SAND2021-8839PE

The Center for Cyber Defenders

Expanding computer security knowledge

RAMSeS : GroundTruth

Matthew Robinson
New Mexico Tech

Project Mentors: Samuel A. Mulder, Org. 1462; Ryan Vrecenar, Org. 5823

Problem

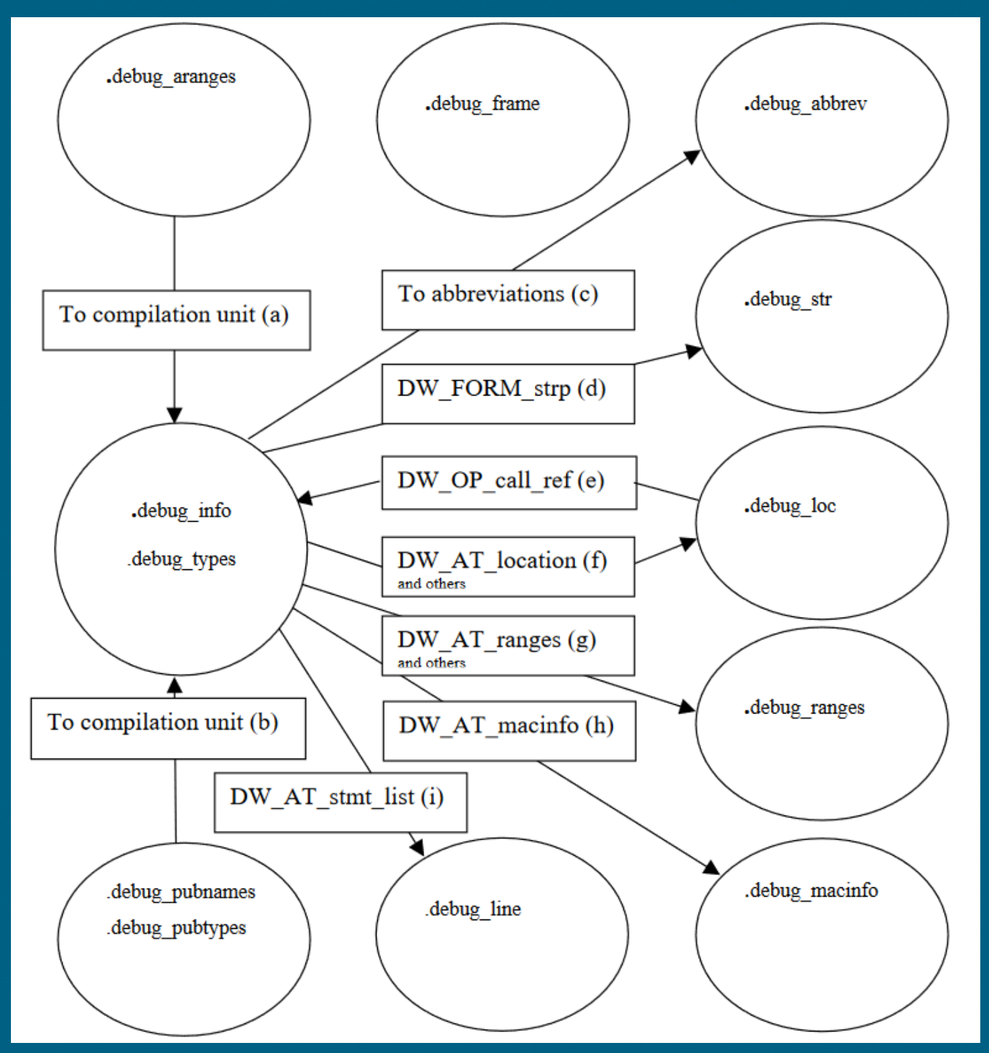
Binary analysts gain confidence from the ability to account for every byte of a binary and their purpose. Currently when given a binary, there is a lack of central information detailing what everything is.

Objective

- Create modules that can parse a binary to map bytes of a file to their meaning
- Parse binary bytes not directly tied to code to gain information
- Specifically, create a module that can parse DWARF symbols, display their interpreted meaning, and be used to gain further information

Approach

- Use DWARF documentation to parse respective binary bytes
- Parse every DWARF section and all it's contained information
- Use this to learn high-level information about the program



- **The .debug_info section**
- This section is, along with .debug_abbrev, are the main DWARF sections that describe how to decode other sections, or have information contained directly within.

```
    <10>   DW_AT_language    : 4    (C++)
    <11>   DW_AT_name        : (indirect string, offset: 0xb9): test.C
    <15>   DW_AT_comp_dir    : (indirect string, offset: 0x9e): /home/ma
    <19>   DW_AT_ranges      : 0x0
    <1d>   DW_AT_low_pc      : 0x0
    <25>   DW_AT_stmt_list   : 0x0
<1><29>: Abbrev Number: 2 (DW_TAG_typedef)
    <2a>   DW_AT_name        : (indirect string, offset: 0x2f): size_t
    <2e>   DW_AT_decl_file   : 2
    <2f>   DW_AT_decl_line   : 216
    <30>   DW_AT_type        : <0x34>
<1><34>: Abbrev Number: 3 (DW_TAG_base_type)
    <35>   DW_AT_byte_size   : 8
    <36>   DW_AT_encoding    : 7    (unsigned)
    <37>   DW_AT_name        : (indirect string, offset: 0x219): long ur
<1><3b>: Abbrev Number: 3 (DW_TAG_base_type)
```

- **The .debug_abbrev section**
- Below is a dump of what's contained in the .debug_abbrev section. It contains a series of abbreviation codes and their corresponding attribute names and values. Each attribute name and value pair helps to describe information about the source level binary.

```
abbrev_code 1 tag 0x11 has_children 1
attr_name:  0x25 attr_form:  0xe
attr_name:  0x13 attr_form:  0xb
attr_name:  0x3 attr_form:  0xe
attr_name:  0x1b attr_form:  0xe
attr_name:  0x55 attr_form:  0x17
attr_name:  0x11 attr_form:  0x1
attr_name:  0x10 attr_form:  0x17
abbrev_code 2 tag 0x16 has_children 0
attr_name:  0x3 attr_form:  0xe
attr_name:  0x3a attr_form:  0xb
attr_name:  0x3b attr_form:  0xb
attr_name:  0x49 attr_form:  0x13
abbrev_code 3 tag 0x24 has_children 0
attr_name:  0xb attr_form:  0xb
attr_name:  0x3e attr_form:  0xb
attr_name:  0x3 attr_form:  0xe
abbrev_code 4 tag 0x24 has_children 0
```

Results

DWARF symbols contain quite a bit of information on source level programs. Things like variable names, classes, source language, function names, and more can be buried in the binary bytes. Using this information, it's possible to map functions to their addresses.

Impact and Benefits

This work aids binary analysts by helping them understand and account for every byte in a binary file.

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC., a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525. SAND2017-XXXXC.

U.S. DEPARTMENT OF ENERGY

National Nuclear Security Administration

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.