Chapter 1 (actually chapter 7 in the book)

The Building Security in Maturity Model as a Research Tool

CONTENTS

1.1 Introduction

Measurement of software security is difficult; it is next to impossible to take two pieces of code and decide and decide which is "more secure" than the other [1]. To tackle this problem, bright minds had the idea to instead try to measure second-order effects, i.e., to study the software security related activities performed by successful software development organisations.

The Building Security In Maturity Model (BSIMM)[2] has been used suc-

cessfully for years by the software security company Cigital[1] to measure the software security maturity level of their clients. The BSIMM report and framework is released with a Creative Commons Attribution-ShareAlike license[2], which implies that it is freely available to anyone who wants to use it for whatever purpose, including self-assessment.

In this paper we try to establish whether BSIMM is also suitable as an academic research tool, and discuss possible adjustments that could make it more tractable. The remainder of the paper is structured as follows: In Section 1.2 we present relevant background related to BSIMM. In Section 1.3 we present a case study where BSIMM was used by a third party to perform a maturity assessment of a set of software development organisations. We discuss further in Section 1.4, and conclude in Section 1.5.

1.2 Background

The starting point for the first BSIMM survey in 2008 [2] was to study the software security activities performed by nine selected companies. The nine companies were presumably far ahead in software security, and the activities that were observed here formed the basis of the framework in Table 1.1. Representatives from Cigital physically visited each company, and these first surveys were done by Gary McGraw and Sammy Migues personally, using a whole day for each company.

The purpose of BSIMM is to quantify the software security activities performed in real software development projects in real organisations. As these projects and organisations use different methodologies and different terminology, a framework that allows describing all initiatives in a unified manner has been created. The BSIMM framework consists of twelve practices organised into four domains; Governance, Intelligence, SSDL Touchpoints and Deployment (see Table 1.1). Each practice has a number of activities on three levels, with level 1 being the lowest maturity and level 3 is the highest. For example, for practice Strategy and Metrics, SM1.4 is an activity on level 1, SM 2.5 is an activity on level 2, and SM 3.2 is an activity on level 3. In total, there are currently[3] 112 BSIMM activities.

[1] `http://www.cigital.com`

[2] `https://creativecommons.org/licenses/by-sa/3.0/`

[3] New activities are added as they are observed in the field, and activities are promoted or demoted as their relative importance is determined to change. In the latest update of the BSIMM report, from BSIMM V to BSIMM 6, no new activities were added, but 4 activities were assigned new levels.

Table 1.1 The BSIMM Software Security Framework

Governance	Intelligence	SSDL Touchpoints	Deployment
Strategy and Metrics	Attack Models	Architecture Analysis	Penetration Testing
Compliance and Policy	Security Features and Design	Code Review	Software Environment
Training	Standards and Requirements	Security Testing	Configuration Management and Vulnerability Management

1.3 A Case Study

Jaatun et al.[3] performed a study on software security maturity of 20 public[4] organizations in a small European country using the BSIMM activities as a basis for a questionnaire. The method used in Jaatun et al.'s study can be characterized as "assisted self-evaluation"; the respondents from the various organisations indicated in a questionnaire which software security activities that they do, and then they participated in a follow-up interview with the purpose of clarifying uncertainties and correcting possible errors in the questionnaire. However, the researchers did synchronize their assessment criteria, both before and during the interview phase, in order to ensure that they had an as similar as possible perception of what is required to receive a "yes" for the various activities in the questionnaire. However, it is still possible that researchers may have made different assessments related to what should be approved as an activity.

Since the study was based largely on self-evaluation, there is reason to believe that the resulting "BSIMM-score" is higher than it would be with a review in line with the one made by Cigital in the original BSIMM study, since they were not in a position to verify the claims made by each organisation. In concrete terms, this implies that we must assume that it has been easier for the organisations to get an activity "approved" in that study than it would be if Cigital had done the survey in accordance with its usual practice. This means that although these results provide some indications of the maturity level of the evaluated organisations, none of the organisations in this study can claim that they have established their "BSIMM Score". It would also be misleading to compare their results directly with the official BSIMM reports. On the other hand, the validity of the answers in the study were increased because of the follow-up interviews, compared with the results from a pure survey.

One thing that is clear is that the organisations studied vary dramatically,

[4] government departments, government-owned or municipality-owned organisations, etc.

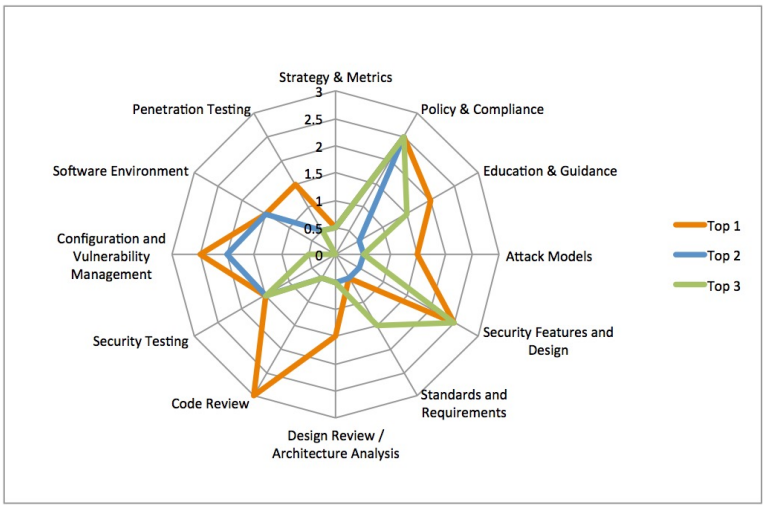

Figure 1.1: Conservative maturity for the three most mature organisations

bot in maturity level and in what kind of activities they perform. Figure 1.1 illustrates this for the three organisations that received the highest total maturity score among the 20 surveyed. This figure uses the so-called "conservative" BSIMM measure defined by Jaatun et al. [3], where 0.5 points are given if only some activities on level 1 are performed within a practice, 1 point means all activities on level 1 are performed, 1.5 points means all activities on level 1 pluse some on level 2 are performed, and so on. We see that the top organisation gets a top score in the practice "Code Review", but the next two organisations do only a few activities on the lowest maturity level. None of the three organisations do all of the activities even on the first level in the practice "Strategy and Metrics", whereas the third organisation does all the level 1 activities and some level 2 activities in the practice "Standards and Requirements", where the first and second organisation do not even do all the level one activities.

The BSIMM framework is based on the idea that there is a formally defined software security group (SSG), and the activities are centered around this group. Few of the surveyed organisations had such a formally defined group. Several organisations had a manager with more or less explicit responsibility for software security, but then usually as part of an overall security responsibility in the organisation.

1.4 Discussion

In personal communication, Konstantin Beznosov stated that he abandoned software security [4] as a research topic because he was unable to get access to the inner workings of the development organisations, and thus were unable to do real empirical research. This may be where the main differentiator lies, since Cigital typically fills the role of a consultant with the target companies, and at the point of performing the BSIMM assessment, they have already convinced the target organisation ”what's in it for them”. As BSIMM gains name recognition among mainstream European businesses, it may be that this may also spill over to more academic endeavors; many businesses are interested in knowing more about where they stand when it comes to software security, and many are interested to know how they compare with other, similar organisations.

“The real BSIMM” is not performed using a questionnaire, but using a questionnaire approach significantly lowers the threshold for initiating a software security maturity study. As Jaatun et al. [3] has shown, much of the ambiguity can be resolved by a simple follow-up interview. However, more work is necessary to compare the level of information than can be extracted from an organisation using questionnaire and follow-up, vs. embedding one or more researchers in the organisation for a day. Although self-assessment is frequently used in other fields such as medicine [5], we cannot overlook that optimistic bias will lead some respondents to overstate their practices [6]. However, it may be equally possible that some respondents may downplay their maturity because they realise that they could be even better; in a larger statistical sample these effects may cancel themselves out.

Another important aspect of BSIMM is that it is the actual performance of activities that is is important, not just having the procedures in place. Thus, depending on who is bing asked, the answer may be “yes” (because we have the procedures) or “no” (because we never use the procedures). Clearly, selection of respondents must be done carefully, and strategies to mitigate a situation of sub-optimal respondents must be in place. One could be tempted to say that more explanations of each activity would remove ambiguity and doubt, but with 112 activities the questionnaire is already quite long, and takes about 1 hour to fill out. In the case of an online questionnaire, an alternative might be to first only display the level 1 activities, and only then display the level 2 activities if all the level 1 activities are “fulfilled”. The disadvantage of this approach is that it only covers the conservative metric introduced by Jaatun et al., and not the weighted or high-water-mark metrics, the latter of which is used for comparison purposes in the BSIMM report [2].

BSIMM claims to be descriptive rather than normative, but by ranking activities in maturity levels, there is an implicit statement that some activities are “better” (or more mature) than others. However, a given organisation may have good reasons for not doing a certain activity, but this will not be reflected in a

study that blindly follows the BSIMM framework. A concrete example of this could be an organisation that develops and runs a service that runs in the cloud. In this case, activity SE2.4 "Use code signing" does not make sense, since the source or binaries are never transferred out of the organisation's cloud. Sometimes checklists have an option to specify "Not relevant" to a given question, and it could be worth considering adding this to the BSIMM yardstick as well. Looking at this from a different angle, maybe an organisation should first establish the set of software security activities that represent the "holy grail" for them, i.e., the 112 minus any activities deemed to be not relevant. The results should then be compared with this modified yardstick.

From a psychological point of view, it is tempting to ask if there is a threshold where a BSIMM score becomes de-motivating rather than inspiring. If an organisation is "flatlining" in almost every practice, management might not even want to tell the employees. This is troublesome on many levels, not least if it leads to the assessment report to be filed and promptly forgotten. If we accept that the BSIMM activities represent "good software security practice", organisations should most likely strive to implement more activities; simply ignoring the immaturity problem does not make it go away.

1.5 Conclusion and Further Work

The BSIMM Software Security Framework represents a comprehensive list of good practice software security activities which is a good foundation to build a software security program in a development organisation. It may not be possible to replicate the BSIMM study method as it is done by Cigital, but even a questionnaire-based approach can produce useful results when studying software security practices in the real world.

Acknowledgment

(To be added after review)

References

The published version has additional references

[1] Jaatun, M.G.: Hunting for Aardvarks: Can Software Security Be Measured? In Quirchmayr, G., Basl, J., You, I., Xu, L., Weippl, E., eds.: Multidisciplinary Research and Practice for Information Systems. Volume 7465 of Lecture Notes in Computer Science. Springer Berlin Heidelberg (2012) 85–92

[2] McGraw, G., Migues, S., West, J.: Building Security In Maturity Model (BSIMM 6) (2015) http://bsimm.com.

[3] Jaatun, M.G., Cruzes, D.S., Bernsmed, K., Tøndel, I.A., Røstad, L.: Software security maturity in public organisations. In Lopez, J., Mitchell, C.J., eds.: Information Security. Volume 9290 of Lecture Notes in Computer Science. Springer International Publishing (2015) 120–138

[4] Beznosov, K., Kruchten, P.: Towards agile security assurance. In: Proceedings of the 2004 New security paradigms workshop, ACM (2004) 47–54

[5] Fitzgerald, J.T., White, C.B., Gruppen, L.D.: A longitudinal study of self-assessment accuracy. Med Educ **37**(7) (Jul 2003) 645–649

[6] Rhee, H.S., Ryu, Y.U., Kim, C.T.: Unrealistic optimism on information security management. Computers & Security **31**(2) (2012) 221 – 232