Ethical Control of Unmanned Systems

Focus Topic:

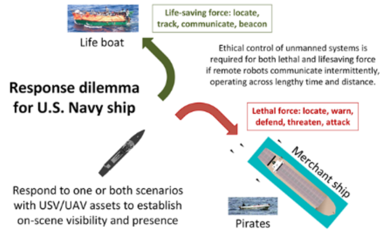
Data-Centric Security Considerations

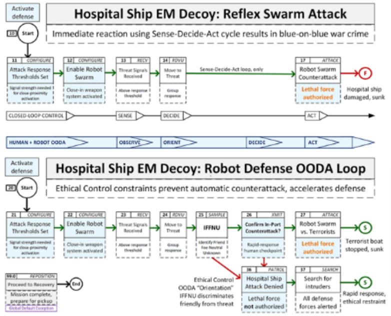
Don Brutzman and Curt Blais

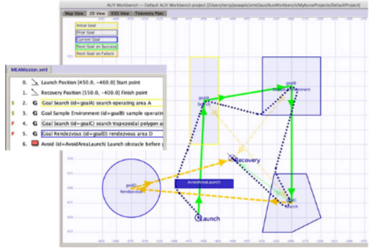
Naval Postgraduate School (NPS)

16 April 2020

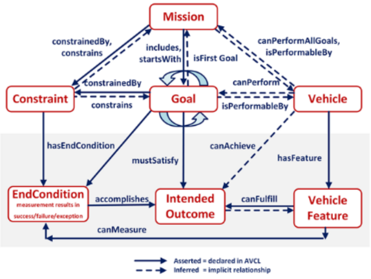
Ethical Control of Unmanned Systems:
Keeping Warfighters in Charge of Autonomy


Life-saving force: locate, track, communicate, beacon
Life boat
Ethical control of unmanned systems is required for both lethal and lifesaving force if remote robots communicate intermittently, operating across lengthy time and distance.
Response dilemma for U.S. Navy ship
Lethal force: locate, warn, defend, threaten, attack
Merchant ship
Respond to one or both scenarios with USV/UAV assets to establish on-scene visibility and presence
Pirates


Activate defense
Hospital Ship EM Decoy: Reflex Swarm Attack
Immediate reaction using Sense-Decide-Act cycle results in blue-on-blue war crime
Start
Sense-Decide-Act loop, only
Hospital ship damaged, sunk
Activate defense
Hospital Ship EM Decoy: Robot Defense OODA Loop
Ethical Control constraints prevent automatic counterattack, accelerates defense
Start
Terrorist boat stopped, sunk
Rapid response, ethical restraint



Mission Execution Ontology (MEO) evolving



Milestones and Transitions

- CRUSER development led to first project selection under CRADA with Raytheon Missile Systems (RMS).
- Successful progress on test missions entering TRL 5 with simulation and Web-sharable 3D visualization.
- Expressing multiple robot mission plans consistently, coherently for diverse UAV, USV, UUV platforms.
- Use Semantic Web Standards to support warfighters.
- Evaluate NAVSEA Unmanned Maritime Autonomy Architecture (UMAA) evolution for robot qualification.

Why / Objectives

- Ethical control of unmanned systems can be accomplished through structured mission definitions that are trusted, consistently readable, validatable, repeatable and understandable by humans and robots.
- Orders must be lawful. Unmanned systems must behave ethically and comprehensibly if they are to support manned military units effectively.
- Well-structured mission orders can be tested and trusted to give human commanders confidence that offboard systems *will do what they are told to do*, and further *will not do what they are forbidden to do*.
- Demonstrate that no technological limitations exist that prevent applying the same kind of ethical constraints on robots and unmanned vehicles that already apply to humans, in lethal and life-saving scenarios.

https://savage.nps.edu/EthicalControl

What / Deliverables

- Update Mission Execution Ontology (MEO) concepts demonstrated in tests and simulation, building to perform field experimentation (FX).
- Supervise thesis work to explore canonical exemplar missions that are expected to utilize unmanned systems, looking across the full range of Naval warfare communities. Example scenarios include UAV for sailor overboard, UAV for refugee/lifeboat escort, and adept scouts. All must observe Law of Armed Conflict (LOAC), Rules of Engagement (ROE), and moral guidance of commanders despite long durations/distances.
- Define, simulate, and test combination of real-world goals and ethical constraints to robot mission tasking across set of canonical scenarios.
- Illustrate how human-robot teams meet moral and legal requirements if deploying unmanned systems with potential for lethal, life-saving force.

NPS PRAESTANTIA PER SCIENTIAM

Building on Concept Demonstrations

Principal Investigator: Don Brutzman
brutzman@nps.edu 1.831.656.2149
Co-Investigator: Curtis Blais
clblais@nps.edu 1.831.656.3215

Synopsis: Ethical Control of Unmanned Systems

- **Project Motivation**: ethically constrained control of unmanned systems and robot missions by human supervisors and warfighters.
- **Precept:** well-structured mission orders can be syntactically and semantically validated to give human commanders confidence that offboard systems
  - *will **do** what they are told to **do***, and further
  - *will **not do** what they are **forbidden to do***.

Paraphrase: can qualified robots correctly follow human orders?

- **Project Goal:** apply Semantic Web ontology to scenario goals and constraints for logical validation that human-approved mission orders for robots are semantically coherent, precise, unambiguous, and without internal contradictions.
- **Long-term Objective:** demonstrate that no technological limitations exist that prevent applying the same kind of ethical constraints on robots and unmanned vehicles that already apply to human beings.

Table of Contents

- Flyer and Synopsis
- Background History
- Mission Representation using Autonomous Vehicle Command Language (AVCL)
- Semantic Web and Mission Execution Ontology (MEO)
- Ethical Control Mission Design
- Exemplar Missions
- Multiple Programming Languages
- Simulation and Visualization
- **Data-Centric Security and Trust**
- Related Ethical Activities
- Related Resources of Interest
- Conclusions and Recommendations
- Contact

Data-Centric Security and Trust

Compression, authentication, encryption, composability, blockchain ledger, and asymmetric advantages enable group communication of secure mission orders and responses. This is a Chain of Trust for distributed Command Authority.

XML Security for Data: **Structure** 1

1. Extensible Markup Language (XML) provides formal structure for data models and information exchange.
   a. "XML is a markup language that defines a set of rules for encoding documents in a format that is both human-readable and machine-readable." – Wikipedia
   b. Declarative and self-describing data structures, not program source code.
   c. Data validation through XML Schema includes strong typing of values and correct parent-child hierarchical relationships.
   d. Avoids Garbage In Garbage Out (GIGO) pathologies when communicating between multiple systems and across related protocols.
   e. Similarly applicable using JavaScript Object Notation (JSON) and other formats.
   f. Offers complete precision of expressive power when defining human orders and system responses, e.g. via Autonomous Vehicle Command Language (AVCL).

XML Security for Data: **EXI Compression** 2

2. Efficient XML Interchange (EXI) provides best-possible compression of XML documents, reducing size and speeding up decompression.
    a. Years of work by exceptionally competent working group, proven results.
    b. EXI Recommendations by World Wide Web Consortium (W3C).
    c. Multiple open-source and commercial implementations in Java, C++
    d. Preserves sufficient structure for lossless composition of compressed XML.

Thus even signed and encrypted data documents shown in this work can get best-proven compression for use on limited, disadvantaged and challenged communications links facing deployed Naval forces.

Network Optional Warfare (NOW): Efficient Messaging

Navy networks afloat are very different than networks ashore. Bandwidth is a precious and finite resource, latency can be huge, connectivity can be intermittent, environmental effects dominate, channels are limited in varying ways, and mobile relays are rare. Manned and unmanned naval systems need efficient messaging for networks afloat - but rarely have it. Failing to properly utilize communications capacity directly limits tactical effectiveness.

Efficient messaging is needed to take maximum advantage of severely constrained data links. The key to our strategies for achieving efficient messaging is first to use of Extensible Markup Language (XML) for structured data languages, and then use EXI for compressing XML. Since XML provides a flexible and validatable way to define regular data structures for any language, it provides a practical opportunity to compatibly capture and convert all manner of diverse data formats used for military messaging. The economics of Web technologies are undeniable and usually provide industry-wide best practices as well. As a result, this use of open standards is scalable and repeatable, avoiding the "stove pipes" which commonly prevent system-wide interoperability between Navy platforms and coalition partners.

"Efficiency" means both size and speed. EXI has demonstrated compaction that *always* meets or beats the most commonly used compression techniques (zip and gzip). Additionally, because EXI decompression goes straight into memory rather than string characters, which then require significant additional parsing, decoding EXI is many times faster than other techniques. This approach also reduces memory requirements and power consumption on small devices. Because Navy tactical traffic is usually highly structured and highly numeric, EXI provides major advantages that might well impact all afloat Navy communications. Alternative bit-centric compression schemes cannot take full advantage of those characteristics.

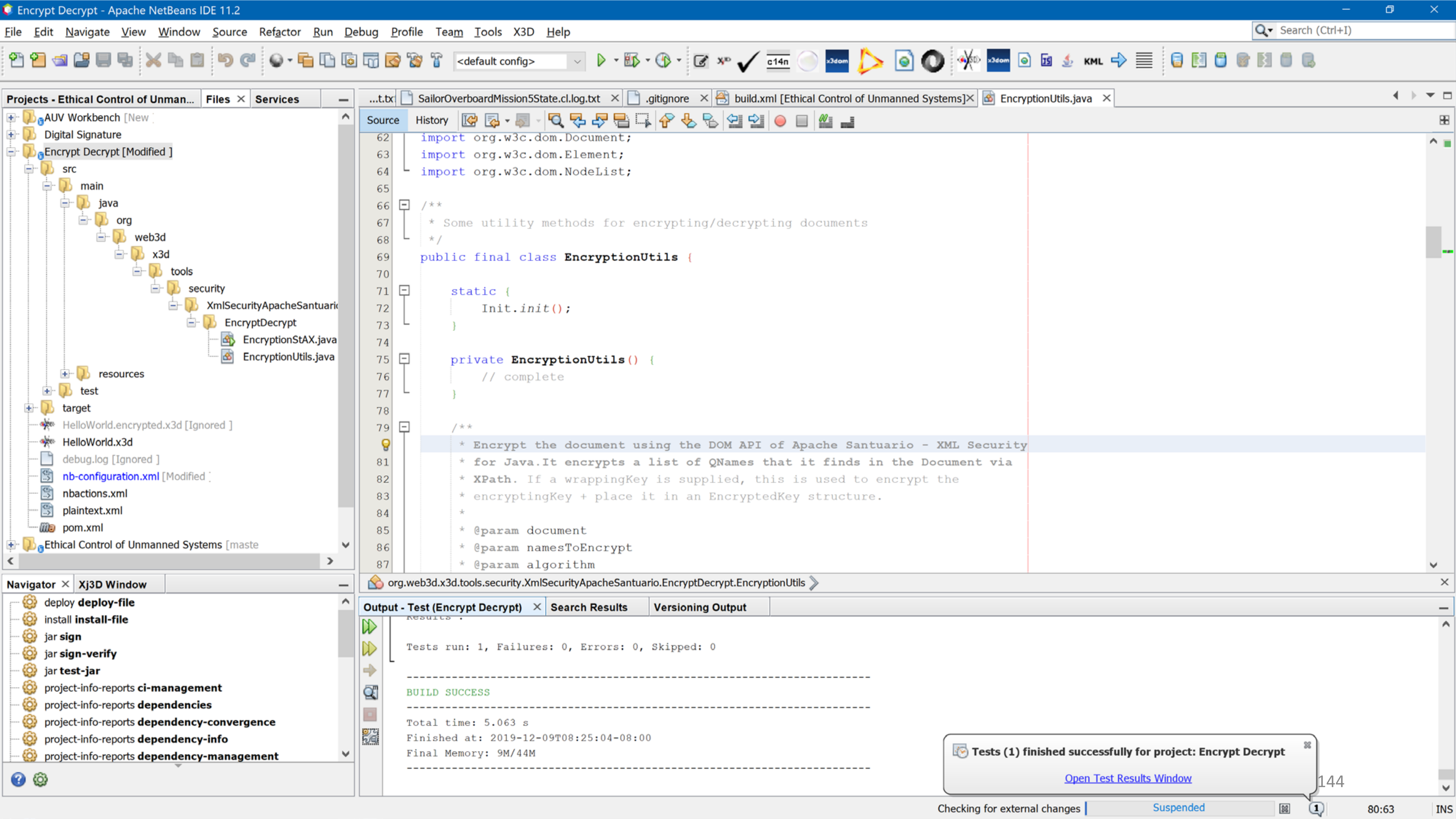
"Efficiency" is compatible with Data-Centric Security. Demonstrated thesis work has shown that digital signature (for authentication) and XML Encryption (privacy and access control) can coexist with efficient compression, when applied in the correct order. Such interoperability for Information Assurance (IA) Is necessary when working with coalition partners, and also for safeguarding data within deployed unmanned systems that are beyond the reach of network-centric security.

XML Security for Data: **Digital Signature** 3

3. XML Digital Signature (DS) defines XML syntax for digital signatures.
   a. W3C Recommendation, stable since 2013, international adoption.
      - https://www.w3.org/TR/xmldsig-core1
   b. Public-private key pairs for signature/authentication, key distribution is separate.
   c. Applicable to entire documents or to fragments (subsections).
   d. Requires XML Canonicalization of input documents to regularize formatting so that identical documents are uniquely expressed.
   e. Can sign any data resource for identity verification, non-repudiability, confirmation that original information has not been tampered with, etc.
   f. Completely compatible for data handling within trusted networks.
   g. 2019 NPS has adapted open-source Java version of Apache Santuario as utility classes and test suite for further use. Prior examples from years ago still work.
      https://sourceforge.net/p/x3d/code/HEAD/tree/www.web3d.org/x3d/tools/security/XmlSecurityApacheSantuario

XML Security for Data: **Encryption** 4

4. XML Encryption (XML-Enc) defines how to encrypt XML data.
   a. W3C Recommendation, stable since 2013, international adoption.
      - https://www.w3.org/TR/xmlenc-core1
   b. Public-private (i.e. shared-secret) key pairs, key distribution is separate.
   c. Applicable to entire documents or to fragments (subsections).
   d. Different from Transport Layer Security (TLS) used by http/https for sending encrypted traffic over the Internet.
   e. Some vulnerabilities were reported publicly, but each was performed via exhaustive attacks against server, incrementally analyzing error responses. Not likely or practical mode of attack against unmanned systems.
   f. Completely compatible for data handling within already-trusted networks.


Encrypt Decrypt - Apache NetBeans IDE 11.2
File Edit Navigate View Window Source Refactor Run Debug Profile Team Tools X3D Help
Search (Ctrl+I)
<default config>
Projects - Ethical Control of Unman... Files Services
AUV Workbench [New
Digital Signature
Encrypt Decrypt [Modified]
src
main
java
org
web3d
x3d
tools
security
XmlSecurityApacheSantuari
EncryptDecrypt
EncryptionStAX.java
EncryptionUtils.java
resources
test
target
HelloWorld.encrypted.x3d [Ignored]
HelloWorld.x3d
debug.log [Ignored]
nb-configuration.xml [Modified
nbactions.xml
plaintext.xml
pom.xml
Ethical Control of Unmanned Systems [maste
...t.tx SailorOverboardMission5State.cl.log.txt .gitignore build.xml [Ethical Control of Unmanned Systems] EncryptionUtils.java
Source History
Navigator Xj3D Window
deploy deploy-file
install install-file
jar sign
jar sign-verify
jar test-jar
project-info-reports ci-management
project-info-reports dependencies
project-info-reports dependency-convergence
project-info-reports dependency-info
project-info-reports dependency-management
Output - Test (Encrypt Decrypt) Search Results Versioning Output
Checking for external changes
Suspended
80:63
INS

```java
import org.w3c.dom.Document;
import org.w3c.dom.Element;
import org.w3c.dom.NodeList;

/**
 * Some utility methods for encrypting/decrypting documents
 */
public final class EncryptionUtils {

    static {
        Init.init();
    }

    private EncryptionUtils() {
        // complete
    }

    /**
     * Encrypt the document using the DOM API of Apache Santuario - XML Security
     * for Java.It encrypts a list of QNames that it finds in the Document via
     * XPath. If a wrappingKey is supplied, this is used to encrypt the
     * encryptingKey + place it in an EncryptedKey structure.
     *
     * @param document
     * @param namesToEncrypt
     * @param algorithm
```

org.web3d.x3d.tools.security.XmlSecurityApacheSantuario.EncryptDecrypt.EncryptionUtils

```
Results :

Tests run: 1, Failures: 0, Errors: 0, Skipped: 0

------------------------------------------------------------------------
BUILD SUCCESS
------------------------------------------------------------------------
Total time: 5.063 s
Finished at: 2019-12-09T08:25:04-08:00
Final Memory: 9M/44M
------------------------------------------------------------------------
```

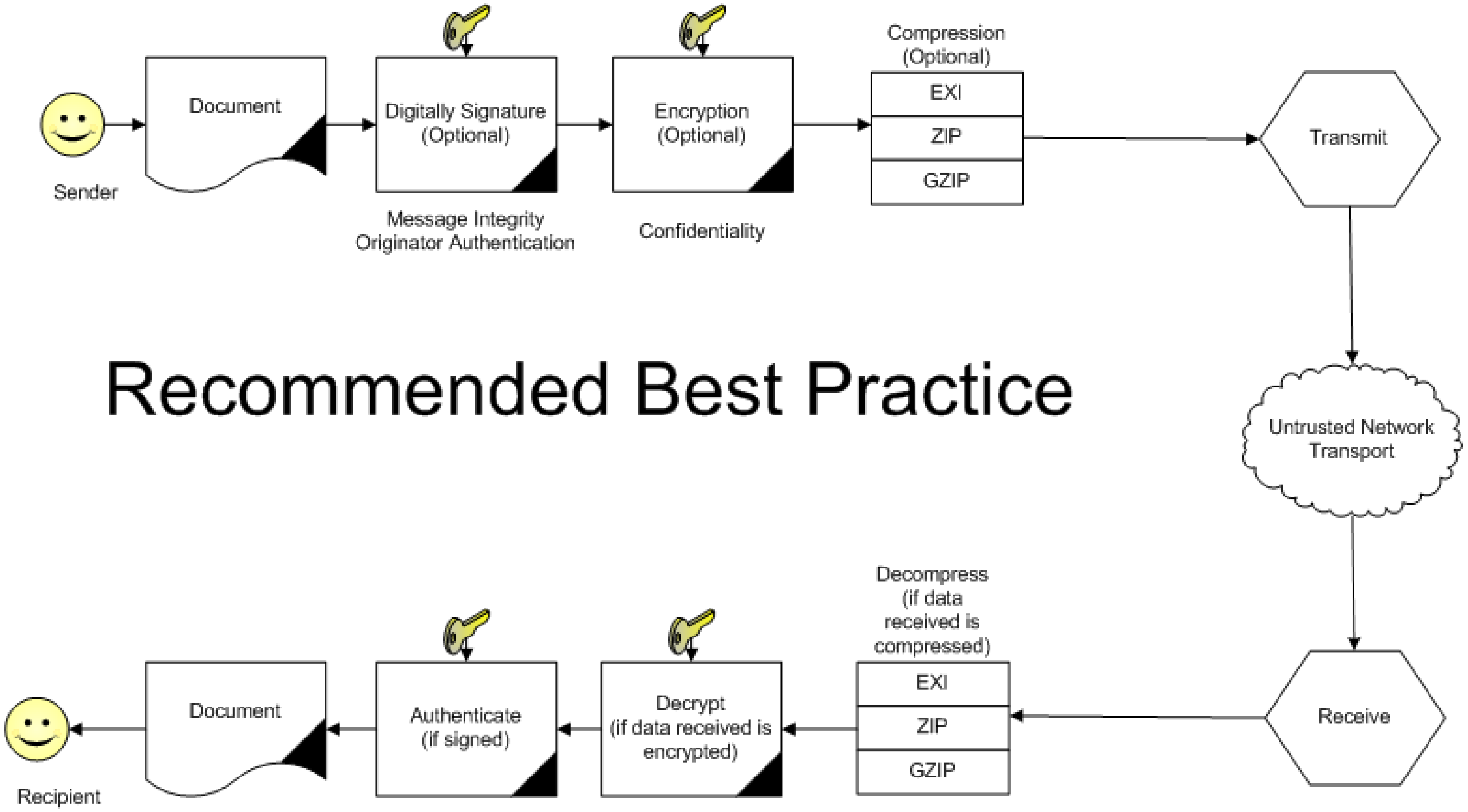
Tests (1) finished successfully for project: Encrypt Decrypt

Open Test Results Window

XML Security: **Composition** 5

5. EXI Compression, XML Digital Signature and XML Encryption can be composed for applying to data in single files/documents/messages.
    a. Each technology works on data formatted as valid XML.
    b. Multiple NPS theses have examined EXI characteristics in combination with XML Security.
    c. Such composition is partially demonstrated, appears completely feasible.
    d. Williams, Jeffrey S., Document-based message-centric security using XML authentication and encryption for coalition and interagency operations, Masters Thesis, Naval Postgraduate School, Monterey, California, 2009.
    e. Each is usable in concert for data-centric security, compatibly within any secure network or within fixed/mobile data storage of unmanned systems.

Williams Thesis: Composition of EXI Compression, XML Authentication, and XML Signature


Document
Digitally Signature (Optional)
Encryption (Optional)
Compression (Optional)
EXI
ZIP
GZIP
Transmit
Sender
Message Integrity
Originator Authentication
Confidentiality
Recommended Best Practice
Untrusted Network Transport
Decompress (if data received is compressed)
EXI
ZIP
GZIP
Document
Authenticate (if signed)
Decrypt (if data received is encrypted)
Receive
Recipient

Security Assertion Markup Language (SAML)

- Security Assertion Markup Language (SAML) is an open standard for exchanging authentication and authorization data between parties, in particular, between an identity provider and a service provider.
- SAML is an XML-based markup language for security assertions (statements that service providers use to make access-control decisions).
- With some adaptation work, SAML might be used to formally describe policies and requirements for data-centric security of mission orders.

References

- RDML Danelle Barrett USN, "The data-driven Department of the Navy," *CHIPS*, January-March 2018
- Latest version is SAML 2.0 standardized by OASIS

OASIS
Open standards. Open source.

Blockchain distributed ledger characteristics

- "A **distributed ledger** is a consensus of replicated, shared, and synchronized digital data geographically spread across multiple sites, countries, or institutions. There is no central administrator or centralized data storage." – Wikipedia
- Design characteristics can be tuned to match system needs and include strict sequencing of ledger entries, nonrepudiability of message indexes, consensus algorithm (proof of work or stake), etc.
- Implementation is often accomplished via a blockchain system.

Future work: ships, aircraft and ground systems might maintain a strong distributed ledger of all messages sent and received, reducing risk of spoofing or counterfeit messages compromising unmanned systems.

Significant protections from hostile takeover… is possible for deployed friendly-force robots

Accountability for actions requires a traceable, provable decision tree.

The following vulnerability "anti-pattern" provides an interesting use case, whereby non-repudiability of mission orders can prevent an opponent from falsely claiming a "rogue robot" or "rogue commander" scenario:

- Opponent captures control of a friendly unmanned system (physically or through cyber attack).
- Opponent has no key, unable to decrypt previously recorded sensor data.
- Opponent disables onboard security interlocks, directs unmanned system to execute hostile act (e.g. attack on friendly or neutral force).
- Post-incident investigation reveals and proves that mission orders were not authenticated or authorized by original friendly commander.
- Block-chain ledger of all issued authenticated orders reveals that no gaps occurred in shipboard records of approved missions.

Data-Centric Security and Command Authority

Data-centric security that includes authentication of ordered missions for unmanned systems provides a military, legal, ethical and moral basis for non-repudiability and accountability of human commanders.

- Authorized humans remain in charge, accountable for robot actions.
- Collected robot data is encrypted in asymmetric manner, greatly reducing vulnerabilities following any robot capture or compromise.
- Data-centric security can coexist within all levels of network security.

Such reliability provide excellent rationale to link data-centric security to design considerations for Ethical Control, compatibly across all networks.

- Once again, Ethical Control leads to *more-effective warfighting*.

Trust

Trusted Mission Orders

- Formal shared meaning between robots and human commanders
- Controlled vocabulary of terms with well-defined conditions, outcomes
- Syntax validation, well-formed data
- Numerical validation, in bounds
- Semantic confirmation of tactical prerequisites, coordination steps
- No logical contradictions present

Trusted Mission Execution

- Portable tasking across diverse unmanned systems, C4I networks
- Data-centric encryption for transmission across any network
- Digital-signature authentication that confirms command identity
- Blockchain ledger authoritatively confirms completeness, no gaps
- Testable in simulation, eventually formalized as robot qualification

Blockchain ledger for distributed accountability

Given a trusted chain of message exchange among participating human commands and distributed systems, there are additional vulnerabilities that still need to be considered. Blockchain technology is relevant.

Obvious tactical accountability issue is missing gaps or jammed messages

- Failure to receive even one message (perhaps requiring human permissions) can invalidate any subsequent actions, thereby losing control of lethal force.

Extrapolating further needs: after-action analysis, investigation, improvement.

- Having a ledger of all received/sent messages can provide accountability and verifiable chain of trust for authoritative reconstruction and progress.

Important future work: custom blockchain providing assurances that scale among diverse participants and over time, without needing a central hub.

Related work: Zero Trust Architecture (ZTA)

Zero Trust Architecture, Scott Rose (NIST), Oliver Borchert (NIST), Stu Mitchell (Stu2Labs), Sean Connelly (DHS), 2nd Draft, NIST SP 800-207, February 2020. zerotrust-arch@nist.gov

- "Zero trust refers to an evolving set of network security paradigms that narrows defenses from wide network perimeters to individual resources. Its focus on protecting resources rather than network segments is a response to enterprise trends that include remote users and cloud-based assets that are not located within an enterprise-owned network boundary."
- https://csrc.nist.gov/publications/detail/sp/800-207/draft

Data-centric security seems like logical conclusion of such an approach.

Next steps: implement, evaluate, deploy tests

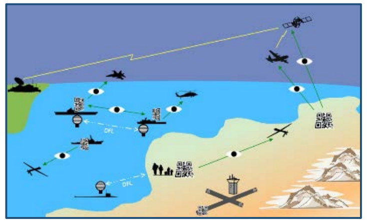
- Data-centric security can provide guarantees of command authority over the application of lethal lifesaving force by unmanned systems.
- Open standards and implementations exist for each component: compression, signature, encryption, assertion metadata, etc.
    - Alternative technologies are also available.
- Composition testing with robots during field experimentation (FX) can extend laboratory results with real-world experience, risk analysis and red-team testing.
- Further work recommended.

Related resources of interest

This project draws on multiple relevant activities and capabilities.
The following synopses are distilled from each respective resource.

Network Optional Warfare (NOW)

Overall Concept



Naval forces do not have to be engaged in constant centralized communication. Deployed Navy vessels have demonstrated independence of action in stealthy coordinated operations for hundreds of years.

- Littoral operations, deployable unmanned systems, and a refactored force mix for surface ships pose a growing set of naval challenges and opportunities. Network-optional warfare (NOW) precepts include Efficient Messaging, Optical Signaling, Semantic Coherence and Ethical Human Supervision of Autonomy for deliberate, stealthy, minimalist tactical communications.
- https://wiki.nps.edu/display/NOW/Network+Optional+Warfare

Trusting Software and Trusting Data

- Network Optional Warfare (NOW) Blog, January 2016
- In 1983, Dennis Ritchie and Ken Thompson jointly received the Turing Award for their development of generic operating systems theory, and specifically for the implementation of the UNIX operating system.
- Ken Thompson's lecture was Reflections on Trusting Trust, with the subtitle "*To what extent should one trust a statement that a program is free of Trojan horses? Perhaps it is more important to trust the people who wrote the software.*" This talk can still surprise: he describes source code that looks like it does one thing, but actually performs things that are quite different.
- So in effect, Ken Thompson chose his Turing Award moment to reveal to the world that he had superuser and user access for every Unix system and server on the planet. Further he revealed that, even with a great many people scrutinizing and rebuilding the source code, and even despite users banging on Unix daily everywhere, anyone else might use a super password for each and every account. Meanwhile no one else knew that the super password existed, much less that it quietly insisted on re-propagating itself in each fresh new copy of Unix. **No kidding.**
- How does the Navy get beyond software barriers to reach the next level of capability: **trust for shared data**?

DoD Digital Modernization Strategy

DoD CIO Digital Modernization Strategy provides Information Resource Management Strategic Plan FY 19-23, guiding IT transformation (interview)

- Advancement of digital environment to ultimately ensure competitive advantage for warfighters. Addresses Cybersecurity, AI, Cloud,C3.
- Four initiatives: Innovation, Optimization, Cybersecurity, Talent.
- Relevant sections include: Establish JAIC, Enterprise Cloud, Modernize C4, Treat Data as Strategic Asset, International Collaboration Partnerships and Allied Interoperability, Protect Positioning Navigation & Timing (PNT), End-To-End Airborne ISR Data Transport, Info Sharing to Mobile Users, Drive Standards into IT Systems, Transform Cybersecurity Architecture, End-To-End Identity Credential Asset Management, Risk Management

CRUSER · NEWS
Consortium for Robotics and Unmanned Systems Education and Research

NPS
PRAESTANTIA PER SCIENTIAM
1909

ETHICAL MISSION DEFINITION AND EXECUTION FOR MARITIME ROBOTS UNDER HUMAN SUPERVISION

- Lethality requires ethical and legal basis, supervised by military teams.
- Executable robot tasking can resemble tactical tasking of humans afloat.
- Careful application of goal constraints makes ethical control feasible.
- Robot missions then complement and extend naval operation orders.
- Semantic Web logic can confirm ethical correctness and completeness.
- Next steps: continue 2 decades of work with realistic scenario testing.

> "Ethical constraints on robot mission execution are possible today. There is no need to wait for future developments in Artificial Intelligence (AI). It is a moral imperative that ethical constraints in some form be introduced immediately into the software of all robots that are capable of inflicting unintended or deliberate harm to humans or property." Robert McGhee, April 2016

- IEEE Journal of Oceanic Engineering (JOE) paper along with online references.
- Authors Don Brutzman, Curtis Blais, Duane Davis and Robert McGhee, NPS.
- Feedback and recommendations always welcome. Contact: brutzman@nps.edu