A conceptual model and simulation model for phishing

Gerd Wagner

Brandenburg University of Technology, Konrad-Wachsmann-Allee 5, 03046 Cottbus, Germany

Abstract

We present a conceptual model for phishing, which is the basis of a simulation model, using the Object Event Modeling and Simulation approach. Both the conceptual model and the simulation model help to clarify the real-world semantics of phishing and provide a basis for more elaborate models, e.g., for capturing suitable concepts of trust and susceptibility or for capturing the organizational aspects needed for modeling spear phishing.

Keywords

conceptual phishing model, phishing simulation model, ontological grounding

1. Introduction

Phishing is an attempt to steal exploitable data via message-based communication with targets, typically in the form of user names, passwords, or other account information. Phishers, who disguise themselves as a trusted source, e.g., by impersonating a reputable brand such as Microsoft, either use the stolen information themselves, for instance to take over the victim's accounts, or sell the stolen information.

For tricking the targets to trust the phishing message, phishers often use a spoofed email address. The phisher's *lure message* contains a weblink (often with a spoofed URL) that leads to a forged website consisting of one or more *hook pages*, which trick the target to enter exploitable data.

The main concept of phishing, described in Figure 1, consists of the following four steps:

1. A phisher sends a lure message to phishing targets.
2. Targets follow the deceptive link in the lure message leading to the phishing website.
3. Targets provide exploitable data on a hook page of the phishing website.
4. The phisher exploits the scammed data after obtaining it.

This process description subsumes various messaging channels, such as email, texting, WhatsApp, etc., but the main phishing channel is email with deceptive links (which is the #1

Proceedings of the Joint Ontology Workshops (JOWO) - Episode X: The Tukker Zomer of Ontology, and satellite events co-located with the 14th International Conference on Formal Ontology in Information Systems (FOIS 2024), July 15-19, 2024, Enschede, The Netherlands.

g.wagner@b-tu.de

0000-0003-3049-8885

CEUR Workshop Proceedings ceur-ws.org ISSN 1613-0073

email threat category according to CloudFlare's 2023 *Phishing Threats Report*, appearing in 35.6% of their detections). While the main concept of phishing refers to this 4-step information-stealing scenario, there is also a broader definition of phishing, which includes the case where the target is tricked into downloading malicious software. In this paper, however, we restrict our attention to the main concept of phishing.

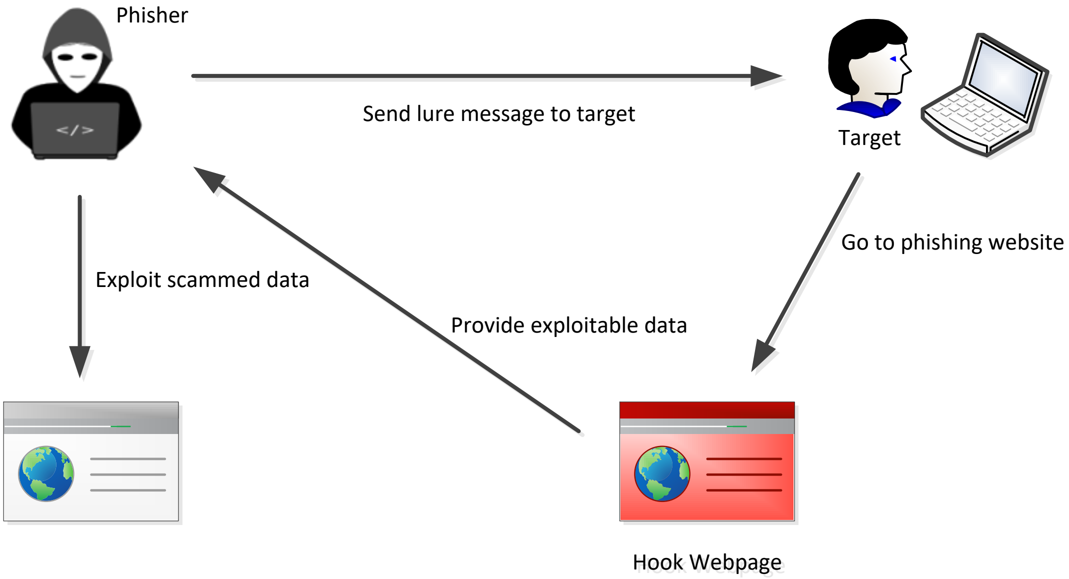

Figure 1: Phishing as a four-step process.

In *bulk phishing*, which allows attacks using mass email lists with private email addresses, a popular brand such as Microsoft, Adobe, DHL or Amazon, is used as the impersonated sender, while in *spear phishing*, the attack is more personalized and therefore requires special information about potential victims, including their professional email addresses.

2. Conceptual domain model

In *Object Event Modeling (OEM)* [1], a conceptual domain model consists of a conceptual information model, e.g., in the form of an *Object Event (OE) Class Model*, and a conceptual process model, e.g., in the form of a *BPMN Process Model*.

Since phishing is an interaction between phishers and their targets, we model phishers and their targets as *agents* employing the agent concepts of OEM:

1. Agents are special objects that interact with each other via ***communication*** and with their inanimate environment via ***perception*** and ***action***.
2. Agents may *perceive* objects and events in their environment and, in response, take certain actions. ***Perception events*** may lead to an update of the information state of the perceiver. ***Action events*** may change the environment.
3. A *communication event* is composed of two successive atomic events: an ***out-message action event*** (corresponding to a message send action of the sender) and a correlated

in-message event (or message reception by the receiver). In-message events may lead to an update of the information state of the receiver.

Analyzing the phishing process sketched in Figure 1, we can identify the agent types “phisher” and “phishing target”, the object types “lure message” and “hook page”, as well as the (out-message) action and in-message/perception event types listed in Table 1.

Table 1
Action and Perception Event Types

(Out-Message) Action Event	In-Message/Perception Event
send lure message	receive lure message
open lure message body	read lure message header
visit hook page	read lure message body
provide exploitable data	look at hook page
exploit scammed data	perceive exploitable data

All these (agent, object and event) types, and the associations between them, are described in the form of a conceptual OE class diagram shown in Figure 2.

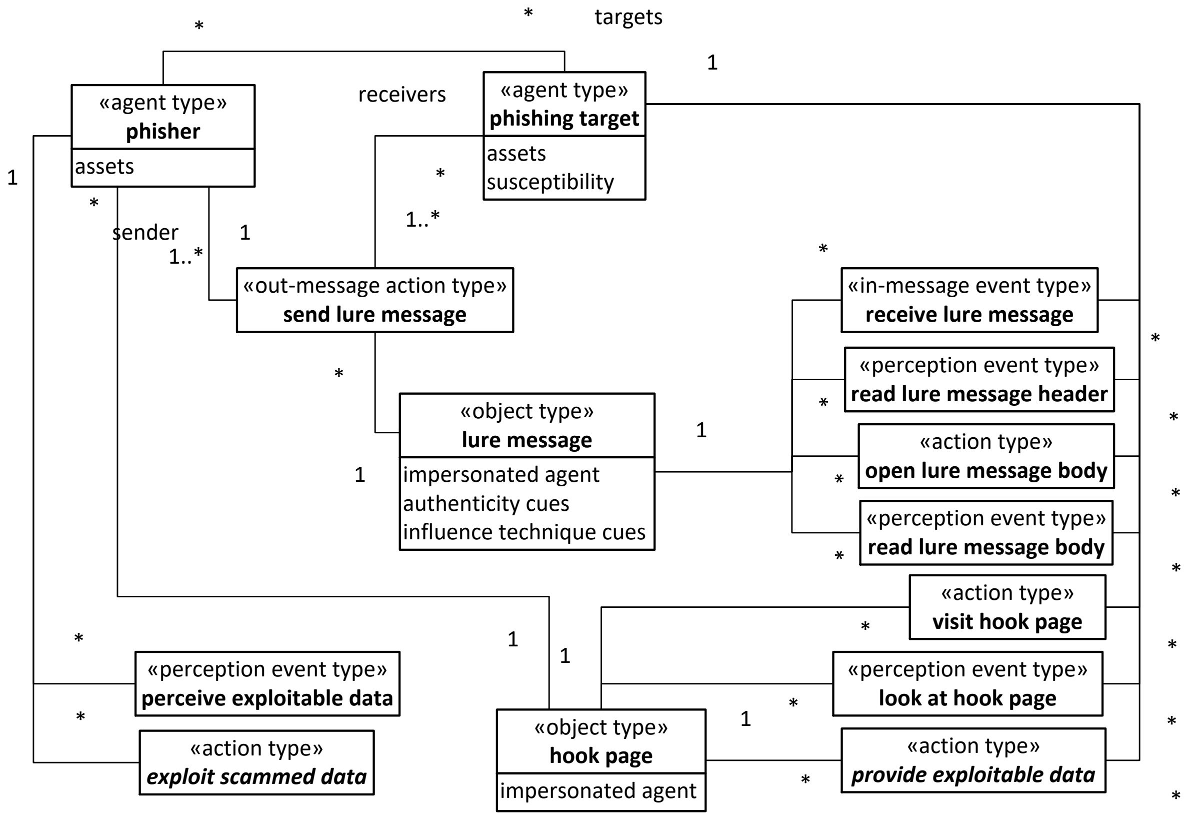

Figure 2: A conceptual OE class diagram describing agent, object and event types.

For the *participation associations* between object types and event types, OE class models show a *history multiplicity* at the side of the event types. For instance, for the association between *phishers* and *send lure message* events, the multiplicity 1..* shown at the side of *send lure message* events implies that *over time*, a phisher participates in *one or many* events of type *send lure message*. By default, if it is not shown as a second multiplicity annotation of such an association end, the *snapshot multiplicity* is 0..1, which means that *at a time*, a phisher may participate in at most one *send lure message* event.

By default, the event types in an OE class model are types of instantaneous events, which implicitly have an attribute *occurrence time*.

While the OE class model shown in Figure 2 describes the types of agents, objects and events involved in a phishing scenario, the BPMN process diagram shown in Figure 3 describes the possible sequences of events and their effects in such a scenario. It consists of two BPMN Pools, one for agents of type *phisher* and one for agents of type *phishing target*. BPMN Pools are container rectangles representing an agent type and including events and actions concerning that agent type.

In the diagram of Figure 3, we use the BPMN shape for "tasks" (rectangles with rounded corners) for representing actions, and the shape for "signal events" (an Event circle with an enclosed triangle) for representing perception events.

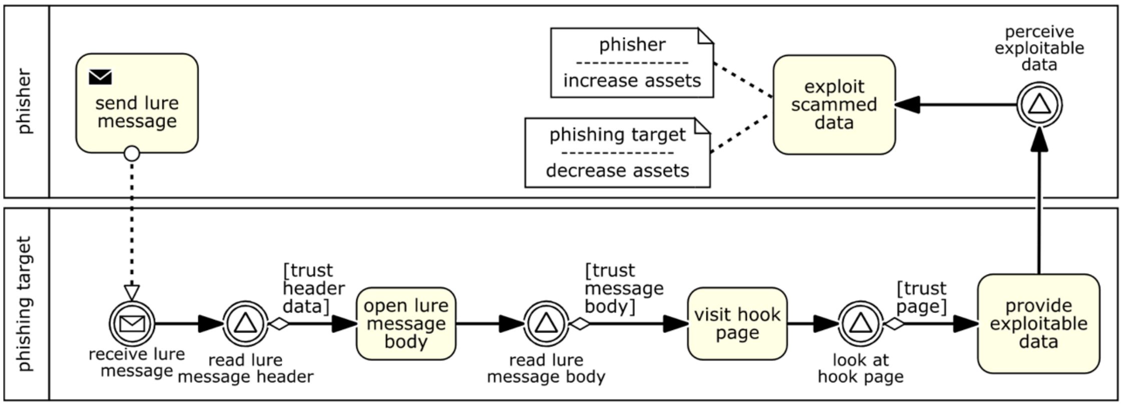

Figure 3: A conceptual process model describing the possible sequences of events in a phishing scenario.

The model of Figure 3 includes three trust-based decisions of the target: only if targets first trust the lure message header data (sender address and subject line), and then trust the message body and finally trust the hook page, they fall victim to the phishing attempt by providing exploitable data.

Notice that in this model the only event that affects the objects involved (by changing their state) is *exploit scammed data*, which leads to a transfer of assets from the target to the phisher.

3. Phishing concepts

The model of Figure 2 includes the following phishing-specific concepts and attributes:

1. A *phisher* may be a private individual or an agent working for a criminal or state organization.
2. A *phishing target* may be a private individual or a person working for an organization. The *susceptibility* of a phishing target denotes their behavioral disposition to be victimized by a phishing attempt.
3. A *lure message* has
    a. an *impersonated agent* (such as Microsoft) as an alleged sender,
    b. *authenticity cues* (such as the sender's email address, spelling and grammar, subject line and the domain name of the hook page) for signaling trustworthiness,
    c. *influence technique cues* (such as statements of urgency or authority) for pushing the target to activate one of its deceptive links,
    d. one or more *deceptive links*.
4. A *hook page* also has an *impersonated agent* and *authenticity cues* for signaling trustworthiness corresponding to those of the lure message.

According to [2], the target's level of attention to authenticity cues is negatively related, while the level of attention to the influence technique cues is positively related, to the likelihood to respond to phishing emails

4. Ontological grounding of phishing concepts

In [3], a *Phishing Attack Ontology (PHATO)* has been proposed. PHATO is grounded in the *Reference Ontology for Security Engineering (ROSE)* [4] and the *Common Ontology of Value and Risk (COVER)* [5], which are both founded on the *Unified Foundational Ontology (UFO)* [6].

PHATO covers most of the concepts listed in the previous section, except for

1. the susceptibility of targets,
2. authenticity cues of lure messages and hook pages,
3. influence technique cues of lure messages.

5. Simulation design model and implementation

An information design model in the form of an OE class design model on the basis of the conceptual information model of Figure 1 is shown in Figure 4.

A simulation design model is obtained from the conceptual model by making certain simplifications (such as dropping irrelevant elements) and by enriching it with computational details such that the result is a computationally complete model (an executable specification). A simulation design model consists of an information design model and a process design model.

An information design model in the form of an OE class model defines a number of classes (for object and event types) that can be implemented with an Object-Oriented simulation language such as OESjs [7], which is a JavaScript-based framework for Object Event Simulation.

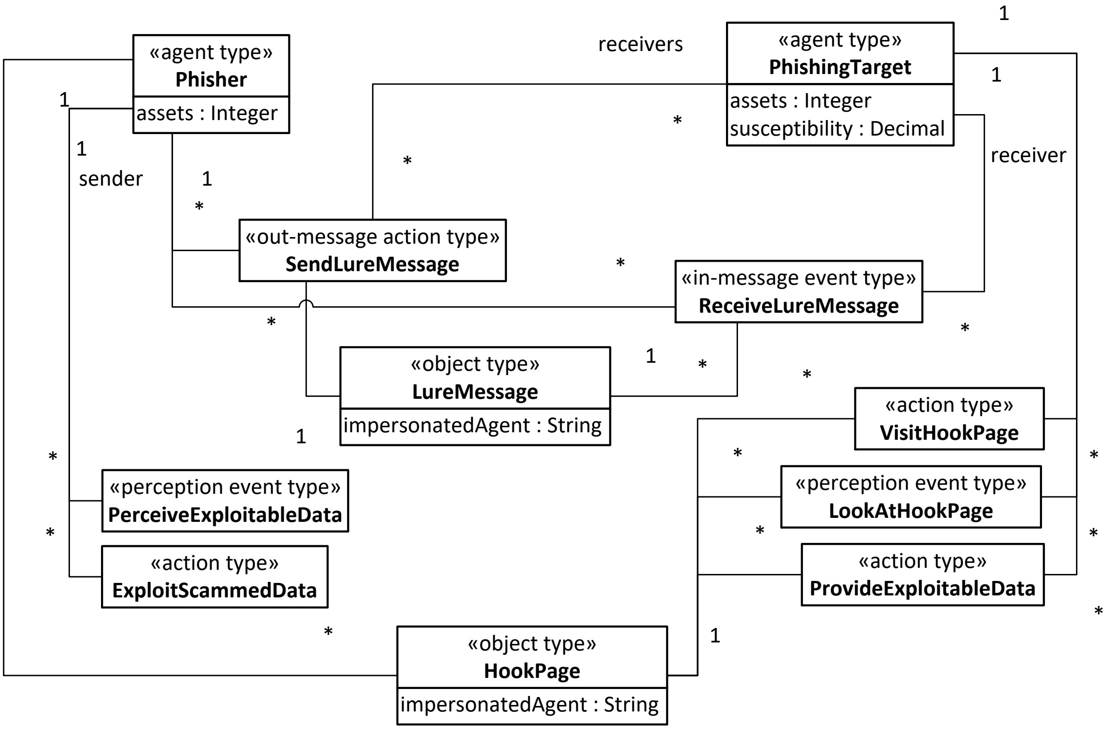

Figure 4: An information design model in the form of an OE class diagram defining eleven classes: two agent types, two object types, and seven event types.

In the OE class design model of Figure 4, the *susceptibility* attribute of phishing targets has been modeled as a decimal-valued attribute holding a probability value, while for simplicity the attributes *authenticity cues* and *influence technique cues* of lure messages have been dropped.

An OESjs implementation of this simulation design model is available at https://gwagner57.github.io/oes/phishing-1.

6. Conclusions

We have presented a conceptual model for (agent-based) phishing simulation, which is the basis of a simulation design that has been implemented with the simulation framework OESjs. We have also briefly discussed how to obtain an ontological grounding of the phishing concepts included in our conceptual model with UFO, pointing out some open issues.

Both the conceptual model and the simulation design model provide a basis for more sophisticated models, which may aim to capture elaborate concepts of the susceptibility of targets related to their trust in messages and websites or the organizational aspects needed for modeling spear phishing.

References

[1] G. Wagner, Object Event Modeling and Simulation, 2023. URL: https://sim4edu.com/reading/OEMS.

[2] A. Vishwanath et al, ”Why do people get phished? Testing individual differences in phishing vulnerability within an integrated, information processing model”, Decision Support Systems 51 (2011) 576–586.

[3] Í. Oliveira, R.F. Calhau, G. Guizzardi, Toward a phishing attack ontology, ER2023: Companion Proceedings of the 42nd International Conference on Conceptual Modeling: ER Forum, 7th SCME, November, 2023, Lisbon, Portugal.

[4] Í. Oliveira, T. P. Sales, R. Baratella, M. Fumagalli, G. Guizzardi, An ontology of security from a risk treatment perspective, in: International conference on conceptual modeling, Springer, 2022, pp. 365–379.

[5] T. P. Sales, F. Baião, G. Guizzardi, J. P. A. Almeida, N. Guarino, J. Mylopoulos, The common ontology of value and risk, in: Conceptual Modeling. ER 2018, volume 11157, Springer, 2018, pp. 121–135.

[6] G. Guizzardi, A. Botti Benevides, C. M. Fonseca, D. Porello, J. P. A. Almeida, T. P. Sales, UFO: Unified Foundational Ontology, Applied Ontology 17 (2022) 1–44.

[7] G. Wagner, Object Event Simulation with OESjs, URL: https://sim4edu.com/reading/des-engineering/es-with-oesjs