DNV GL Safety & Security フォーラム

AIのセキュリティとトラスト に関する国際学術動向

櫻井　幸一
九州大学
大学院　システム情報科学研究院　& サイバー セキュリティセンター

背景/目的
国内の＊＊学会にAI-Security研究会を立ち上げること：
その準備としての国内外の学会動向を調査し報告する（研究の現状と課題）

専門: 暗号とサイバーセキュリティ

学歴

- 1986: 九州大学　理学部　卒業
  - 数学専攻
- 1988 : 九州大学大学院
  - 工学研究科修士課程修了
- 1993 : 博士(工学)学位(九州大学)
  - ゼロ知識証明
- 1997: (一年間)米国研修留学
  - コロンビア大学　(NY)
    - 計算機科学科 訪問研究員

職歴

- 1988 :三菱電機(株)
  - 情報電子研究所勤務 暗号理論の研究に従事
- 1995 : 九州大学工学部
  - 情報工学科助教授就任
- 2002 年 4月 九州大学大学院
  - システム情報科学研究院情報工学部門教授
- 　現在に至る
- 2018: (兼) ATR/国際電気通信基礎研究所
  - 先端セキュリティ研究室・客員研究員

@Press 2020.03.23 IEEEメンバー

サイバー社会暗号技術

『九州大学 櫻井幸一が提言』

- インターネットなどのネットワークや人工知能(AI)への不正を仕掛ける攻撃は年々高度化
- 暗号技術が重要さを増している
- セキュリティーとプライバシーとの関係性に注意すべき

- 仮想通貨(暗号資産)やそれを支えるブロックチェーン(分散型台帳)の研究者
- それら技術の不可欠な要素であるサイバー社会暗号技術

AI Safety Landscape

https://www.ai-safety.org

- About **CLAIS**: The Consortium on the Landscape of AI Safety is a global not-for-profit organization which oversees the production and use of an AI safety "view" of the current needs, challenges and state of the art and the practice of this field, as a key step towards developing an AI Safety body of knowledge.

- **Events**
  - **AI Safety**
    - **with IJCAI-PRICAI 2020**
      - Japan, Jan 2021 (Virtual)
  - **WAISE**
    - **with SAFECOMP 2020**
      - Virtual, September 15, 2020
  - **Safe AI**
    - **with AAAI 2021**
      - Virtual, Feb 7/8, 2021
  - **Safety & AI**
    - **DATAIA Workshop**
      - Paris-Saclay, Sept 23, 2020

AI Safety 2020

- **ORGANIZING COMMITTEE**
- Huáscar Espinoza, Commissariat à l Energie Atomique, **France**
- John McDermid, University of York, **UK**
- Xiaowei Huang, University of Liverpool, **UK**
- Mauricio Castillo-Effen, Lockheed Martin, **USA**
- Cynthia Chen, University of Hong Kong, **China**
- José Hernández-Orallo, Universitat Politècnica de València, **Spain**
- Seán Ó hÉigeartaigh, University of Cambridge, **UK**
- Richard Mallah, Future of Life Institute, **USA**

- **ROGRAMME COMMITTEE**
- UC Berkeley, USA, University of York, UK, Jonas Nilson, NVIDIA, USA
- University of Kentucky, USA, University of Warwick, UK
- Partnership on AI, USA, ETH Zurich, Switzerland
- University of Massachusetts Amherst, USA
- Imperial College London, UK, UC Berkeley, USA
- Affiliate at University of Oxford, China
- Intel, Germany, University of Toronto, Canada
- NASA Ames Research Center, USA, CNRS LAAS, France
- University of York, UK, Kansas State University, USA
- BMW, Germany,, CEA LIST, France, Atos, Spain
- Defence Science and Technology Laboratory, UK
- Johns Hopkins University, USA, Frazer-Nash Consultancy, UK
- IBM and University of Padova, Italy, Google DeepMind, UK
- Renault, France, LAAS-CNRS, France, Sorbonne University, France
- Tsinghua University, China
- CEA LIST, France, RISE SICS, Sweden
- Defence Science and Technology Laboratory, UK
- Broad Institute of MIT and Harvard, USA, Carnegie Mellon University, USA
- Argo AI, Germany. IBM Research, USA
- **Toshihiro Nakae, DENSO Corporation, Japan**
- University of Bristol, UK, California Institute of Technology, USA
- Safe Perspective, UK, MIT, USA, Fraunhofer ESK, Germany
- Space and Naval Warfare Systems Center Pacific, USA, Conmy, Adelard, UK
- Technical University of Munich, Germany, University of York, UK
- Johns Hopkins University, USA

日本は？

- 第2回 AI/IoTシステム安全性シンポジウム（2020. 11/10～11/12）
  - 「STAMP Workshop」（11/11）
  - 「FRAM Workshop」（11/12）
- 主催：
  - JST未来社会創造事業「機械学習を用いたシステムの高品質化・実用化を加速する“Engineerable AI”技術の開発」プロジェクト,
  - AI/IoTシステム安全性シンポジウム実行委員会,
  - 機械学習研究会（MLSE）機械学習システムセーフティ＆セキュリティWG,
  - 株式会社エヌ・ティ・ティ・データ,
  - 有人宇宙システム株式会社

- **2020年11月10日(火)**
  機械学習システムのセーフティ・セキュリティWG（MLS^3）セッション
  自動運転の事例にみる機械学習システムの安全性の課題等

  基調講演：製造業におけるAI/IoT技術を活用したDXの取組み
  浦本 直彦氏（株式会社三菱ケミカルホールディングス 執行役員）
  招待講演１：機械学習品質マネジメントガイドライン
  大岩 寛 氏（国立研究開発法人産業技術総合研究所
  招待講演２：自動運転の標準化の世界動向（仮）
  東道 徹也 氏（株式会社デンソー）
  招待講演３：Microsoft の Responsible AI への取り組み
  女部田 啓太 氏（マイクロソフト）
  講演：機械学習の品質ガイドラインと機械学習工学研究について
  石川 冬樹氏（国立情報学研究所・准教授）
  講演者によるパネルディスカション
  「機械学習応用システムの安全性は担保できるのか？」 司会：丸山 宏氏（PFN）
  交流会

  ・11/11（水）Asian STAMP Workshop
  「Asian STAMP Workshop」は、システム理論に基づく安全性分析手法STAMP（※1）
  の提唱者のMIT ナンシーレブソン教授を迎えアジアで初めてグローバルに開催します。第5回となるSTAMPワークショップ（日本語）と共に開催します。

  基調講演　ナンシーレブソン氏 （米国　マサチューセッツ工科大学教授）
  Asian STAMP Workshop（English5件、日本語5件、ショートトークセッション）

  ・11/12(木）FRAM　Workshop（※2）
  「FRAM Workshop」は、昨年に引き続きレジリエンス・エンジニアリングの提唱者
  のエリック・ホルナゲル教授を迎え、議論します。
  レジリエンスエンジニアリング発表(3件)
  基調講演　エリック・ホルナゲル氏 （スウェーデン　ヨンショーピング大学教授）

AI

Safety, Security, Privacy, Trust, Dependability,

- 13th ACM WORKSHOP ON ARTIFICIAL INTELLIGENCE AND SECURITY
  - November 13, 2020 — Orlando, USA
  - with the 27th ACM Conference on Computer and Communications Security

- 3rd DEEP LEARNING AND SECURITY WORKSHOP
  - with the 41st IEEE Symposium on Security and Privacy
  - May 21, 2020

Safety, Security, Privacy, Trust, Dependability,

- Safety and/or/vs Security
- Dependability vs. Security
- Privacy vs. Security
- Trust vs. Security

- **関係と違いは？**

Security vs Safety
日本語ではどちらも“安全”と呼びますか？
安全・安心？？

- *Attack with Avdersary*
  - **悪意の攻撃者**
  - Cf. Natural Disasters/自然災害
    - ハッカー vs. 地震
    - セーフティ・ベルト
      - セキュリティベルト（？）

Dependability and/vs Security

- 南谷（東大名誉教授）

@インドIIIT-dmj集中講義

Privacy and Security: What's the Difference? [March 5, 2019]

https://it.umn.edu/news-alerts/news/privacy-security-whats-difference

Security => Confidentiality ? <= Privacy

Security

- How information (data) is protected at all stages
- Defines the "how" of protecting data
- Sets up safeguards and controls to allow or restrict
- access to data, protecting from unauthorized disclosure, theft, alteration, or loss

Privacy

- The right to keep personal information from being accidentally or maliciously disclosed, or from unauthorized access
- Defines the "who, what, and when" of protecting data
- Outlines the conditions under which information can be accessed, used, or shared
- Establishes the **right** for information to be protected

Security and/or/vs Trust

安全/安心（？）/信頼

ESCAR: Europe(2003~ 19th 2021Nov.) /USA(8th 2021) Asia(2014~: 7th 2021)

- **E**mbedded **S**ecurity In **C**ars Conference
- The World's Leading **Automotive Cyber Security** Conference!
- started in 2003 in Germany
- Its founder and organizer, isits AG was consistently supported by its event partner ESCRYPT.

- Escar欧州19th　(2021 Nov)
  - Hybrid: Virtual & Frankfruit
- Escarアジア　2021？2020??
  - 2019: 東京 (11月　品川)
- Escar USA
  - 2020 canceled

38th SCIS2021@電子情報通信学会
暗号と情報セキュリティシンポジウム

- 自動車セキュリティ(1)&(2)
  - 車線検出機能に対する色調改変攻撃とその対策(横浜国立大学)
  - 車載制御システム向けサービス探索の脅威と保護手法(名古屋大学) & (オートネットワーク技術研究所)
  - SOME/IPの周期的なサービスにおける異常検知手法の検討(住友電気工業株式会社)
  - 車載Ethernet向けフローベースIDSの提案：SOME/IPへの攻撃例に対するフロー分析(パナソニック株式会社)
  - Intelligent Impact Assessment for Product Security Incident Response Team in the Automotive Industry (Hitachi Ltd.)
  - 車載 Event Data Recorder のデジタル・フォレンジックに関する調査及び検証 (警察大学校)& (名古屋大学) & (パナソニック)
- AI セキュリティ(1)---(4)
  - 説明可能なAIに対するデータ収集を必要としないModel Stealing攻撃 (NTTセキュアプラットフォーム研究所)
  - ランダムパターンノイズを用いたブラックボックス Adversarial Examples攻撃 (茨城大)
  - Adversarial ExampleのTransferabilityに基づく特徴抽出層の同一性判定　(NTTテクノクロス株式会社/情報セキュリティ大学院大学)
  - 深層学習モデルで安全に推論するためのモデルパラメータ暗号化の検討(立命館大学)
  - Trusted Execution Environmentによる省メモリな深層学習モデル保護方式(三菱電機株式会社)
  - 時系列をもつマルウェアデータセットにおける有効な特徴の変化の調査
  - (トレンドマイクロ株式会社/情報通信研究機構)、津田侑(情報通信研究機構)

真の背景/目的は国内の***学会にAI-Security研究会を立ち上げること：その準備としての国内外の学会動向を調査し、報告する（産学官連携の現状と課題）

目標
AIセキュリティ研究会発足へ向けて

- **研究分野の特質**
  - **AI vs. Security**

研究会立ち上げ準備状況

賛同者/幹事協力者を求む
現在（２名）
溝口 誠一郎@DNV GL
ビジネス・アシュアランス・ジャパン Cybersecurity Laboratory スペシャリスト

& 櫻井
（sakurai@inf.kyushu-u.ac.jp）

ASIACCS 2022

17TH ACM SYMPOSIUM ON INFORMATION,
COMPUTER AND COMMUNICATIONS SECURITY

17th ACM ASIACCS 2022. May 長崎出島メッセ

Wrap Up
研究：継続は力なり

- **石/暗号の上にも３年/33年　！**
  - "a rolling stone gathers no moss"
  - persistence pays off
  - slow and steady wins the race

- 1988: 人工知能
- 1990: 量子計算機
- 　-- 携帯電話
- 今や：自動運転
  - 機能安全からサイバー攻撃対策
  - 海➔ 陸➔ 空

Take-Away
{海, 陸} から空へ