（別紙様式２－２）

宇宙地球フロンティア実地研修 報告書

Report for Onsite Training in Earth-Space Frontier Science

氏名/Name	中山　悠平
所属部局/ Affiliation	理学系　研究科　物理学　専攻 Department of , Graduate School of
研究機関・企業名 /Hosting Institution	NTT 研究所, 日立製作所
期間/Period	2022 年 8 月 1 日　～　2022 年 9 月 16 日　＊西暦で記入 2022 / 8 / 1　～　2022 / 9 / 16　mm/dd /yyyy

私の現在の専門は素粒子理論であるが，普段から数値計算や研究室の Web サイト運営などで IT 技術に慣れ親しんでいることもあり，以下の企業インターンシップに参加してその分野の研究開発に従事した．

1. ソフトウェアに混入した悪意あるコードの検出手法の検討（NTT 研究所）
   オープンソースソフトウェア（OSS）は現代社会を支える重要な技術の一つであるが，近年その利用者を標的とした攻撃が発生しており問題となっている．攻撃者は様々な経路でソフトウェア・サプライチェーンに侵入し，利用者の許諾を得ない形でバックドアを仕掛ける．パッケージの利用者が多いほど被害が大きくなる一方で，有名な OSS ほど規模が大きくパッケージを全数検査する人的・時間的リソースがないというジレンマを抱えている．そのためバックドアを自動的に検知し効率的に排除する手法の開発が望まれている．

   このインターンシップでは，データセットとしてバックドアを含むパッケージ群と含まないパッケージ群を用意し，ソフトウェアの静的解析技術によって次の２つを行った．

   1. 先行研究で示されたバックドアの特徴に基づいた検知ルールを定義し，バックドアを正しく検知できる割合（True positive rate）と誤ってバックドアと判定してしまう割合（False positive rate）を割り出す．
   2. 新たな検知ルールを独自に定義し，同様に true positive rate と false positive rate を割り出してルールの評価を行う．

   今回新たなルールを幾つか提案したが，その中で最良のものは false positive rate を 1%程度に抑えながら 60%近くのバックドアを検知することに成功した．これは先行研究と比較して大幅な改良となっている．

   またこのインターンを通して得られた知見や成果について 30 分程度の発表を行った．内容のみならず，プレゼンの分かりやすさについても好評であった．

2. ハイブリッドクラウドストレージサービスの研究開発（日立製作所）
   近年 IT インフラではクラウドを活用する事例が増えている．日立製作所では，従来の強み技術であるストレージサービスをベースに，クラウド活用の促進に向け，日立のストレージとクラウドとの透過的なデータ管理・運用を可能なハイブリッドクラウドソリューションの提供を進めている．インターンシップでは，ハイブリッドクラウドストレージの構築や運用に必要な要素技術の調査，および自動化手法の検討を行った．特に行ったのは以下の 3 点である:
   1. ハイブリッドクラウドストレージの構築・運用に必要なネットワーク要件の調査

   2. ハイブリッドクラウドストレージを構築・運用した場合に生じる問題点の指摘とその解決策の提案
   3. ハイブリッドクラウドストレージの構築・運用自動化プロトタイピング

   また，インターンシップで得られた知見と成果について 20 分程度で発表を行った．

Although I am currently majoring in particle physics, I participated in the following internships and was engaged in R&D in the IT industry.

1. Towards detecting malicious injected code in software (NTT R&D)
   The Open-Source Software (OSS) is one of the most important technologies to contribute to the development of modern society. Recently, many people have reported attack on the OSS users and developer's community think of them as thread to be purged. The attackers hack into a software supply chain and inject backdoors into the victim's computer without any permission of them. While injection into packages having many users will result in a severe damage, the community does not have enough manpower and time to check all the source codes in the package. Therefore, the developers and end users desire methods to realize an automatic detection and removal of backdoors.
   In this internship, I performed the following two tests on the dataset composed of malicious and benign software based on the static analysis on the source code.
   1. I define detection rules based on features of backdoors shown in literatures and estimate these rules on true/false positive rates.
   2. I define some original detection rules and compute true/false positive rates in the same way.
   Among the rules I suggested, the best one succeeded in detecting around 60% of malicious packages with the false positive rate about 1%. This result shows the new rule is far better than that of literatures.
   In addition, I gave a presentation about knowledge and achievement I gained in this internship in 30 minutes. I got a high evaluation in not only the contents but also the clarity of the presentation.
2. Research and Development on Hybrid cloud storage services (Hitachi)
   Recently, the use of cloud computing in IT infrastructure has been increasing. To promote it, Hitachi is going to provide the hybrid cloud storage solution based on Hitachi's excellent storage technologies, which enables us to manage data transparently between cloud services and Hitachi's storage. In this internship, I investigated underlying technologies to build and manage hybrid cloud storages and considered how to automate them. Especially, I worked on the following three points:
   1. To investigate network requirements for building and operating hybrid cloud storage services
   2. To point out issues occurring when building and operating hybrid cloud storage services and suggest solutions to them
   3. To write a protype code to build and operate hybrid cloud storage services automatically

   I gave a presentation about knowledge and achievement I gained in this internship in 20 minutes.