Про исследования и реверс

О чём речь?

- Роль исследователя в ИТ не выражена явно
- В банальном ИТ (CRUD и сайты на wordpress) её выполняет техлид/тимлид/техдир
- Она явно выражена лишь в сложных проектах с Research & Development составляющей

О роли исследователя

- Это технолог
- Он делает сложное простым
- Он ищет способы решения проблем

Задачи исследований

- поиск решения или технологии
  - когда есть готовое
  - когда нет готового (полностью новое)
- поиск фундаменталки под технологию (есть исходники, но непонятно как оно работает)
- Итог исследований - статья и/или proof of concept

Задачи исследований

Reverse engineering, или обратная разработка — один из частных случаев исследований

Задачи реверса

- анализ малвари
  - снятие защит (больше почти не встречается)
- поиск уязвимостей
- реверс прошивок
- восстановление знаний (утеряны исходники, утеряны знания по проекту)
- интеграция с закрытым ПО
- конкурентная разведка (анализ решений конкурента)

Обзор рынка

- Позиция Researcher — это среднее между аналитиком, проектировщиком, экспертом по алгоритмам и технологом
- Позиция Reverse Engineer — таких мало, в основном ИБ. Часто требуется такой навык для других позиций

Качества исследователя

- Начитанность — чем больше кода классических проектов, тем лучше (ядро Linux, утилиты GNU, заголовки библиотек, исходники игр, итд)
- Архитектор наоборот
- Этому учат?
- Это продукт личного поиска и целеустремлённости

Исследования в Enterprise

- К концу третьего дня на новой работе вам дали доступ к git
- Вы скачали очень интересный, но крайне непонятный проект
- Поздравляю! Вы — исследователь!
- Ближайшие недели и месяцы вы потратите на поиски смысла в этом проекте

Манифест Agile

Мы постоянно открываем для себя более совершенные методы разработки программного обеспечения, занимаясь разработкой непосредственно и помогая в этом другим. Благодаря проделанной работе мы смогли осознать, что:

- *Люди и взаимодействие важнее процессов и инструментов*
- ***Работающий продукт важнее исчерпывающей документации***
- *Сотрудничество с заказчиком важнее согласования условий контракта*
- *Готовность к изменениям важнее следования первоначальному плану*

То есть, не отрицая важности того, что справа, мы всё-таки больше ценим то, что слева.

https://agilemanifesto.org/iso/ru/manifesto.html

Качество проекта

Наличие непосредственно в репозитории:

- README

Менее важно, но показатель:

- ChangeLog
- Roadmap
- Комментарии в коде с описанием мотивов решений, принципов работы сложных кусков

В энтерпрайзе этого я не видел никогда

Восстановление знаний

В принципе, для этого есть всё:

- Git
- Трекер задач
- База знаний
- Люди

Но данные в них несвязны и мозаичны, и не всегда есть все знания. Это — теневое ИТ

Восстановление знаний

Основные вызовы:

- Восстановление мотивов
- Поиск и объяснение эвристик
- Анализ API
- Описание архитектуры

Восстановление знаний

Итогом работы по восстановлению знаний должен быть README с:

- Инструкцией по сборке
- Инструкцией по развёртыванию
- Кратким описанием задач и архитектуры программы

Методика исследований

- Постановка задачи: на какие вопросы нужно ответить? если затрудняемся, то вопросы:
  - Как это устроено?
  - Что здесь происходит?
- Рабочий журнал (воспроизводимость результата)
- Исследование — это итеративный процесс
- Откладываем отладчик на самое потом

Методика исследований

- Такую задачу уже решали!
- Ищем статьи, монографии на тему
- Ищем программы-аналоги с исходниками
- Как бы мы решили данную задачу? Проектируем сами

Методика исследований

- Читаем справку по программе
- Смотрим окрестности: конфиги, файлы с ресурсами (.xml .json .yaml и прочие)
- А давайте спросим у разработчика?

Методика исследований

Зацепки:

- Строки
- Сетевой трафик
- Поведение (файлы, реестр)

Методика исследований

Зацепки:

- API (интерфейсы в .h, импорт, экспорт модулей)
- Точки расширения

Инструменты

- Windows: утилиты Русиновича
- Linux: полно встроенных инструментов, начиная с /proc

Инструменты

binwalk: определение структуры и содержимого файла

Инструменты

Отладчики:

- gdb
- WinDbg
- Visual Studio
- IDA Pro тоже отладчик

Инструменты

Снифферы API:

- strace
- API Monitor

Инструменты

Снифферы трафика:

- WireShark
- Tcpdump
- mitmproxy

Инструменты

Зависимости:

- depends
- ldd

Инструменты

Дизассемблеры и декомпиляторы

- IDA Pro
- dotPeek
- деобфускаторы js
- Ghidra (?)

Конец теории