



**HAL**  
open science

## Specification-Based Protocol Obfuscation

Julien Duchene, Eric Alata, Vincent Nicomette, Mohamed Kaâniche, Colas Le Guernic

► **To cite this version:**

Julien Duchene, Eric Alata, Vincent Nicomette, Mohamed Kaâniche, Colas Le Guernic. Specification-Based Protocol Obfuscation. DSN 2018 - 48th Annual IEEE/IFIP International Conference on Dependable Systems and Networks, Jun 2018, Luxembourg City, Luxembourg. pp.1-12, 10.1109/DSN.2018.00056 . hal-01848573

**HAL Id: hal-01848573**

**<https://inria.hal.science/hal-01848573v1>**

Submitted on 24 Jul 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Specification-based Protocol Obfuscation

Julien Duchêne  
CALID, Paris, France  
& LAAS-CNRS,  
Univ. de Toulouse, CNRS, INSA,  
Toulouse, France  
julien.duchene@intradef.gouv.fr

Eric Alata, Vincent Nicomette,  
and Mohamed Kaâniche  
LAAS-CNRS,  
Univ. de Toulouse, CNRS, INSA,  
Toulouse, France  
firstname.lastname@laas.fr

Colas Le Guernic  
DGA Maîtrise de l'Information  
Rennes, France  
& Univ. Rennes, Inria, CNRS, IRISA  
Rennes, France  
colas.le-guernic@intradef.gouv.fr

**Abstract**—This paper proposes a new obfuscation technique of a communication protocol that is aimed at making the reverse engineering of the protocol more complex. The obfuscation is based on the transformation of protocol message format specification. The obfuscating transformations are applied to the Abstract Syntax Tree (AST) representation of the messages and mainly concern the ordering or aggregation of the AST nodes. The paper also presents the design of a framework that implements the proposed obfuscation technique by automatically generating, from the specification of the message format, a library performing the corresponding transformations. Finally, our framework is applied to two real application protocols (Modbus and HTTP) to illustrate the relevance and efficiency of the proposed approach. Various metrics recorded from the experiments show the significant increase of the complexity of the obfuscated protocol binary compared to the non-obfuscated code. It is also shown that the execution time and memory overheads remain acceptable for a practical deployment of the approach in operation.

## I. INTRODUCTION

Reverse engineering is aimed at extracting knowledge from a component that is, a priori, complex to understand, in order to infer its main characteristics and behavior. It is used for many different purposes, both by legitimate people or attackers. For instance, attackers motivations could be to steal intellectual property and generate counterfeit, whereas legitimate people use reverse engineering to analyze malware in order to develop protection countermeasures. The target of the reverse engineering may be for instance a binary program or a communication protocol. In this paper, we are mainly concerned by the development of efficient countermeasures against malicious protocol reverse engineering activities. Several complementary solutions are available to fulfill this objective, such as cryptography or obfuscation. An obfuscation is a transformation applied on a component (either a software or a communication protocol) to make the inference of the transformed component behavior difficult without knowing its specification. Of course, the transformed component must still ensure the service for which it was developed. Inevitably, reverse engineering and obfuscation activities are closely linked.

This paper focuses on the obfuscation of communication protocols. Several solutions have been proposed recently, based e.g., on randomization, mimicry or tunneling techniques with the objective to make the communication indistinguishable from noise or other protocols (see e.g., the discussion of

related work in [1]). Most of these techniques have been developed in order to circumvent network censorships. However, the proposed transformations have not been designed to provide enhanced protection against communication protocols reverse engineering. Furthermore, the obfuscations are integrated a posteriori in the binary. They are implemented through a dedicated function between the transformation layer and the core application, that can be easily identified by an attacker to understand the obfuscation logic.

The main objective of this paper is to present a new protocol obfuscation technique that is aimed at increasing the effort needed by an adversary, having access to network traces or to the application binary, to successfully reverse the protocol. For that purpose, the transformations are applied to the specification of the protocol, focusing on the message format. The transformations are, by construction, invertible to avoid ambiguities when the messages are parsed. We are not aware of similar obfuscation techniques that operate at the protocol specification level.

Cryptography could be another solution. Indeed, it guarantees several security properties including confidentiality. Confidentiality does imply protection against protocol reverse engineering. However, confidentiality is lost if the attacker can intercept the buffer before encryption in the process memory. In that case our approach offers some additional protection. Finding a single buffer with a very specific access pattern is arguably easier than reversing the code or the message format produced by our approach. Note that a higher level of protection can also be obtained by combining both techniques: e.g., messages can be obfuscated before being encrypted and sent through the secured communication channel.

To implement our new technique, we developed a framework with the following design characteristics: 1) the framework automatically generates, from the specification of the message format, a library code performing the transformations, that can be easily linked to the core application to provide an obfuscated binary; 2) this library code can be easily regenerated with new transformations, at regular intervals, to produce new versions of the obfuscated core application; 3) the generated code is designed to make the protocol difficult to reverse for an attacker that would capture network traces or reverse the binary code of the application itself. One of the objectives of the framework is to make the interface

between the transformation layer and the core application difficult to identify and understand by an attacker. Moreover, the framework generates obfuscated protocols that behave according to non regular models which are known to be difficult to reverse by existing reverse engineering tools.

We applied the proposed obfuscation framework to two application protocols (Modbus and HTTP). Various metrics are presented to illustrate the significant increase of the complexity of the obfuscated protocol binary compared to the non-obfuscated code. It is also shown that the execution time and memory overhead remains acceptable for a practical deployment of the approach in operation.

This paper is organized as follows. Section II presents basic background about protocol reverse engineering methods, associated tools and inference models. It also discusses the main challenges faced by reverse engineering analysts. Then, Section III discusses some related work addressing obfuscation techniques and outlines the main motivations and original characteristics of our obfuscation techniques. Section IV and Section V respectively present the architecture of the framework we designed to implement our obfuscation technique and a detailed description of the main transformations applied to the message format specification supported by this framework. Section VI presents and justifies some choices we made for the implementation of the framework and Section VII presents the different experiments we have carried out in order to assess the relevance of our obfuscation technique. Then Section VIII concludes and outlines some future work.

## II. PROTOCOL REVERSE ENGINEERING (PRE)

Reverse engineering is the process of analyzing a subject system to extract relevant knowledge or design information about its components, their interrelationships and behavior, and to create representations of the system based on the extracted information [2]. Historically, it has initially targeted hardware products and then its main concepts were applied to software applications and communication protocols. Software applications reverse engineering mostly applies to "closed-source" programs and usually requires to disassemble the application binary with tools such as IDA [3] or radare2 [4].

In this paper we focus on communication protocols reverse engineering (PRE). PRE is the process in which protocol parameters, format, and semantics are inferred in the absence of the formal protocol specification [5]. It can be achieved by focusing either on: the vocabulary (types of messages which can be exchanged), the message format (encoding language of message types) or on the protocol grammar (encoding language of message exchanges). PRE is useful in many domains such as interoperability, protocol simulation, security audits or conformance testing. Unfortunately, it is also useful for attackers to steal intellectual property or to make counterfeit software. PRE also raised some legal concerns. These are not discussed in this paper. The remainder of this section presents: i) the different methods used to perform communication protocols reverse engineering, ii) some state-of the art PRE tools, and iii) the associated challenges.

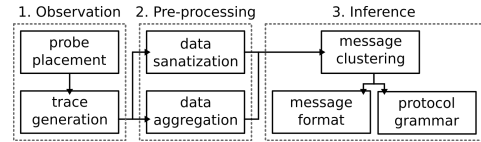


Fig. 1. Protocol reverse engineering steps.

### A. PRE methods

In order to reverse engineer a protocol, an analyst needs to have access either to a network protocol execution trace, or to the application binary. The analysis of the traces is carried out by so-called "network based inference" techniques. Binary analysis, carried out by so-called "application based inference" techniques, focus on the instructions of the binary that parse or generate messages. It can be done using static code analyses or dynamic analyses if the binary can be properly executed to trigger communications.

The reverse engineering activity is divided into several steps. These steps are quite similar for both network based and application based inference tools. They are summed up in figure 1.

The first step, called observation, is aimed at gathering raw information resulting from the protocol execution. Probes are placed to collect data the less noisy possible according to the method used by the reverser. For instance, a network probe monitors traffic that is fully encapsulated in many protocols (IP, TCP, etc.). On another side, a probe deployed in the application, using a debugger, can dump messages without any noise. If data are noisy, a preprocessing step is required. With network traces, this preprocessing step consists in removing the consequences of network protocol encapsulations, using data sanitization and data aggregation. For instance, some network traces that have been fragmented by the TCP layer must be aggregated in order to retrieve messages. The last step is dedicated to the inference process that begins by the classification of sanitized messages into different classes, representing different message types. Finally, either a message format inference is done on each message class, or a protocol grammar inference is done on sequences of message types. This last step is based on language learning algorithms.

### B. PRE tools

Various surveys of protocol inference tools are available [6], [5], [7]. Before 2004, PRE was mainly performed manually. It was error prone and time consuming. In 2004, PI Project PRE tool [8], [9] proposed a sequence alignment algorithm for message classification and message format inference based on network traces. Shortly afterwards, several tools were developed using this algorithm while inference algorithms based on regular languages were used to retrieve the protocol grammar. For instance, ReverX [10] uses regular language inference algorithm for both message format inference and grammar inference. Netzob [11], [12] uses active inference to guess the message semantics.

The number of application based inference tools is more important. The main tools are FFE/x86 [13], Dispatcher [14], [15], [16], Prospex [17] and MACE [18], [19]. FFE/x86 is based on a static analysis of the application to retrieve messages format as a hierarchical finite state machine. Polyglot [20] introduces dynamic binary analysis for message format inference. This technique was widely used and improved by following tools. Prospex measures the impact of message processing on the system to classify the messages and infer the protocol grammar with classic regular language learning algorithms. MACE infers the protocol grammar based on a symbolic execution using a regular model.

Almost all PRE tools rely on regular models to retrieve the protocol specification (message format and protocol grammar). In addition, the message classification step is important for a coherent format inference.

### C. Challenges

In the following, we focus on some of the challenges faced during the protocol reverse engineering process, that we have considered in our study to guide the selection of proposed obfuscation approach.

1) *Observation*: The placement of probes to capture relevant information required for protocol reverse engineering is critical. When the application uses a cryptographic library, most of the time, the interface between this library and the core application is easy to locate and understand. So, it is still possible through this interface to dump messages using a debugger and hooks on the interface. Recent work [16], [21] has introduced techniques to automatically identify the cryptographic library and to perform PRE on encrypted protocols. Thus, making the placement of such probes difficult for a reverser will make the reverse engineering of message format more complex. This objective can be fulfilled by ensuring that the code used for the generation of the messages is not easy to identify by the reverser. Serialization projects naturally answer to this requirement as they provide an interface based on accessors (setters and getters) to manipulate data stored in an internal abstract representation.

2) *Fields delimitation*: When performing message format inference, fields delimitation is generally based on a sequence alignment algorithm and well known delimiters like '\r\n', '\0' or 'SP'. Thus, the PRE process will be more tedious if the delimiters are removed. Furthermore, sequence alignment algorithms are very efficient when applied to messages of the same types, as these messages have many sub-sequences in common. If messages of the same type do not fulfill this property, the classification will be more complex.

3) *Classification*: Classification in PRE is mainly based on similarity measures. It is a key step in PRE as the efficiency of the inference depends on the quality of this classification. This quality can be degraded if 1) two messages of the same type seem different or 2) if two messages of different type seem very close. In the first situation, the number of classes obtained after the classification exceeds the real number of message types. In the second situation, the number of classes

is lower compared to the effective number of message types. With a mix of the two approaches, the classification is likely to provide meaningless classes.

4) *Inference models*: To perform message format inference, most PRE tools rely on regular models (automata, trees, etc), that are possibly annotated to represent dependencies such as a field which is the length of another field. Therefore, PRE tools are likely to be less efficient when the message formats are not regular. The inference algorithm may not converge, or it may lead to overfitting (the model accepts a message that does not belong to the protocol) or underfitting (the model doesn't recognize messages that belong to the protocol).

## III. OBFUSCATION BACKGROUND

The objective of a program obfuscation is to make it "unintelligible" while preserving its functionality [22], [23], [24]. It is implemented by means of a set of transformations that are used to transform a component P (a software or a communication protocol) into an equivalent component P' (providing the same service) such that the behavior of P' is less understandable than the behavior of P, without having their specification. To obfuscate a communication protocol, these transformations can be applied to the application implementing the protocol itself, or on the way messages are transmitted. The chosen transformations must be adapted to the considered attacker model.

### A. Software obfuscations

For software obfuscations, it is commonly assumed that the attacker has access to the software binary. He can use static analyzes and possibly dynamic analyzes if he is able to properly execute the software to trigger communications.

In [25], Collberg *et al.* propose a taxonomy of obfuscating transformations for software programs. This taxonomy distinguishes four transformation targets: 1) layout obfuscation; 2) data obfuscation; 3) control obfuscation and 4) preventive transformation. In particular, the data obfuscation category, that is relevant to protocol obfuscation, contains three sub-categories: 1) Storage & Encoding; 2) Aggregation and 3) Ordering. To measure the effect of an obfuscating transformation, three metrics are defined: 1) *potency* describing how much a program is more complex to understand by a human being; 2) *resilience* describing how it resists to automatic tool analysis; and 3) *cost* assessing the execution time/space penalty which a transformation incurs on an obfuscated application.

Initially, software obfuscation has focused on hardening decompilation steps [26], [27], [28], [29]. In [30], Wroblewski proposed obfuscation transformations specific to binary code instead of transformations that apply to higher level languages. In [31], Linn and Debray introduce the replacement of direct calls by so-called branching functions. This work is extended in [32] by Cappaert and Preneel. They formalize the notion of control flow graph flattening to prevent information leakage.

Recently, some solutions have been proposed to mitigate dynamic analysis. As an example, software diversification is applied in [33] to increase the complexity of dynamic analysis.

Most of dynamic analyses are based on data tainting [34], thus in [35], transformations are proposed to increase the risk of obtaining a wrong taint analysis.

### B. Communication protocol obfuscations

For communication protocol obfuscations, the frequently considered adversary model is an attacker who can eavesdrop a communication channel to collect transmitted data, without having access to the binary of the application.

Many obfuscation techniques have been proposed to mitigate network censorships. In [1], four categories are distinguished to classify protocol obfuscations: *Randomization*, *Mimicry*, *Tunneling/Covert Channel* and *Programmable*. This classification differs from the one in [25] by Collberg *et al.* which considers transformations that must be integrated into the design process of the application while Dyer *et al.* classification considers transformations applied after the application development.

1) *Randomization*: The goal of *Randomization* is to transform a message sequence into a network traffic seemingly random. This transformation must prevent fingerprinting and any inference of any statistical characteristics of the protocol.

The main projects dealing with obfuscation by randomization are used in Tor as Pluggable Transports plugins<sup>1</sup>, *e.g.* ScrambleSuit [36], obfproxy [37]. These projects modify the application layer encoding and some part of the transport layer (connection characteristics) that are often used in firewall rules. These techniques are very effective against firewalls based on blacklists.

2) *Mimicry*: The goal of *Mimicry* is to change the communication characteristics (notably, message format) to mimic characteristics of other legitimate protocols, *e.g.* Skype or HTTP.

With this technique, firewalls based on whitelists of authorized protocols can be bypassed. As an example, StegoTorus [38] project embeds information into the headers and body of a set of predefined HTTP messages, using steganographic techniques. SkypeMorph [39] uses the facts that Skype traffic is encrypted and focuses on mimicry of statistical characteristics of a Skype communication. However, both of these approaches can be distinguished from legitimate protocols using semantics, dependencies between connections and error connections [40]. Furthermore, *mimicry* incurs a higher overhead (time and memory usage) compared to *randomization*.

3) *Tunneling/Covert Channel*: The goal of *Tunneling* is to use a legitimate layer protocol as a new transport layer protocol. The tunneling strategy can be integrated in an application using a library implementing the legitimate protocol. Thus, the observed behavior corresponds to the behavior of a legitimate application which uses the legitimate protocol. However, the overhead of this solution is higher compared to *Mimicry*. Skype has been widely used for this purpose in the Freewave [41] and Facet [42] projects. In [43], a solution

based on online videogames communications is proposed to reduce the overhead. Their solution is also easily adaptable to different online videogame protocols.

4) *Programmable*: The goal of this technique is to combine benefits of both *Randomization* and *Mimicry* by allowing the system to be configured to accommodate either strategy. FTE [44] project is categorized as a programmable system by the authors because the obfuscation techniques are parameterized by the user with a regular expression. However, it only considers message format. Thus, they developed Marionette [1] to take into account communication channel properties.

5) *Cryptography*: *Cryptography* is a specific type of *Randomization* that also ensures security properties: privacy and integrity. Encrypted traffic is difficult to process by the reverser, the resilience metric is therefore very high for this category. On the other hand, this category is cumbersome and costly. It often requires the use of keys that have to be distributed, managed and revoked, and the use of a cryptographic algorithm that is costly at runtime. Moreover, one can question the robustness of these techniques if the attacker model is extended by considering that attacker also has a copy of the application binary. Indeed, tools such as Dispatcher [16] and Reformat [21] have shown their efficiency in identifying the interface between the cryptographic functions and the core of the application. A debugger placed at this interface can then dump the plain messages, thus, bypassing the cryptographic algorithms.

### C. Discussion and contribution

As pointed out by Dyer *et al.* in [1], most existing protocol obfuscation techniques have been developed in order to circumvent network censorships. These techniques were not designed to provide efficient protection against protocol reverse engineering. Indeed, they can be easily bypassed especially when the attackers have access to a network trace and to the binary of the application.

Usually, the obfuscation transformations are integrated into the binary and are applied a posteriori. Accordingly, the reverse engineering process can be facilitated if probes can be successfully placed by the adversary at the interface between the core application and the transformation layer.

As far as we know, none of the state-of-the art techniques have investigated the possibility to obfuscate the specification of the communication protocol to provide protection against protocol reverse engineering. The main contribution of this paper consists in defining and implementing a framework for communication protocols obfuscation based on such approach, considering transformations that are applied to the specification of the format of the messages. The transformations are, by construction, invertible to avoid ambiguities while parsing a message. Also, the definition of the transformations is guided by the reverse engineering challenges discussed in Section II-C, to make the reverse engineering process more cumbersome and complex. In particular the following observations are taken into account in our approach:

<sup>1</sup><https://www.torproject.org/docs/pluggable-transport>

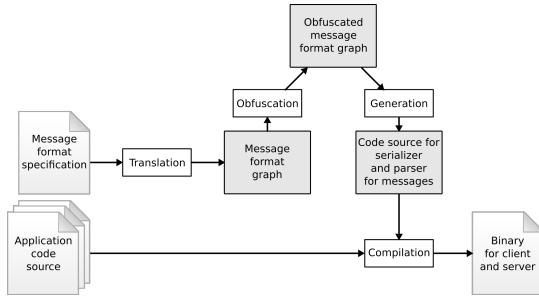


Fig. 2. Architecture of the framework ProtoObf.

- Inference algorithms used by PRE tools to retrieve the protocol grammar or the message format rely on a classification of messages. An obfuscation that could lead to a bad classification will likely affect the efficiency of the reverse engineering activity.
- The specification of communication protocols is generally based on regular models (automata, tree, etc.) that are simple to implement and for which messages can be parsed and generated quickly. Naturally, PRE tools usually adopt similar models to infer the protocols messages format or grammar. These tools are likely to be inefficient if more complex models are used to generate the obfuscated messages (e.g., pushdown automata, context-free grammar, etc.) without sacrificing processing time [45].
- The development, debug and maintenance of obfuscated protocols should not result in significant overheads to the users, thus, building a message should use the same interface, even in presence of obfuscations.

The obfuscation framework presented in the remaining sections is aimed at fulfilling these requirements. In this paper we only address the obfuscation of the protocol message format. The following advantages of our approach can be highlighted: 1) operating at the protocol specification level allows the definition of transformations that are aware of the semantics of the message fields (in other words, the transformations are coherent with respect to the organization of the message); 2) transformations are generated using non-regular languages (e.g., context-free language such as  $a^n b^n$  or context sensitive language such as the copy language) to make the syntax of protocol messages appear more complex than the syntax of regular languages; 3) obfuscated messages are more complex to infer with acceptable parsing and processing time; 4) our approach is integrated directly into the development process of the application. The core application doesn't build the non-obfuscated message to send. The obfuscated message is directly constructed when it is serialized. This strategy complicates the work of reverse engineering tools even if the attacker has access to the binary of the application; 5) our approach is orthogonal to existing solutions, thus can be used in conjunction with them.

## IV. ARCHITECTURE

The architecture of our framework, named *ProtoObf*, is presented in figure 2. The input of the framework is the message format specification of the protocol (noted  $S$  in the following). This specification is translated into a graphical representation named a message format graph and noted  $G_1$  in the following.

According to criteria established by the developer, the framework selects a number  $n$  of transformations to be applied to  $G_1$ . These transformations are either aggregation transformations or ordering transformations according to the taxonomy defined in [25]. Each of the transformations noted  $\tau_i$  takes a graph  $G_i$  as an input and provides a modified graph  $G_{i+1}$  as an output. The chosen transformations are composed and applied to the initial graph  $G_1$ . Note that all the transformations must be invertible so that the receiver is able to inverse the transformation.

The framework is used during the design and the development of the application to generate the source code that will perform the obfuscation or deobfuscation of messages during the execution of the application, based on  $G_{n+1}$ . Therefore, the output of the framework is the source code for the message parser and the corresponding message serializer. These source codes must be integrated within all the applications that communicate, so that they use the same obfuscations.

During the execution, the message serializer analyzes an abstract syntax tree (AST) of a message, which is an instantiation of  $G_1$  (i.e., it belongs to the language generated by  $G_1$ ). This AST is serialized by performing transformations on the fly while constructing the obfuscated message.

Let us note that the graph is an abstraction of the format of the messages and does not contain the values of the message fields. These values are defined in each AST corresponding to the instantiation of the graph for a specific message.

## V. MODELS AND TRANSFORMATIONS

This section first presents the different models we adopted for the formalization of the message format of the protocol as well as the obfuscations of the messages. Then, detailed information is provided for the proposed elementary obfuscations chosen in our approach. Finally, the main principles of the serializer and parser behavior are presented.

### A. Message format graph

This section provides more details on the abstract syntax tree of messages and on the associated message format graph. These models are illustrated, in figure 3, with a simplified example.

An AST is structured as a tree containing nodes and edges. A leaf of this tree represents a value of a message field. The overall message corresponds to the concatenation of these values using an ordered depth-first search. The intermediate nodes of the AST describe the message structure. Figure 3 presents an example of two types of messages derived from the Modbus protocol, denoted as  $M_1$  and  $M_2$ , with the associated

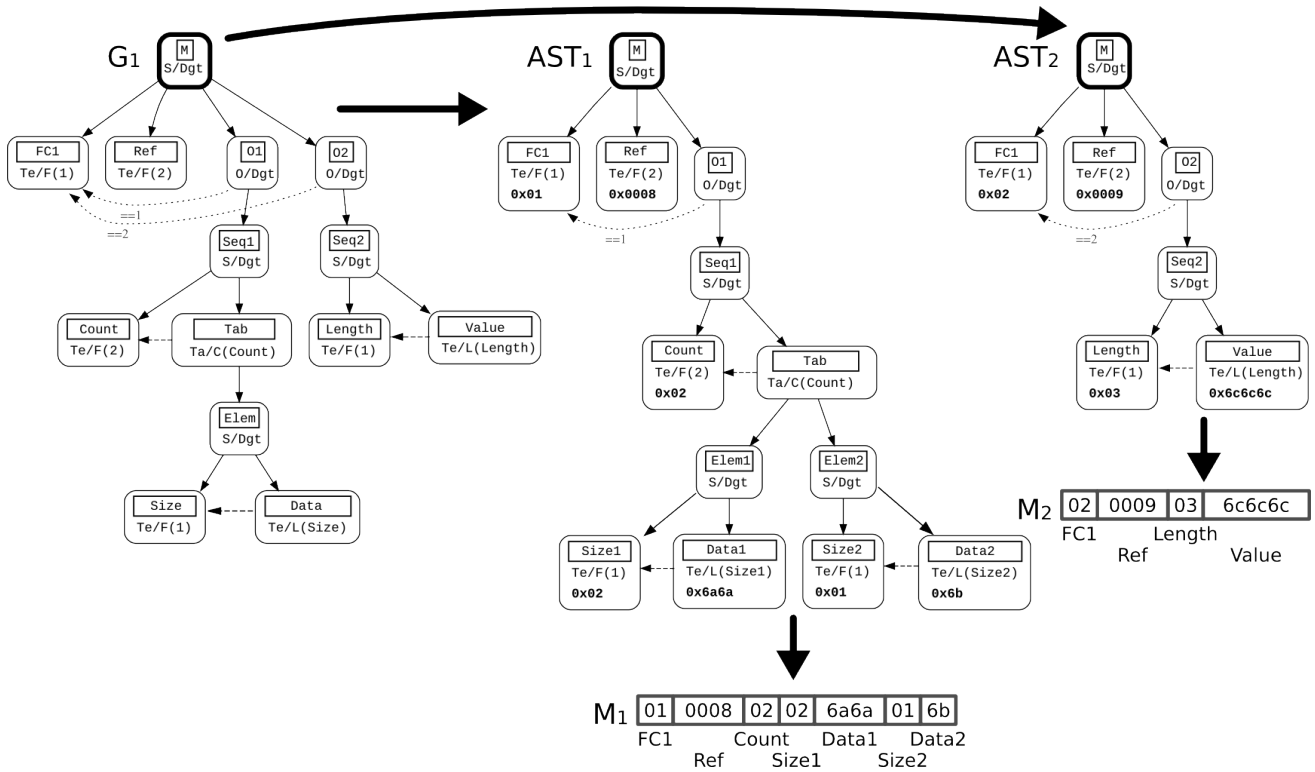


Fig. 3. Message format graph and abstract syntax trees.

abstract syntax trees,  $AST_1$  and  $AST_2$ , and the corresponding sequence of bytes.

A message format graph  $G_1$  describes all AST that are compliant to the specification of  $S$ . A node of the graph describes a node in the corresponding AST. In figure 3, the graph  $G_1$  describes both  $AST_1$  and  $AST_2$ .

A node is defined by five attributes: 1) a *Name*; 2) a *Type*; 3) a list of sub-nodes named *SubNodes*; 4) a parent node named *Parent* (none for the root node) and 5) a boundary method named *Boundary*. The *Type* or the *Boundary* attributes may contain an implicit reference to another node.

The type of a node can be:

- *Terminal* if the node of the AST contains user data or message related information, e.g. the size of another node;
- *Sequence* if the node of the AST contains a sequence of sub-nodes;
- *Optional* if the node of the AST is optional, depending on the value of another node in the AST;
- *Repetition* if the node of the AST consists of a repetition of the same sub-node;
- *Tabular* if the node of the AST consists of a repetition of its sub-node, and the number of repetitions is given by another node in the AST.

The *Boundary* attribute indicates the method used to define the length of the associated field. It can be:

- *Fixed* if it has a fixed size defined in  $S$ ;
- *Delimited* if it ends with a predefined byte or sequence of bytes (for instance  $\backslash r \backslash n$  in *HTTP*);

- *Length* if the length of the field is defined by another node;
- *Counter* if the node is a *Tabular*, the number of repetitions of the sub-node in the AST is defined by another node;
- *End* if the field corresponds to the remaining of the message;
- *Delegated* if the length of the field corresponds to the sum of the length of the sub-nodes.

The *Boundary* attribute must be consistent with the type of the field. For instance, a *Terminal* field must be delimited either with a *Fixed* boundary, a *Delimited* boundary, a *Length* boundary or an *End* boundary.

This graph is well suited to describe classical protocols that rely on regular models in language theory: *Optional* type can be used to represent the “|” operator; *Sequence* type is used for the concatenation “.”; and *Tabular* and *Repetition* types can be used to represent closure “\*”.

In the representation of such graph in the figure 3, nodes are represented by their name. The type of node for *Terminal*, *Sequence*, *Tabular* and *Optional* fields is specified under the node by using notation  $Te$ ,  $S$ ,  $Ta$  and  $O$ . *Boundaries* are shown for *Delimited*, *Delegated* and *End* by the notation  $De$ ,  $Dgt$  and  $E$ , for *Fixed* by the notation  $F(n)$  ( $n$  stands for the fixed size), and for *Counter* and *Length* by the notations  $C(n)$  and  $L(n)$  (where  $n$  stands for the node that helps to define the size, identified with a dashed arrow in the figure).

TABLE I  
SUMMARY OF GENERIC TRANSFORMATIONS

<i>SplitAdd</i> A <i>Terminal</i> node with a value $v$ is split into a sequence of two sub-nodes with values $v_1, v_2$ : $v = v_1 + v_2$ .
<i>SplitSub</i> and <i>SplitXor</i> Same as <i>SplitAdd</i> with a subtraction or a xor.
<i>SplitCat</i> A <i>Terminal</i> node with a value $v$ is split into a sequence of two sub-nodes with values $v_1, v_2$ : $v = \text{concatenate}(v_1, v_2)$ .
<i>ConstAdd</i> A <i>Terminal</i> node with a value $v$ is substituted by a node with value $v + \text{constant}$ ( <i>constant</i> is predefined in the framework).
<i>ConstSub</i> and <i>ConstXor</i> Same as <i>ConstAdd</i> with a subtraction or a xor.
<i>BoundaryChange</i> A <i>Delimited Boundary</i> is changed into a <i>Length Boundary</i> : the node is replaced by a sequence of two-nodes $n_1, n_2$ ( $n_1$ is the length of $n_2$ ).
<i>PadInsert</i> A node with random value is added to a <i>Sequence</i> .
<i>ReadFromEnd</i> A node is read from the end, from right to left.
<i>TabSplit</i> A <i>Tabular</i> with $n$ sub-nodes is replaced by a sequence of <i>Tabular</i> nodes.
<i>RepSplit</i> Same as <i>TabSplit</i> with a <i>Repetition</i> .
<i>ChildMove</i> Permutation of two sub-nodes of a <i>Sequence</i> .

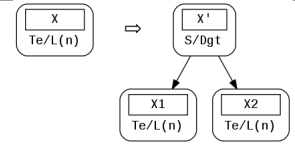
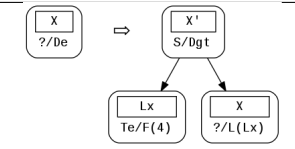

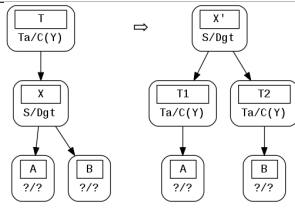
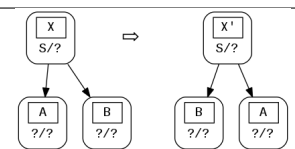
## B. Transformations

A transformation, noted  $\tau_i$ , modifies the structure of a message format graph that leads to a modification of the abstract syntax tree of the messages processed during the execution. Thus, it also leads to a change of the message serializer and the message parser behavior. The transformation must be invertible by design to allow the receiver to correctly parse the obfuscated message. The proposed framework is designed to be applied to a large set of message format graphs. Thus, we have defined a set of generic transformations that are presented in table I. They include ordering transformations such as *ChildMove* and *TabSplit* and some aggregation transformations such as *SplitCat* and *ConstAdd*. This set can be extended with new generic transformations.

A generic transformation  $\mathcal{T}$  is a function that consists in changing a graph pattern  $a$  into a graph pattern  $b$ , associated to some applicability constraints. If the graph  $G_i$  being obfuscated by the framework contains the graph pattern  $a$  and if  $G_i$  complies with the constraints of  $\mathcal{T}$ , then the transformation  $\tau_i$  can be derived and the graph  $G_{i+1}$  is obtained by replacing the instantiated graph pattern  $a$  by the instantiated graph pattern  $b$  (with a renaming of nodes if needed).

The proposed transformations do not remove any information; they can only modify the value or the order of the different fields of the message. These transformations are easily inverted. In other words, we have  $\tau_i^{-1} \circ \tau_i = \text{id}$ . The main difficulty lies in the composition of the message parser and the message serializer with the transformations. Therefore, these transformations are constrained to ensure that the composition of the message parser and the message serializer leads to the

TABLE II  
DESCRIPTION FORMAT OF GENERIC TRANSFORMATIONS

<b>SplitAdd</b>	
	<b>Serialization pseudocode</b> ↓ Choose a random value $X1$ Compute $X2 = X + X1$
	<b>Constraints</b> <i>Boundary</i> of parent nodes must be either <i>Delegated</i> or <i>End</i>
<b>Challenge</b> Inference models and classification: more dependencies between fields in message and various representations of the same message	
<b>BoundaryChange</b>	
	<b>Serialization pseudocode</b> ↑ Measure the serialization of $X$ Prefix the result with this length
	<b>Constraints</b> <i>Boundary</i> of parent nodes must be either <i>Delegated</i> or <i>End</i>
<b>Challenge</b> Fields delimitation: delimitation with a length field	
<b>ReadFromEnd</b>	
	<b>Serialization pseudocode</b> ↑ Mirror the serialization of $X$
	<b>Constraints</b> <i>Boundary</i> of parent nodes can be anything but <i>Delimited</i>
<b>Challenge</b> Inference models and classification: subpart of message read in reverse order	
<b>TabSplit</b>	
	<b>Serialization pseudocode</b> ↓ Map <i>fst</i> and <i>snd</i> on $X$ Create the sequence
	<b>Constraints</b> <i>Boundary</i> of parent nodes can be anything but <i>Delimited</i> and <i>Boundary</i> of $X$ must be <i>Delegated</i>
<b>Challenge</b> Inference models: turn a regular language $(AB)^*$ into a context-free language $A^m B^m$	
<b>ChildMove</b>	
	<b>Serialization pseudocode</b> ↓ Switch children in $X$
	<b>Constraints</b> <i>Boundary</i> of parent nodes can be anything but <i>Delimited</i> and no nodes inside $B$ must depend on a node inside $A$
<b>Challenge</b> Classification: meaningful fields are no more at the beginning	

identity.

The framework memorizes, for each applied transformation  $\tau_i$ , the node in the graph that corresponds to the graph pattern  $a$ . Accordingly, it is able to correctly derive the message serializer and the message parser, taking into account the



transformations. Some transformations may change the values of the fields that are needed to correctly serialize (or parse) the remaining of the AST (or of the message), for instance a length field. As a result, the strategy adopted is to process transformations on the fly. The message serializer uses a depth first search on the AST and the transformations are executed during this graph traversal (same for the message parser).

Each generic transformation can be formatted as presented in Table II. This table illustrates the more interesting generic transformations from Table I. Other transformations are small variation (for instance *SplitSub*, etc.). For the generic transformations of figure II, the graph at the left hand side of the  $\Rightarrow$  symbol corresponds to pattern *a* and the one at the right hand side corresponds to the result of the transformation (pattern *b*). The serialization pseudocode is generated by the framework to perform the transformation on the fly. The vertical arrow indicates if this transformation is performed before serializing the children (down arrow), or on the result of the serialization (up arrow). The constraints correspond to the attributes to check on the node of the pattern, the sub-nodes and parent nodes. The last information indicated in the table is the protocol reverse challenge that is emphasized by each transformation. These challenges are presented in section II-C.

Table II shows that most of the challenges are covered by one of these generic transformations. The *SplitTab* and the *ReadFromEnd* transformations change a regular language, that is compatible with most of reverse engineering tools, into a language that does not fit models traditionally supported by these tools (for instance, context-free language as  $a^n b^n$ ). In particular, the *ReadFromEnd* encodes a message from right to left. This practice is unusual and makes the inference of links between fields very difficult. The delimitation of fields that is easier in presence of *Delimited* node, is more difficult with *Length* node and the *BoundaryChange* change from the first towards the second. In addition, this generic transformation is also useful to circumvent some constraints of other generic transformations. The classification is also made more difficult with generic transformations like *SplitAdd* that can be applied on message keywords which are often used to decide classification. The only challenge that is not addressed directly by generic transformation is the *Observation*. This challenge is addressed in the implementation of the framework, presented in section VI.

### C. Serializer and parser behavior

As soon as the message serializer starts the serialization of a node, it inspects the list of transformations to find out if one transformation needs to be applied before serializing the node. If so, this transformation is executed on the current node of the AST. Then the message serializer processes the node and the node with its sub-tree is replaced by a node containing the result of the serialization. At the end of this processing, the serializer inspects again the list of transformations to know if a transformation must also be applied at the end of the serialization of the node. If so, this transformation is in turn applied. The parser works in the same way. However,

the parser has to face an additional challenge: to rebuild a sub-node of AST from the message, it must first delimit the corresponding sub-part in the message.

## VI. IMPLEMENTATION

The framework is implemented using the C language. *Lex* and *Yacc* tools are used to parse the message format specification and generate the message format graph. The structure that represents a message format graph is simply a transcription in the C language of the attributes presented in section V-A. Then, each node of the graph is analyzed to identify compatible generic transformations. A transformation is randomly chosen among them and applied to the node. This routine is applied as many times as indicated by a parameter specified in the framework. Finally, a depth-first search algorithm is executed on the resulting message format graph to generate the source code. Generic transformations presented in the previous section cover all reverse engineering challenges except the *Observation* challenge. This last challenge is taken into account during the generation of the code source used to manipulate, parse and serialize an AST. In the following, we provide more information on the structure used to store the AST and the functions generated by the framework that the core application can use to instantiate this AST (i.e., the accessors of the AST).

First, let us consider a naive implementation that consists in instantiating an non-obfuscated AST, during the execution of the core application, and then, in applying the selected transformations to the complete non-obfuscated AST to generate the obfuscated AST which is then serialized. With such approach, the entire non-obfuscated AST and obfuscated AST are available in the memory during the execution and a unique function is used to obfuscate the first AST. Therefore it is easy to locate this function in the memory to recover the non-obfuscated AST. Obfuscation techniques that process the binary usually obfuscate the code and the internal data. However, the AST is designed to generate a message that will be sent through the network and these obfuscation techniques ignore these data (in fact, they must not modify the format of message sent in the network).

Our framework focuses on the message format specification. It can obfuscate an intermediate representation of the AST that does not correspond neither to the entire non-obfuscated AST nor to the entire obfuscated AST. In the framework, this intermediate representation corresponds to the AST after the application of aggregation transformations and before the application of ordering transformations. When the core application decides to send a message, it generates this temporary AST through a set of setter functions. These setter functions perform aggregation transformations on the fly. When the AST is complete, ordering transformations are applied while serializing this message. Hence, the serialization is spread into multiple function calls.

The code source generated by the framework provides the prototypes of the message parser and serializer, plus the

accessors (setters and getters) and the structures for the intermediate AST. Getters and setters are functions that retrieve or store a value in a field while performing the aggregation transformations, on-the-fly. To make them harder to identify, they can be implemented as macro, and thus inlined in the code. This interface must be stable regardless of the chosen transformations. Accordingly, the set of transformations can be easily replaced by another set of transformations without changing the core application. From a practical point of view, this interface is directly obtained from the non-obfuscated specification of the message format. Accessors will hide the complexity implied by aggregating transformations while the code of the parser and serializer hide the complexity implied by the ordering transformations.

## VII. EXPERIMENTATIONS

To evaluate our framework, we have implemented the specification of two protocols: a binary protocol, TCP-Modbus [46], and a text-protocol HTTP [47]. Modbus contains a *Tabular* field, a *Length Boundary* and a *Counter Boundary*, while HTTP contains an *Optional* field, a *Repetitive* field, as well as *Delimited Boundary*. For Modbus protocol, we have also developed a core application that generates the messages 1, 2, 3, 4, 5, 6, 15 and 16 and their response, as required by *simply modbus*<sup>2</sup> client implementation. This set of messages includes all the different formats of Modbus messages. For HTTP protocol, we have also developed a simplified core application. However, this implementation doesn't create messages with consistent values for the keywords. We consider this verification to be relevant to the server code, not to the parser code.

### A. Experiments

In order to analyze the impact of the framework, several experiments are carried out with a different number of obfuscations (0 to 4) per field, *i.e.*, per node of the graph. For each experiment, the transformations are selected randomly among the set of applicable generic transformations and the code source of the parser and serializer is generated. The core application is compiled with this code source. Then, it is executed to generate different messages with random values. The source code used to initialize the message and invoke the serialization process and the parsing process is the same for all experiments. This validates our new concept of protocol transformations created at the compilation time: the code that uses the protocol is simple and independent of applied transformations.

The results presented in the following subsections for HTTP and Modbus correspond to 5000 experiments each (1000 for each obfuscation scenario, from 0 to 4).

### B. Measures

During the experiments, different measures were collected. The time required for the code source generation (*i.e.*, the parsing of the specification, the application of transformations

and the code generation) is called *Generation time*. This number must be low to allow the developer to easily adjust the number of desired transformations.

The maximum number of obfuscations per node and the total number of applied transformations on the graph are memorized. Multiple experiments with the same number of transformations per node may lead to different numbers of effectively applied transformations on the graph. Indeed, according to the transformations applied (which are randomly chosen), the number of obfuscations may be different because some transformations may create new nodes whereas others do not create any. Also, some randomly selected transformations may not be applicable if the associated constraints are not satisfied. So, we compute the *average*, *min* and *max* for this last metric.

The complexity of the generated code, *i.e.* the potency of the obfuscations, is also considered in the experiments. The number of code lines is the amount of code generated by the framework for the complete protocol specification. It contains code for serializer, parser and accessors functions, the internal structures and sanity checks, *i.e.* the complete serialization library. Let us recall that the main objective of our approach is to make the reverse engineering of the obfuscated protocol significantly more difficult for the attacker than without obfuscation. The increase of the complexity can be reflected *e.g.*, by a higher number of the lines of code or of the number of internal structures used in the library to store data during the parsing process.

The *cflow* tool is used to extract the call graph for the parsing process. This graph reflects the complexity of function invocations in the code. We retrieve the size of this graph (the number of nodes) and its depth.

Finally, we have evaluated the cost of our solution by measuring the time required to serialize and parse a message, and the space overhead associated to the serialized message size (through the evaluation of the buffer size).

### C. Results

Results are summarized for each considered protocol in Tables III, IV, respectively. Three values are indicated, with the following syntax: *average*, [*min*, *max*]. The results reported for the potency metrics are normalized by the values associated to the non-obfuscated version. The cost metrics are provided in absolute values.

For the simple case where at most one obfuscation is applied per node (which nevertheless corresponds to an average of 10.1 applied transformations on the HTTP graph and 47.8 applied transformations on the Modbus graph), the complexity of the generated code is about twice the complexity of the code without obfuscation. In particular, the increase in the number of structures reflects a significant difference between the initial specification and the result of the transformation. For the other obfuscation cases, these metrics increase as expected. The highest increase is observed for the call graph size.

To have better insights on the impact of obfuscations, Figures 6 and 7 plot the evolution of the potency metrics relative

<sup>2</sup><http://www.simplymodbus.ca>

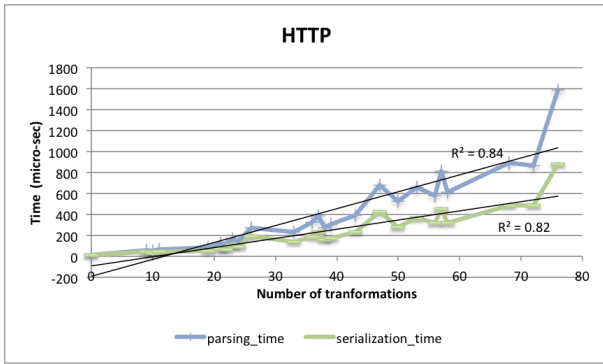


Fig. 4. HTTP: Parsing and serialization time

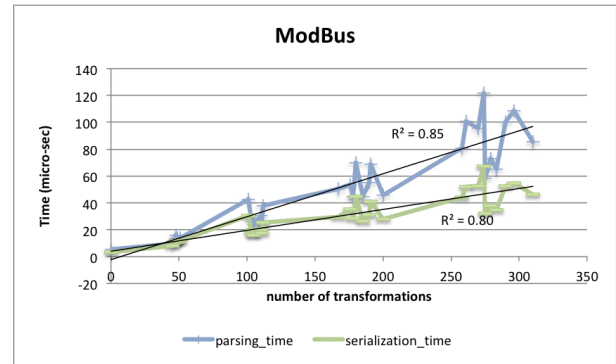


Fig. 5. ModBus: Parsing and serialization time

increase compared to the non-obfuscated case, according to the number of obfuscations applied on the graph. Generally, we observe a linear increasing trend of the number of lines, the number of structures and the size of the call graph. The increase of the call graph depth and of the buffer size is slower and tend to stabilize. The increase of buffer size is kept very low which is very important especially for application contexts where network packet resources are usually more crucial than application execution time.

The cost of the obfuscations is illustrated in figures 4 and 5 which present the evolution of the parsing and serialization times according to the number of transformations applied on the graph. The straight lines report the result of the linear regression between these times and this number of transformations (the correlation coefficient is also indicated). These figures show that inevitably the processing time in the presence of transformations increases. However, this increase is linear with the number of transformations applied and the slope is smooth. This indicates that the overhead due to these transformations is not important and could be reduced with a more optimized implementation of the framework. Note that these results are achieved with a high number of transformations. A developer may consider it sufficient to make only a limited number of transformations. It is also noteworthy that in all the experiments that we have carried out, the parsing and serialization times did not exceed 0.5 ms for Modbus and 2.8 ms for HTTP. The average values are significantly lower.

Finally, as regards the cost of our obfuscation framework associated to the generation of the obfuscated code, it remains low. Indeed, the generation time is kept under 4 ms in the worst case. This worst case corresponds to a succession of *SplitOp* obfuscations applied on a large data field. It is noteworthy that the overhead associated to the generation of the obfuscated code is less critical as this operation is performed offline.

#### D. Resilience Assessment

To analyse the resilience of our framework, we asked an expert of (and a contributor to) *Netzob* [12], a popular protocol reverse engineering tool based on network trace analysis, to perform PRE. We have sent to him a network trace containing

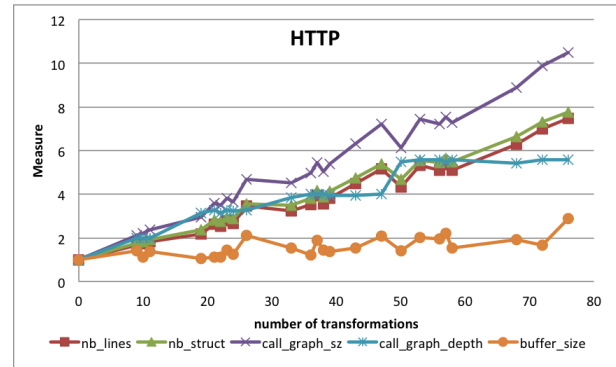


Fig. 6. HTTP: normalized potency metrics

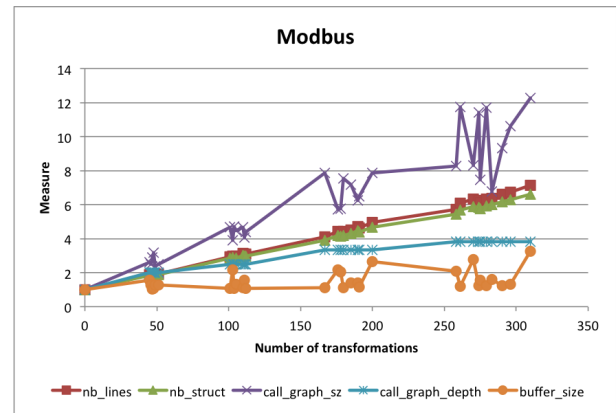


Fig. 7. ModBus: normalized potency metrics

4 different messages and their corresponding answers of Modbus protocol. In less than half an hour, he was able to retrieve the exact format of the messages for the non-obfuscated protocol. For a version generated with one obfuscation per field, he was not able to obtain any relevant results after more than two hours of work. He confirmed that the obfuscated code was more difficult to analyze with classic PRE tools. Of course, this assessment is not sufficient, and more significant experiments are needed to validate the resilience of the framework. It is noteworthy that such experiments are not easy to perform as they require the contribution of independent protocol reverse

TABLE III  
A COMPARATIVE RESULTS FOR HTTP PROTOCOL

<i>Nb. transf. per node</i>	1	2	3	4
<i>Nb. transf. applied</i>	10[9; 11]	22[19; 26]	39[33; 47]	59[50; 76]
<b>Potency (normalized)</b>				
<i>Nb. lines</i>	1.7[1.6; 2.0]	2.7[2.2; 3.5]	4.0[3.2; 5.2]	5.6[4.3; 7.5]
<i>Nb. structs</i>	1.8[1.7; 2.1]	2.9[2.4; 3.6]	4.3[3.5; 5.4]	5.9[4.7; 7.8]
<i>Call graph size</i>	2.2[2.0; 2.6]	3.7[3.0; 4.7]	5.6[4.5; 7.2]	7.9[6.1; 10.5]
<i>Call graph depth</i>	2.0[2.0; 2.0]	3.2[3.1; 3.3]	4.0[3.9; 4.0]	5.5[5.4; 5.6]
<b>Costs (absolute)</b>				
<i>Generation time (ms)</i>	2.10[1.92; 2.41]	3.17[2.59; 4.03]	4.80[3.84; 6.36]	8.93[5.41; 26.08]
<i>Parsing time (ms)</i>	0.06[0.04; 0.12]	0.15[0.08; 0.47]	0.37[0.22; 1.00]	0.79[0.47; 2.80]
<i>Serialization time (ms)</i>	0.04[0.02; 0.10]	0.10[0.05; 0.34]	0.22[0.13; 0.75]	0.43[0.25; 1.57]
<i>Buffer size (bytes)</i>	137[95; 244]	154[101; 284]	181[112; 297]	219[119; 404]

TABLE IV  
A COMPARATIVE RESULTS FOR TCP-MODBUS PROTOCOL

<i>Nb. transf. per node</i>	1	2	3	4
<i>Nb. transf. applied</i>	47[45; 51]	107[101; 112]	184[167; 200]	279[258; 310]
<b>Potency (normalized)</b>				
<i>Nb. lines</i>	1.9[1.8; 2.0]	3.0[2.8; 3.2]	4.5[4.1; 4.9]	6.4[5.7; 7.1]
<i>Nb. structs</i>	1.9[1.8; 1.9]	2.9[2.7; 3.1]	4.3[3.9; 4.7]	6.0[5.4; 6.6]
<i>Call graph size</i>	2.6[2.1; 3.2]	4.3[3.4; 5.5]	6.8[4.7; 8.6]	9.8[6.8; 12.2]
<i>Call graph depth</i>	2.0[2.0; 2.0]	2.5[2.5; 2.5]	3.3[3.3; 3.3]	3.8[3.8; 3.8]
<b>Costs (absolute)</b>				
<i>Generation time (ms)</i>	6.39[5.97; 6.72]	12.53[9.66; 31.06]	16.34[14.56; 17.74]	24.29[21.76; 27.01]
<i>Parsing time (ms)</i>	0.01[0.00; 0.06]	0.03[0.01; 0.14]	0.05[0.01; 0.25]	0.09[0.02; 0.52]
<i>Serialization time (ms)</i>	0.01[0.00; 0.06]	0.02[0.00; 0.10]	0.03[0.01; 0.16]	0.05[0.01; 0.31]
<i>Buffer size (bytes)</i>	30[3; 195]	33[3; 293]	38[3; 381]	42[3; 478]

engineering experts and to have an easy access to automatic PRE tools which is not the case today.

### VIII. CONCLUSION

This paper presented a novel protocol obfuscation framework that is aimed at increasing the effort needed by an adversary to successfully reverse the protocol. The main contribution consists in obfuscating the specification of the messages format. The specification is formalized as a graph on which generic transformations are automatically applied to generate a library code that can be easily linked to the core application. The obfuscated messages are scattered throughout the memory so that it is difficult for the reverser to easily reconstruct the message. A proof of concept prototype of the framework is implemented and a set of experiments are carried out on two protocols to illustrate the feasibility of the proposed approach and evaluate its impact on the complexity of the generated code and its overhead. The results show a significant increase of the complexity of the obfuscated protocol binary compared to the non-obfuscated code. It is also shown that the execution time and memory overhead remains acceptable for a practical deployment of the approach in operation.

Our approach can be applied to any protocol for which the specification of the messages can be represented according to the proposed message format graph. We believe that this can be

easily achieved for most common protocols, including binary and text protocols. The proposed framework also provides the opportunity to enhance the protection of the considered protocol as new obfuscated versions of the protocol can be easily generated. The deployment of new versions, at regular intervals, should decrease the likelihood that the protocol can be successfully reversed and compromised.

It is noteworthy that the proposed framework is designed to resist to attacks aimed at reverse engineering the protocol, rather than extracting partial information concerning e.g., specific data fields or keywords. Cryptographic techniques are more suitable in this latter case.

Several extensions of this work can be investigated. In particular, in the current implementation the obfuscations are selected randomly. A more efficient approach could be defined by taking into account the grammar of the protocol. Another open question concerns the definition of the number of obfuscations needed to achieve an acceptable level of resilience of the protocol against reverse engineering attacks. Finally, a more significant validation of the proposed approach needs to be carried out, using e.g. different automated reverse engineering tools and independent experts. Such evaluation is not easy to achieve.

## REFERENCES

- [1] K. P. Dyer, S. E. Coull, and T. Shrimpton, "Marionette: A Programmable Network Traffic Obfuscation System," in *Proc. 24th USENIX Security Symp.* USENIX Assoc., 2015, pp. 367–382.
- [2] E. J. Chikofsky and J. H. C. II, "Reverse engineering and design recovery: A taxonomy," *IEEE Software*, vol. 7(1), pp. 13–18, 1990.
- [3] C. Eagle, *The IDA Pro Book: The Unofficial Guide to the World's Most Popular Disassembler*. San Francisco, CA, USA: No Starch Press, 2008.
- [4] "Radare2 github repository," <https://github.com/radare/radare2>, 2017.
- [5] J. Narayan, S. K. Shukla, and T. C. Clancy, "A survey of automatic protocol reverse engineering tools," *ACM Computing Surveys (CSUR)*, vol. 48, no. 3, p. 40, 2015.
- [6] X. Li and L. Chen, "A Survey on Methods of Automatic Protocol Reverse Engineering," in *2011 7th Int'l Conf. Computational Intell. and Security (CIS)*. Hainan, China: IEEE, 2011, pp. 685–689.
- [7] J. Duchêne, C. Le Guernic, E. Alata, V. Nicomette, and M. Kaâniche, "State of the art of network protocol reverse engineering tools," *J. Comput. Virology and Hacking Techniques*, pp. 1–16, Jan. 2017.
- [8] M. Beddoe, "Protocol Informatics Project," <http://www.4tphi.net/~awalters/PI/PI.html>, 2004.
- [9] —, "Network Protocol Analysis using Bioinformatics Algorithms," <http://www.4tphi.net/~awalters/PI/pi.pdf>, 2004.
- [10] J. Antunes, N. Neves, and P. Verissimo, "Reverse Engineering of Protocols from Network Traces," in *2011 18th Working Conf. Reverse Eng. (WCRE)*. New York, NY, USA: IEEE, 2011, pp. 169–178.
- [11] G. Bossert, F. Guihery, and G. Hiet, "Towards automated protocol reverse engineering using semantic information," in *Proc. 9th ACM Conf. Comput. & Commun. Security*. Kyoto, Japan: ACM, Jun. 2014, pp. 51–62.
- [12] G. Bossert, "Exploiting Semantic for the Automatic Reverse Engineering of Communication Protocols." Ph.D. dissertation, Suplec, Dec. 2014.
- [13] J. Lim, T. Repts, and B. Liblit, "Extracting Output Formats from Executables," in *13th Working Conf. Reverse Eng., 2006. WCRE '06*. Benevento, Italy: IEEE, 2006, pp. 167–178.
- [14] J. Caballero, P. Poosankam, C. Kreibich, and D. Song, "Dispatcher: enabling active botnet infiltration using automatic protocol reverse-engineering," in *Proc. 16th ACM Conf. Comput. & Commun. Security*, ser. CCS '09. New York, NY, USA: ACM, 2009, pp. 621–634.
- [15] J. Caballero Bayerri, "Grammar and model extraction for security applications using dynamic program binary analysis," Ph.D. dissertation, Carnegie Mellon Univ., Pittsburgh, PA, USA, 2010.
- [16] J. Caballero and D. Song, "Automatic protocol reverse-engineering: Message format extraction and field semantics inference," *Comput. Networks*, vol. 57, no. 2, pp. 451–474, Feb. 2013.
- [17] P. Comparetti, G. Wondracek, C. Kruegel, and E. Kirda, "Prospex: Protocol Specification Extraction," in *2009 30th IEEE Symp. Security and Privacy*. Berkeley, USA: IEEE, 2009, pp. 110–125.
- [18] C. Y. Cho, D. Babić, E. C. R. Shin, and D. Song, "Inference and Analysis of Formal Models of Botnet Command and Control Protocols," in *Proc. 17th ACM Conf. Comput. & Commun. Security*, ser. CCS '10. New York, NY, USA: ACM, 2010, pp. 426–439.
- [19] C. Y. Cho, D. Babić, P. Poosankam, K. Z. Chen, E. X. Wu, and D. Song, "MACE: model-inference-assisted concolic exploration for protocol and vulnerability discovery," in *Proc. 20th USENIX Conf. Security*, ser. SEC '11. Berkeley, CA, USA: USENIX Assoc., Aug. 2011, p. 19.
- [20] J. Caballero, H. Yin, Z. Liang, and D. Song, "Polyglot: automatic extraction of protocol message format using dynamic binary analysis," in *Proc. 14th ACM Conf. Comput. & Commun. Security*, ser. CCS '07. New York, NY, USA: ACM, 2007, pp. 317–329.
- [21] Z. Wang, X. Jiang, W. Cui, X. Wang, and M. Grace, "ReFormat: Automatic Reverse Engineering of Encrypted Messages," in *Comput. Security ESORICS 2009*, ser. LNCS, M. Backes and P. Ning, Eds. Saint Malo, France: Springer Berlin Heidelberg, Jan. 2009, no. 5789, pp. 200–215.
- [22] B. Barak, O. Goldreich, R. Impagliazzo, S. Rudich, A. Sahai, S. Vadhan, and K. Yang, "On the (im) possibility of obfuscating programs," in *Ann. Int'l Cryptology Conf.* Springer, 2001, pp. 1–18.
- [23] —, "On the (im)possibility of obfuscating programs," *J. ACM*, vol. 59, no. 2, pp. 6:1–6:48, May 2012.
- [24] H. Xu and M. R. Lyu, "Assessing the security properties of software obfuscation," *IEEE Security & Privacy*, vol. 14, no. 5, pp. 80–83, 2016.
- [25] C. Collberg, C. Thomborson, and D. Low, "A taxonomy of obfuscating transformations," Dept. Comput. Sci., The Univ. of Auckland, New Zealand, Tech. Rep., 1997.
- [26] W. Cho, I. Lee, and S. Park, "Against intelligent tampering: Software tamper resistance by extended control flow obfuscation," in *Proc. World Multiconference on Systems, Cybern., and Informatics*, 2001.
- [27] C. Collberg, C. Thomborson, and D. Low, "Manufacturing cheap, resilient, and stealthy opaque constructs," in *Proc. 25th ACM SIGPLAN-SIGACT Symp. Principles of programming languages*. ACM, 1998, pp. 184–196.
- [28] T. Ogiso, Y. Sakabe, M. Soshi, and A. Miyaji, "Software obfuscation on a theoretical basis and its implementation," *IEICE Trans. Fundamentals of Electron., Commun. and Comput. Sci.*, vol. 86, no. 1, pp. 176–186, 2003.
- [29] C. Wang, J. Hill, J. Knight, and J. Davidson, "Software tamper resistance: Obstructing static analysis of programs," CS-2000-12, Univ. of Virginia, 12 2000, Tech. Rep., 2000.
- [30] G. Wroblewski, "General method of program code obfuscation," Ph.D. dissertation, Inst. of Eng. Cybern., Wrocław Univ. of Technology, 2002.
- [31] C. Linn and S. Debray, "Obfuscation of executable code to improve resistance to static disassembly," in *Proc. 10th ACM Conf. Comput. & Commun. Security*. ACM, 2003, pp. 290–299.
- [32] J. Cappaert and B. Preneel, "A general model for hiding control flow," in *Proc. 10th annual ACM workshop on Digital rights management*. ACM, 2010, pp. 35–42.
- [33] S. Schrittwieser and S. Katzenbeisser, "Code obfuscation against static and dynamic reverse engineering," in *Inform. Hiding*. Springer, 2011, pp. 270–284.
- [34] M. Egele, T. Scholte, E. Kirda, and C. Kruegel, "A survey on automated dynamic malware-analysis techniques and tools," *ACM computing surveys (CSUR)*, vol. 44, no. 2, p. 6, 2012.
- [35] S. Blazy, S. Riaud, and T. Sirvent, "Data tainting and obfuscation: Improving plausibility of incorrect taint," in *Source Code Anal. and Manipulation (SCAM), 2015 IEEE 15th Int'l Working Conf.* IEEE, 2015, pp. 111–120.
- [36] P. Winter, T. Pulls, and J. Fuss, "ScrambleSuit: A Polymorphic Network Protocol to Circumvent Censorship," in *Proc. 12th ACM Workshop on Privacy in the Electron. Society*, ser. WPES '13. New York, NY, USA: ACM, 2013, pp. 213–224.
- [37] Tor team, "Obfsproxy," <https://trac.torproject.org/projects/tor/wiki/doc/PluggableTransports/obfs4proxy>, 2017.
- [38] Z. Weinberg, J. Wang, V. Yegneswaran, L. Briesemeister, S. Cheung, F. Wang, and D. Boneh, "Stegotorus: a camouflage proxy for the tor anonymity system," in *Proc. 2012 ACM Conf. Comput. & Commun. Security*. ACM, 2012, pp. 109–120.
- [39] H. Mohajeri Moghaddam, B. Li, M. Derakhshani, and I. Goldberg, "SkypeMorph: Protocol Obfuscation for Tor Bridges," in *Proc. 2012 ACM Conf. Comput. & Commun. Security*, ser. CCS '12. New York, NY, USA: ACM, 2012, pp. 97–108.
- [40] J. Geddes, M. Schuchard, and N. Hopper, "Cover Your ACKs: Pitfalls of Covert Channel Censorship Circumvention," in *Proc. 2013 ACM SIGSAC Conf. Comput. & Commun. Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 361–372.
- [41] A. Houmansadr, T. J. Riedl, N. Borisov, and A. C. Singer, "I want my voice to be heard: Ip over voice-over-ip for unobservable censorship circumvention," in *NDSS*, 2013.
- [42] S. Li, M. Schliep, and N. Hopper, "Facet: Streaming over videoconferencing for censorship circumvention," in *Proc. 13th Workshop on Privacy in the Electron. Society*. ACM, 2014, pp. 163–172.
- [43] H. Bridger, N. Rishab, G. Phillipa, and J. Rob, "Games Without Frontiers: Investigating Video Games as a Covert Channel," in *Proc. 2016 IEEE European Symp. Security and Privacy*, ser. IEEE European Symp. Security and Privacy. IEEE, 2015.
- [44] K. P. Dyer, S. E. Coull, T. Ristenpart, and T. Shrimpton, "Protocol Misidentification Made Easy with Format-transforming Encryption," in *Proc. 2013 ACM SIGSAC Conf. Comput. & Commun. Security*, ser. CCS '13. New York, NY, USA: ACM, 2013, pp. 61–72.
- [45] C. de la Higuera, *Grammatical Inference: Learning Automata and Grammars*. New York, NY, USA: Cambridge Univ. Press, 2010.
- [46] A. Swales, "Open modbus/tcp specification," Schneider Electric, Tech. Rep., 1999.
- [47] R. Fielding and J. Reschke, "Hypertext transfer protocol (http/1.1): Message syntax and routing," Internet Eng. Task Force, Tech. Rep., 2014.