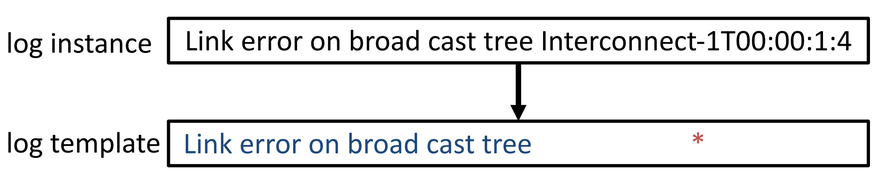
Approach to Better Log Template Generation

Yuya Yamashiro[1], Satoru Kobayashi[2], Kensuke Fukuda[2], Hiroshi Esaki[1]
yuya@hongo.wide.ad.jp, sat@nii.ac.jp, kensuke@nii.ac.jp, hiroshi@wide.ad.jp

1: University of Tokyo, 2: National Institute of Informatics

Introduction

- **Syslog** is widely used for system management
- Huge volume to handle manually
  - 70,000 lines / day (SINET4)
- No “normative grammar”
  - Difficult to extract information from logs
- Translate from raw data to **Log Template**
  - Based on **Software Information**
  - Based on Log Data



Goal

- Automatically generate more accurate template
  - From **open source code**
- Help system operators to extract important information

Dataset

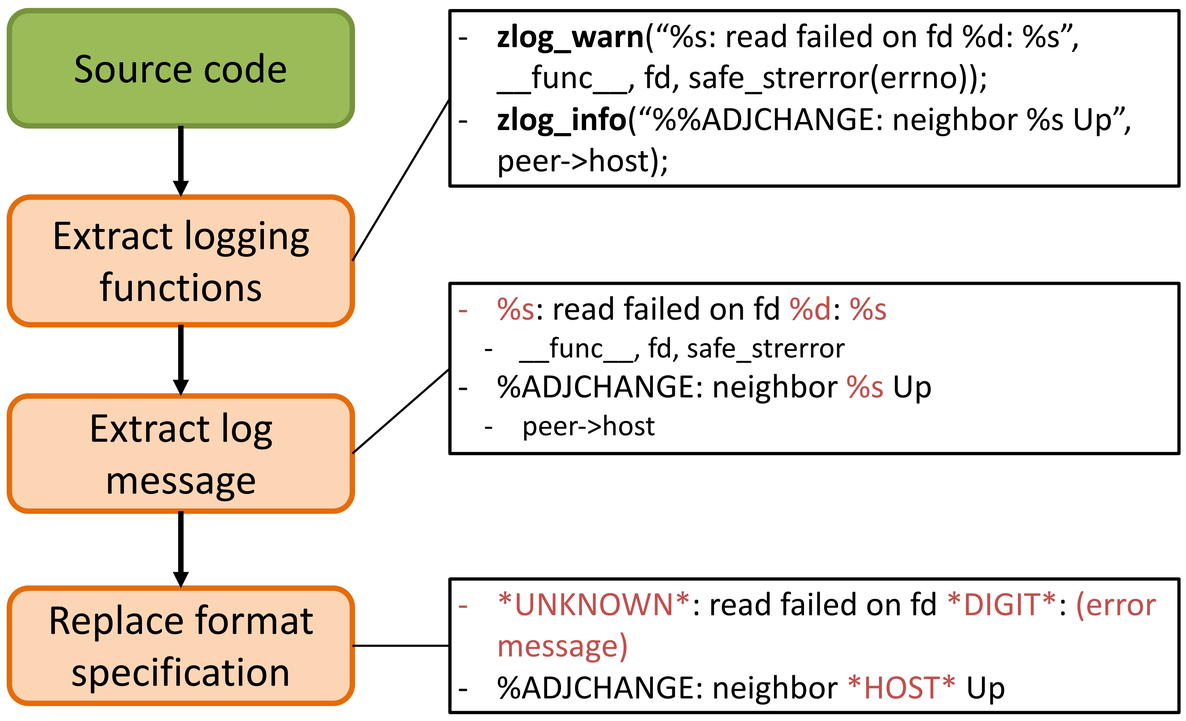
- Vyatta 1.0.8 (napa): Network OS
  - Vyatta Kernel (kernel)
  - Vyatta-quagga (bgpd, ospfd, …)
  - Net-snmp (snmpd)
  - OpenSSH (sshd)
  - Ntp (ntpd)
  - Other many softwares
- Actual vyatta log data of APAN-JP
  - May-June 2018
  - 277,034 lines

Logging Functions

Software	Logging Function
Vyatta Kernel	printk, pr_*
Vyatta-quagga	zlog_*, plog_*
Net-snmp	snmp_log_*, NETSNMP_LOGONCE, DEBUGMSGTL
OpenSSH	fatal, error, sigdie, logit, verbose, debug, debug2, debug3, pam_syslog, helper_log, authlog
Ntp	msyslog

- Difficult to make templates from source code completely automatically

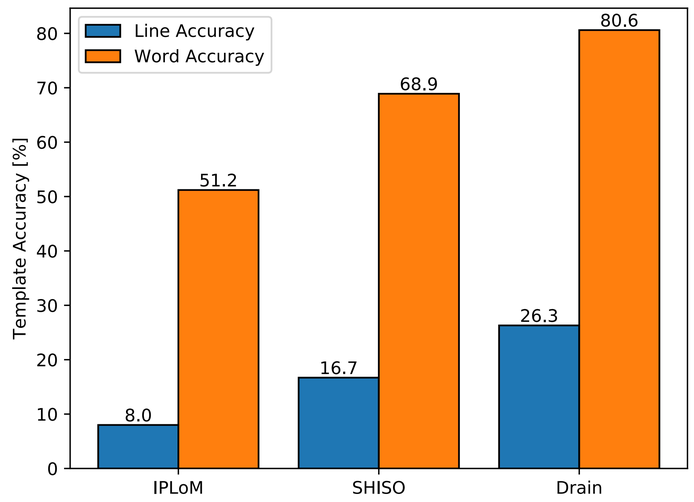
Templates from Source Code



- Total: 189,037 templates

Application of Generated Templates

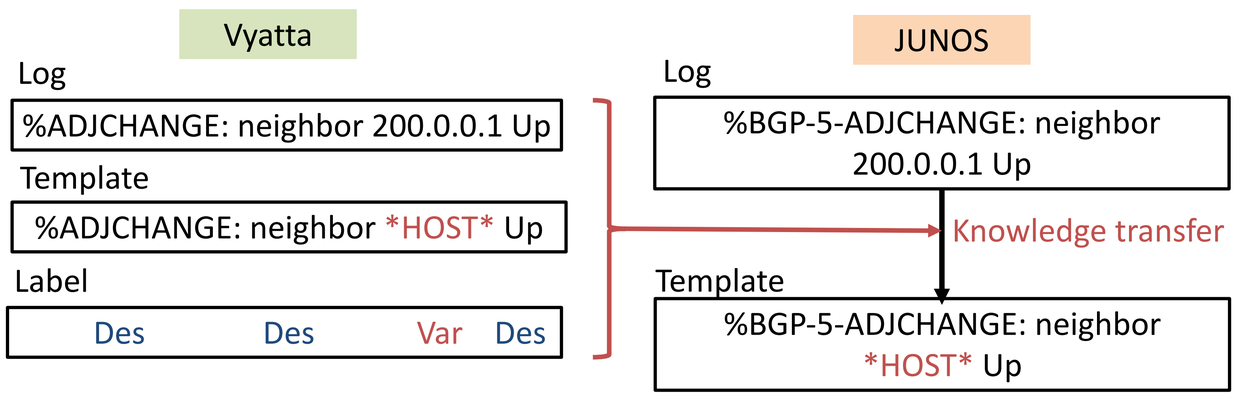
1. Log templates’ accuracy evaluation[1]



   a. State-of-the-art clustering based generation algorithms (IPLoM[2], SHISO[3], Drain[4])

   b. Use our generated templates as ground truth

2. Transfer learning : from vyatta to junos



References

[1] 山城裕陽, 他 “ソースコードからのネットワークログテンプレート自動生成に関する検討”, 電子情報通信学会IA研究会, p.8, 札幌, 2018
[2] Makanju, A., N. Zincir-Heywood, and E. E. Milios. "Iplom: Iterative partitioning log mining." Tech. Rep. CS-2009-07 (2009).
[3] Mizutani, Masayoshi. "Incremental mining of system log format." Services Computing (SCC), 2013 IEEE International Conference on. IEEE, 2013.
[4] He, Pinjia, et al. "Drain: An online log parsing approach with fixed depth tree." Web Services (ICWS), 2017 IEEE International Conference on. IEEE, 2017.