**Ideas:**

* Анализ состояний системы на основе различных уровней мониторинга (pcap + logs analysis).

**-** Формулировка метрик нефункционального тестирования с точки зрения задачи машинного обучения.

**Problem:**

Хочется получить информацию о состоянии системы за определенный период времени. Как это сделать? Проанализировать логи системы. По этому поводу написано большое количество работ: как распарсить[], структурировать[], извлечь features[], натравить алгоритм машинного обучения (supervised[] или unsupervised[]), убедиться что прогноз аномалий адекватный и в конечном счете как-то визуализировать и интерпретировать результат[].  
 Такой подход может хорошо сработать для систем которые прошли тщательное тестирование и можно предположить что система логирования работает корректно (опять же в ней могут быть ошибки). Поэтому единственный (не очень надежный) источник данных для постоения систем поиска аномалий (классический или с использованием машинного обучения) не является надежным (особенно для критически важный систем). И уж тем более такой подход не приемлем для систем которые находятся на этапе тестирования.  
 Также существует подход основанный на парсинге source code системы для извлечения операторов логирования и на их основе структирировать логи [][]. Такой способ может быть эффективнее чем оперирование только с логами системы, но таит в себе две проблемы: 1) В случае независимого тестирования сторонняя компания не имеет доступа к исходному коду (и вообще редко команды тестирования видят исходные коды) 2) В самом исходном коде уже могут содержаться ошибки.

**Solution:**

Использовать для анализа несколько независимых источников данных. Например анализ логов + анализ pcap файлов (+ какие-то метрики системы, память, процессор). Применение такого комплексного подхода позволяет избежать использование единственного источника информации о состоянии системы. А так же найти дефекты в механизмах мониторинга и логирования самой системы.

**Our contribution:**

Our contribution is a general methodology that allows machine learning and information retrieval techniques to be applied for analysis the state of the under testing system using several undependent data sources.

**Abstract**