DARIO FORTE, DAVIDE BERTOLETTI, CRISTIANO MARUTI, AND MICHELE ZAMBELLI

Sebek Web "ITA" interface

AN ALTERNATIVE APPROACH

Dario Forte, CISM, CFE, is adjunct professor of incident response and forensics at the University of Milano, Crema Campus. He is also the founder of the Incident Response Italy Project and the coordinator of the Italian Honeynet Project and is the president of the European Chapter of the HTCIA.

dario.forte@acm.org

Davide Bertoletti, Cristiano Maruti, and Michele Zambelli are graduate students at the University of Milano at Crema and are part of the Incident Response Italy team and the Italian Honeynet Project. Their research interests focus on security and incident response.

IN 2003, A GROUP OF TEACHERS AND undergraduates at the University of Milano at Crema started a project called Incident Response Italy (IRItaly), whose aim was to provide guidelines for incident response and forensics. After five months, the same group founded the Italian Honeynet Project (IHP), which became part of the Honeynet Research Alliance. Since then the IHP has worked on many tasks, like the beta and national deployment of the Honeynet Security Console (from Jeff Dell). The IHP is now working on the development side as well, by contributing to the Sebek Project.

Sebek, a tool created by Ed Balas of Indiana University, is basically a piece of code that lives entirely in kernel space and records either some or all of the data accessed by users on the system. It has the ability to record keystrokes from a session that is using encryption, recover files copied with SCP, capture passwords used to log in to a remote system, recover passwords used to enable Burneye-protected binaries, and accomplish many other forensics-related tasks. You will find more information on the tool in the papers mentioned in the bibliography. In this article we'll discuss our approach to the Web interface.

Our Work

Although the current version (0.8) of the Sebek Web Interface is stable and complete, we have added some additional features that make it even more useful and applicable in a broader variety of situations. The new features are the following:

- XML output
- WAP interface
- Dump database
- Paging

XML OUTPUT

We have written a Web application that displays information collected by Sebek. The application is written entirely in PHP and returns the information as an XML document. We rewrote the tool to output data in XML because it is more flexible than the HTML format previously used. We have also added capabilities to view the status of monitored hosts via cellular phone or WAP-enabled device. We have chosen an XML, Web-based application for a number of reasons:

- XML is becoming the database language for the Web.
- XML is the interchange mechanism between applications.
- XML cleanly separates presentation layers from data layers.

Figure 1. Sebek in Browse Mode

Basically, the architecture works as follows:

The application creates a connection with the remote database, gets and modifies the data, and sends the data back as an XML file. Data is displayed as HTML pages using an XSL stylesheet. The application is also usable in a wireless environment via WAP technology.

The XML file has a two-part structure. The first is located in the header tag and contains information concerning the header of the HTML page; the second, placed between Sebek tags, contains data obtained from the database. Both of these parts are placed between the root tags.

```
<?xml version="1.0" ?>
<root>
<header>
.
.
.
.
</header>
<sebek>
.
.
.
.
</sebek>
</root>
```

Let's have a look at the most common tags in the document:

- `root`: the main tag, within which the other components of the document are placed
- `header`: contains information about the page header
- `sebek`: contains the data created by a database query
- `read data`: contains some information about the database fields

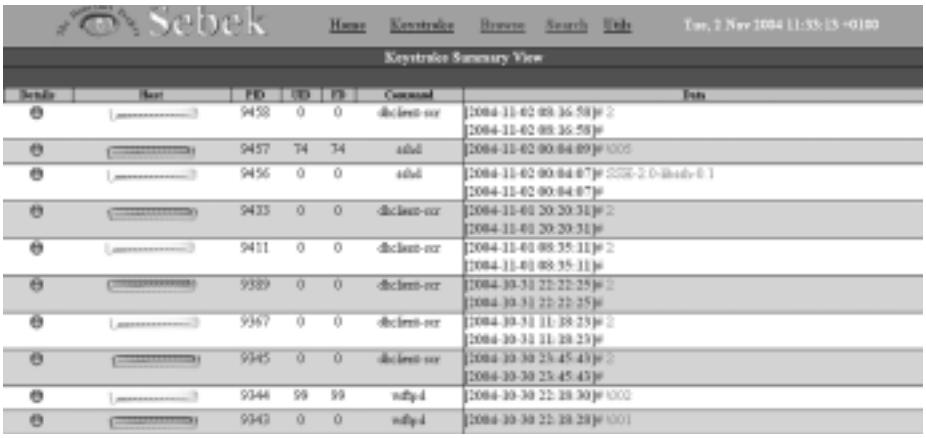

Figure 2. Sebek in Keystroke Mode

There is a relationship between data returned by the Web interface and the data stored in the database. Other data are obtained via aggregation functions or similar operations:

- `ip addr:` IP address of the port where the commands were issued
- `insert time`: insert time of the information in the database
- `command`: command executed by the intruder on the compromised host
- `time`: current host time
- `fd`: file descriptor
- `pid`: process ID
- `uid`: user ID
- `length`: byte length of the recorded activity
- `data list`: a summary list of the keystrokes executed by the intruder
- `rec num`: total number of inserted records
- `start time`: start command time
- `end time`: end command time

In addition to the tags we have explained above, there is another type of tag with a different kind of function:

- `link`: creates a link to a different page. The link can be text type or image type. In addition to the type of link, defined by the `aspect` attribute, there is an additional attribute. The `page` attribute is used to define the link, the frame in which to open the new page, and possibly other parameters.
- `now`: contains information about local server date and time.
- `space`: defines an empty area.

WAP INTERFACE

It is possible to get information about our honeynet via a WAP device, without using a networked PC. By means of a device which supports WAP technology, we are able to verify at any time and from any location whether the network is under attack.

The most useful part of the WAP interface is that it provides a summary page containing the monitored hosts, the number of records for each host, and the latest update. Only these few items were chosen for display because of the limited computational and graphics capabilities of WAP devices. The details of teh situation are reported through the HTML application. In any case, the summary page provides a useful snapshot of the monitored hosts.

It is also possible to create a dump of the database from the WAP device, which may be handy if we are not connected through a "normal" network device.

DUMP DATABASE

After some days of network activity, we observed an overload of the Sebek MySQL database. We therefore created the Utilities section of the interface to allow a dump of the current database. The operation is catalogued with dump date and time information, and an empty database is created. This brilliant solution allows for faster browsing of the collected data entries and a reduction in transmission delay. It is also possible to activate a database dump via a WAP interface.

Figure 3. WAP Interface to Sebek

Further Developments

The Sebek Web ITA Interface can be downloaded from http://www.honeynet.it. Comments and feedback are welcome. Of course, it is just an experiment, but we are pretty confident that this tool can be a valid alternative to the current version of the interface. Meanwhile, Sebek is going to undergo some major changes. Ed Balas has presented the next version of Sebek, which will include a new interface, and the Italian Honeynet Project group was added to the developer team. Sebek and its implementations are proof that the Honeynet Project is maturing rapidly and effectively.

Acknowledgments

The tool was developed by the Italian Honeynet project, a group of people working with the University of Milano at Crema. Special thanks to Michele Zambelli, Cristiano Maruti, and Davide Bertoletti, who worked on the source code and the architecture.

REFERENCES

Ed Balas, "Honeynet Data Analysis: A Technique for Correlating Sebek and Network Data," *Proceedings of the Digital Forensic Research Workshop*: http://www.dfrws.org.

Ed Balas, "Know Your Enemy: Sebek, a Kernel-Based Data Capture Tool": http://www.honeynet.org/papers/sebek.pdf.

Dario Forte, "The Art of Log Correlation: Tools and Techniques for Digital Investigations," *Proceedings of the Information Security South Africa 2004 Conference*: http://www.dflabs.com/images/Art_of_correlation_Dario_Forte.pdf.

The Italian Honeynet Project: http://www.honeynet.it.