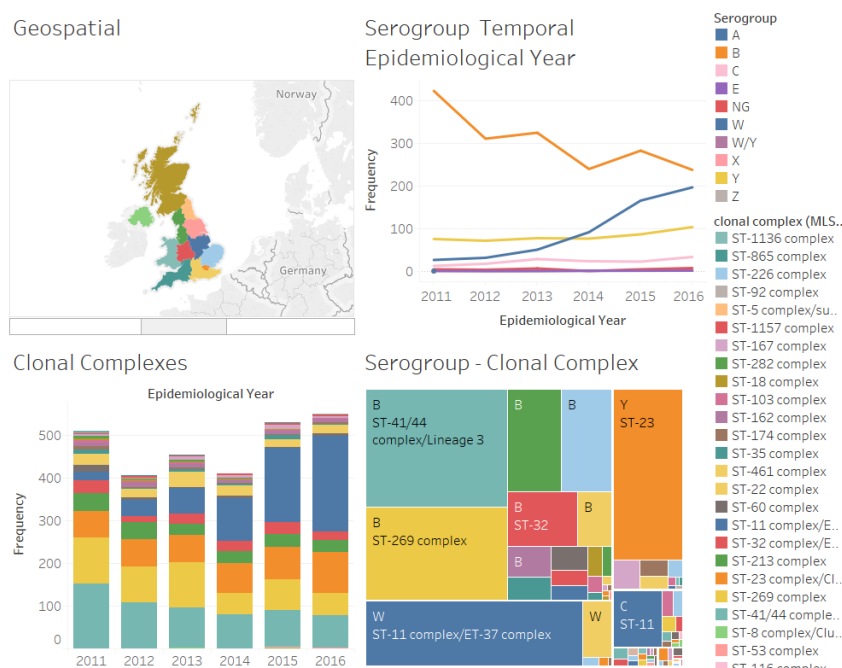


Figure 3: Geo-temporal Analysis of the *Neisseria Meningitidis*

Figure 4 shows 3,506 invasive meningococcal disease isolates from 2010/11 to 2016/17 analysed using the 24 loci OMVT (outer membrane vesicle typing) scheme and visualised using GrapeTree software. The OMVT cluster by clonal complex (cc), represented by different colours, with unfilled nodes representing isolates from the other ccs. The size of the circles is proportional to the number of isolates. [36].



Figure 4: UK meningococcal disease isolates clustered by OMVT [36]

The above examples show that the transformation of the non-visual data into a visual form enables effective analysis of large and complex microbial genomic data in an intuitive and understandable manner.

5.0 SUMMARY

The advances in sequencing technologies have resulted in an explosion of genomic data. Genomic research is vital in the fight against infections: for prevention, surveillance and treatment strategies. The study of microbial genomes has allowed the discovery of anomalies and patterns that have public health implications. Genomic and visual analytics tools can support and enhance our understanding of the biology of infections and lead to new techniques or approaches to diagnosing and treating disease. Visual analytics and visualization of genomic data are now widely used to support the analysis, presentation and understanding of these invaluable, big and complex data, as well as in the communication of the findings and discovery of insights. New technologies and evolving perception and cognitive frameworks enhance the development of better and more contextual and user-centred visual analytics approaches. Genomic and visual analytics combined with machine learning based AI now offer tremendous new opportunities.

6.0 REFERENCES

- [1] Albers, M. J., Human–Information interaction with complex information for decision-making. *Informatics*, 2(2):4-19, 2015.
- [2] Artenstein, M. S., Control of Meningococcal Meningitis with Meningococcal Vaccines. *Yale Journal of Biology and Medicine* 48(3): 197-200, 1975.
- [3] Barry, J., M., The site of origin of the 1918 influenza pandemic and its public health implications, *Journal of Translational Medicine*, volume 2, Article number: 3, 2004.
- [4] Brundage J. F. and Shanks, G. D., What really happened during the 1918 influenza pandemic? The importance of bacterial secondary infections. *The Journal of Infectious Diseases*. 196 (11): 1717–8, December 2007. Author reply 1718–9. doi:10.1086/522355. PMID 18008258.
- [5] Byerly, C. R., The U.S. Military and the Influenza Pandemic of 1918 – 1919, *Public Health Reports*, 2919 Supplement 3/ Volume 125, 2010.
- [6] Endsley, M. R., Toward a theory of situation awareness in dynamic systems. *Human Factors*, 37(1): 32 - 64, March, 1995.
- [7] Endsley, M. R. and Jones, D. G., *Designing for Situation Awareness: An Approach to User-Centered Design*, Second Edition, CRC Press, 2004, ISBN 9781420063554.
- [8] Findlow, J., Meningococcal group B vaccines. *Human Vaccines & Immunotherapeutics*, 9(6):1387–8. doi:10.4161/hv.24689 45, 2013.
- [9] Finne, J., Leinonen, M. and Makela, P. H., Antigenic similarities between brain components and bacteria causing meningitis. Implications for vaccine development and pathogenesis. *Lancet* 2(8346): 355-357, 1983.
- [10] Glover, J. A., Observations on the Meningococcus Carrier-Rate in relation to density of population in Sleeping Quarters. *The Journal of Hygiene* 17(4): 367-379, 1918.
- [11] Gotschlich, E. C., Goldschneider, I. and Artenstein, M. S., Human immunity to the meningococcus IV. Immunogenicity of group A and group C meningococcal polysaccharides. *Journal of Experimental Medicine* 129: 1367-1384, 1969.
- [12] Gray, G. C., Feighner, B., Trump, D. H, Berg S.W., Zajdowicz, M.J. and Zajdowicz, T. R., Disease spread by close personal contact. In: Lenhart MK, Lounsbury D. E, editors. *Military preventive medicine: Mobilization and deployment*, Volume 2. Washington DC: Borden Institute; 2005. p. 1127–8.
- [13] Hall, M. W., Inflammatory diseases of the respiratory tract (bronchitis, influenza, bronchopneumonia, lobar pneumonia). In: Siler JF, ed. *The Medical Department of the United States Army in the World War*. Volume IX: communicable and other diseases. Washington, DC: US Government Printing Office, 1928: 138. <http://www.ibiblio.org/hyperwar/AMH/XX/WWI/Army/Medical/IX/USA-Med-IX-2.html> (accessed 3/10/2019).
- [14] Harrison L. H, Trotter, C. L. and Ramsay, M. E., Global epidemiology of meningococcal disease.

Vaccine. 2009; 27 Supplement 2: B51-63.

[15] Harrison, O. B., Claus, H., Jiang, Y., Bennett, J. S., Bratcher, H. B., Jolley, K. A., Corton, C., Care, R., et al. (2013). Description and Nomenclature of *Neisseria meningitidis* Capsule Locus. *Emerging Infectious Diseases* 19(4): 566-573.

[16] Heer, J. and Shneiderman, B., Interactive Dynamics for Visual Analysis', *ACM Queue* 10(2), pp 1 - 30, February 2012.

[17] Jackson L. A., Schuchat, A., Reeves. M. W., Wenger J. D. Serogroup C meningococcal outbreaks in the United States. An emerging threat. *JAMA*. 1995 Feb 1;273(5):383-9.

[18] Jester, B., Uyeki, T. M., Jernigan, D. B. and Tumpey T. M., Historical and clinical aspects of the 1918 H1N1 pandemic in the United States. *Virology*. 2019 Jan 15;527:32-37. doi: 10.1016/j.virol.2018.10.019.

[19] Jolley, K. A. and Maiden, M. C., BIGSdb: Scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11(1): 595, 2010.

[20] Johnson N P. A. S. and Mueller, Updating the accounts: global mortality of the 1918- 1920 “Spanish” influenza pandemic. *Bulletin of the history of medicine*, 2002; 6:105-15.

[21] Keim, D., Andrienko, G. and Fekete, J. D. & Görg, C., Kohlhammer, J. & Melançon, G., Visual Analytics: Definition, Process, and Challenges. 10.1007/978-3-540-70956-5_7., 2008.

[22] Koppes, G.M., Ellenbogen, C. and Gebhart, R. J., Group Y meningococcal disease in United States Air Force recruits. *Am J Med*. 1977; 62(5):661–6. [PubMed: 404877].

[23] Leca, M., Bornet, C., Montana, M., Curti, C., Vanelle, P. Meningococcal vaccines: current state and future outlook. *Pathologie Biologie (Paris)* (2015) 63(3):144–51. doi:10.1016/j.patbio.2015.04.003

[24] Lurie, N. H., Mason, C.H., Visual representation: Implications for decision making. *Journal of Marketing* 2007, 71(1):160- 177.

[25] Maiden, M. C. J., Bygraves, J. A., Feil, E., Morelli, G., Russell, J. E., Urwin, R., Zhang, Q., Zhou, J., Zurth, K., Caugant, D. A., Feavers, I. M., Achtman, M. and Spratt, B. G. (1998).

Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. *Proceedings of the National Academy of Sciences of the United States of America* 95(6): 3140-3145.

[26] Maiden, M. C., Jansen van Rensburg, M. J., Bray, J. E., Earle, S. G., Ford, S. A., Jolley, K. A. and McCarthy, N. D., MLST revisited: the gene-by-gene approach to bacterial genomics. *Nature Reviews Microbiology* 11(10): 728-736, 2013.

[27] Maiden, M. C., The impact of protein-conjugate polysaccharide vaccines: an endgame for meningitis? *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences* 368(1623): 20120147, 2013.

[28] Makras, P., Alexiou-Daniel, S., Antoniadis, A. and Hatzigeorgiou, D., Outbreak of meningococcal disease after an influenza B epidemic at a Hellenic Air Force recruit training center. *Clin Infect Dis*. 2001;33(6):e48–50. doi: 10.1086/322609. [PubMed: 11512107].

[29] Masterton, R. G., Youngs, E. R., Wardle, J. C., Croft, K. F. and Jones, D. M. Control of an outbreak of group C meningococcal meningitis with a polysaccharide vaccine. *J Infect*. 1988;17(2):177–82. [PubMed: 3141518].

[30] Mehrdadi, S., Acute Bacterial Meningitis: Diagnosis, Treatment and Prevention, *Journal of Archives in Military Medicine*, 6(4):e84749, December 2018.

[31] Morens, D. M., Fauci, A. S., The 1918 influenza pandemic: insights for the 21st century. *The Journal of Infectious Diseases*. 195 (7): 1018–28. April 2007. doi:10.1086/511989. PMID 17330793.

[32] Mougel, N., CVCE, 2011, 2011, English translation, Gratz, J., Centre European Roobert Schuman, Reperes – module 1-0 – explanatory notes – World War I casualties – EN.

[33] Overmars, L., Kerkhoven R., Siezen R. J., Francke, C.: MGcV: the microbial genomic context viewer for comparative genome analysis. *BMC Genomics* 14 (Apr. 2013), 209. doi: 10.1186/1471-2164-14-209.

[34] Ramírez F, Ryan DP, Grüning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dündar F, Manke T. deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Research*. 2016 Apr 13.

- [35] Rosenstein N. E., Perkins B. A., Stephens D. S., et al. Meningococcal disease. *N Engl J Med.* 2001; 344(18): 1378-88.
- [36] Rodrigues, C. M. C., Genomic Epidemiology of Meningococcal Vaccine Antigens in the United Kingdom, PhD thesis, December 2018.
- [37] Sacha, D., Stoffel, A., Stoffel, F., Kwon, B. C., Ellis, G. and Keim, D. A.: Knowledge generation model for visual analytics. *Visualization and Computer Graphics, IEEE Transactions on* 2014, 20(12):1604-1613.
- [38] Sedig, K. and Parsons, P., Design of visualizations for human information interaction: A pattern-based framework. *Synthesis Lectures on Visualization* 2016, 4(1):1-185.
- [39] Smilack, J. D., Group-y meningococcal disease. Twelve cases at an army training center. *Ann Intern Med.* 1974;81(6):740–5.[doi: 10.7326/0003-4819-81-6-740] [PubMed: 4215351]
- [40] Sridhar, S., Greenwood, B., Head, C., Plotkin, S. A., Safadi, M. A., Saha, S., Taha, M. K., Tomori, O. and Gessner, B. D. (2015). Global incidence of serogroup B invasive meningococcal disease: a systematic review. *Lancet Infectious Diseases* 15(11): 1334-1346.
- [41] Taubenberger, J. K. and Morens, D. M., 1918 Influenza: the Mother of All Pandemics, *Emerging Infectious Diseases*, Volume. 12, No. 1, pp 15 -22, January 2006.
- [42] Thomas, J. J. and Cook, K. A. (Eds.), *Illuminating the Path: The Research and Development Agenda for Visual Analytics*, National Visualization and Analytics Center, 2005.
- [43] Tory, M., and Moller, T., *Human Factors In Visualization Research*, *IEEE Transactions on visualization and Computer Graphics*, Volume 10, No. 1, January/February 2004.
- [44] Tufte, E. *The Visual Display of Quantitative Information*, Graphic Press, Second Edition, ISBN-13: 978- 0961392147, May 2001.
- [45] Ware, C., *Information Visualization: Perception for Design*, 3rd Edition. Morgan Kaufmann Publishers, Inc., San Francisco, California, 2012.
- [46] Vaccination in the armed forces – what you need to know, Annex B, TO AFC(H) JIs Dated SEP/OCT 17.
- [47] Viner, R. M., Booy, R., Johnson, H., Edmunds, W. J., Hudson, L., Bedford, H., Kaczmarek, E., Rajput, K., et al. (2012). Outcomes of invasive meningococcal serogroup B disease in children and adolescents (MOSAIC): a case-control study. *Lancet Neurology* 11(9): 774-783.
- [48] Yachdav, G., Wilzbach, S., Rauscher, B., Sheridan R., Sillitoe, I., Procter, J., Lewis, S. E., Rost, B., Goldberg, T.: MSAViewer: interactive JavaScript visualization of multiple sequence alignments. *Bioinformatics* 32, 22 (Nov. 2016), 3501–3503.

